

Word of Mouth: Transcribing an oral history collection

AUGUST 11, 2021

HEIDI MARGOLD AND ERIN ROSE

ILLINOIS HEARTLAND LIBRARY SYSTEM – CATALOGING MAINTENANCE CENTER

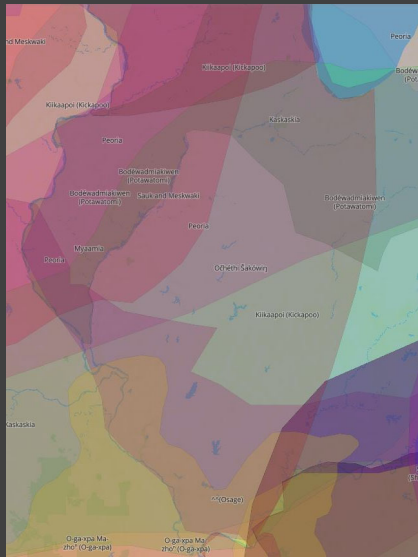
1

Take a couple of minutes to introduce ourselves here:

Emily will be introducing us with the biographies we provided.

Heidi: My name is Heidi Margold and I have been a cataloger with the Cataloging Maintenance Center (CMC) at Illinois Heartland Library System (IHLS) since March of 2020. Prior to being a cataloger with the CMC I attended University of Illinois in Urbana-Champaign to receive my Masters of Library and Information Sciences. During my studies I had my hands in many pots; I was a special collections digitization technician, I worked as a special collections conservation intern, I spent time working on a digitization and metadata project at a museum, and I also served as collections care coordinator after being the graduate assistant for that role.

Erin: My name is Erin Rose and I've been the Metadata Cataloger for the Cataloging Maintenance Center (CMC) at Illinois Heartland Library System since April of 2019. I have an MLIS from the University of Wisconsin – Madison and focused on Archives and Records Management for a Digital Age. I have worked as a digital production technician, a processing archivist, a project archivist, an archival intern (notice a theme here?), and a lot of that work revolved around gathering, organizing, and appending metadata to materials.



<https://native-land.ca/>

Land Acknowledgement

We would like to begin today by recognizing and acknowledging that we are on lands of Native peoples. These lands were the traditional territory of these Native Nations prior to their forced removal; these lands continue to carry the stories of these Nations and their struggles for survival and identity.

2

(Heidi)

We would like to start out by thanking you for attending our presentation today and begin by recognizing and acknowledging that we are on lands of Native peoples. These lands were the traditional territory of these Native Nations prior to their forced removal; these lands continue to carry the stories of these Nations and their struggles for survival and identity.

Moment of silence

Agenda

Brief History

Project Planning

Challenges

Process

Q&A

3

(Heidi)

During the course of the presentation, we will be covering our project from beginning to end as best as we can. We will start with a brief history, discuss project planning, executing our plan, challenges we faced, process, and then a short Q&A at the end. If you have any questions, please post them in the Q&A area along with the slide number and we will do our best to answer them at the end of our presentation. I will now turn it over to Erin who will start by giving you a history of the Marshall Public Library Oral History project.

Friends of Marshall Public Library Oral History Project

Friends of MPL – started to collect interviews in 1994
Volunteers collected interviews

- Volunteers from the Friends of Marshall Public Library group
- Smithsonian exhibits
- Local history projects
- Marshall High School students (1999, 2011)
- Dan Crews (World War II)

Librarian converted tapes to digital and edited digitally captured interviews

Requested transcription and upload to IDA from the CMC

4

(Erin)

The Friends of Marshall Public Library started the project by gathering oral history interviews from local residents as far back as 1994. The project grew to include interviews gathered in conjunction with Smithsonian exhibits, local history projects, 4-H projects, Marshall High School students, and a series of interviews conducted by Dan Crews for the Illinois State Library.

Jamie Poorman, Marshall Public Library's head librarian, converted tapes to digital files and edited both those and digitally-recorded interviews and put them onto Marshall Public Library's YouTube channel

(https://www.youtube.com/channel/UCTcb5In_PdeTau_wJEG6FPQ). It was decided that the interviews should live on the Illinois Digital Archives (IDA) website to enable further accessibility, and transcripts were required. Alyson Thompson, director of the Marshall Public Library, and Jamie contacted the Cataloging Maintenance Center (CMC) to get the process started, and an agreement was made that the CMC would transcribe and upload the materials for Marshall Public Library.

Overall, there were approximately 256 interviews at a total of approximately 187 hours.

Project Planning

Guidelines

Tools

Organization

Transcription Process

Research

IDA

5

(Erin)

I (Erin) had not really worked with oral histories in the past, so this was a fairly new experience for me, and I built the project from the ground up. I needed to complete research on how to transcribe audio files, and quickly realized that I needed to compile some internal standards to follow to ensure consistency. I created a guide with standards, including spellings on different words like geez versus jeez. I found some rules for transcription from the University of Rochester (<https://www.cs.rochester.edu/research/cisd/resources/nonwords.html>) that I felt aligned well with what I had been thinking, so I used their rules as a template for my own, and then embellished or expanded as needed.

I also discovered that file-naming wasn't always consistent (there had been volunteers working on the project too), and there were duplicate interviews, along with audio files that were in two parts – or a set of the two separate files, along with a third file that had been stitched together. I had to comb through to organize them a little better, but even then with the large quantity of interviews and the fact that I'd have to listen to each of them to make sure they matched, it was an ongoing process until we completed the transcription component.

I had done a small bit of research to see what programs or potential outsource options were available, and settled on using ExpressScribe because I had used it once or twice before and found it to be user-friendly and very low-cost. We'll go into more detail on these in later slides.

There was a lot of research to complete, much more than I had anticipated. Different personal and business name spellings were common topics, and without the local contextual knowledge, it took me longer than it would have for somebody who may live in Marshall.

Lastly, I have never worked in IDA before, and only briefly poked at CONTENTdm in graduate school. It's been an ongoing learning curve.

Pricing

Temi: \$.25/minute; \$2,805

Rev: \$1.25/minute; \$14,025

GMR: Varies

- Good audio up to 2 speakers: \$1.25/minute; \$14,025
- Good audio 3+ speakers: \$1.50/minute; \$16,830
- Difficult audio up to 2 speakers: \$1.75/minute; \$19,635
- Difficult audio 3+ speakers: \$2.00/minute; \$22,440

Does not include research of names, creating standards, studying for subject headings, etc.

Good Audio



- Little background noise
- Clear speaker(s)
- Minimal accents

Result: 90-95% Accuracy

[See example →](#)

Difficult Audio



- Heavy background noise
- Crosstalk
- Strong accents

Result: Mostly Unusable

[See example →](#)

Erin – Before really getting into the transcription element, I looked at how much it would cost to have the transcription outsourced. I asked a few former colleagues and classmates what they used and did some research of my own, and Rev, Temi, and GMR were the top three I considered. In the end, I decided that transcribing it ourselves, especially since there wasn't a short turnaround, would be the best route to take. There was a lot of difficult audio, accents, and I would still have to go through each interview to determine subject headings, if parts needed to be cut, and verifying that the spellings of different personal, business, and geographical names were accurate.

Guidelines

Reference other oral history project guidelines

- University of Rochester,
<https://www.cs.rochester.edu/research/cisd/resources/nonwords.html>

Personally-created rules

- Forms of words (geez vs. jeez, etc.)
- Spell out abbreviations
- Noises ([clear-throat], [phone rings], etc.)
- Square brackets instead of angle brackets

File-naming conventions = Sanity

7

(Erin)

As I had mentioned, I discovered almost right away that I would need to create standards to maintain consistency across the transcriptions.

I relied heavily on the University of Rochester's transcription rules (<https://www.cs.rochester.edu/research/cisd/resources/nonwords.html>) because it was the clearest to me and it aligned with different rules that I learned over time. I liked that they spelled numbers out. I made a few exceptions for coding purposes; for instance, they have a category of "other phenomena" which includes mumble, laughter, and more, and they put them in angle brackets. However, I wasn't sure what sort of code I may encounter with uploads and chose to err on the side of caution by using square brackets.

I created a file-naming scheme to be followed in order to keep everything organized; unfortunately, files came to me without any consistent naming conventions, there were duplicates, and more, so I really didn't have a clear idea of what I was working with. Renaming the files as I went along made it easier to track and organize materials.

For the sake of time, I won't go over every detail of how I chose to transcribe, but did include a copy of the guidelines that I created for this project as a handout that anybody is welcome to reference or even use. Sharing is caring!

Transcription Tools

Express Scribe - transcription

MS Word/Excel - transcribing/tracking

OneDrive - share materials with MPL

Adobe Bridge/Photoshop

VLC Media Player - to convert files

Headphones

8

(Heidi)

Erin and I used a variety of tools to successfully transcribe the interviews and I will detail those here.

Our bread and butter was Express Scribe which is software we used to play back our transcription audio. This is a very powerful transcription tool because you can set up custom hot keys on your keyboard to rewind, pause, play, and fast forward the audio file. Express Scribe also has an auto transcribe function, but we found that it didn't work well.

Microsoft Word is the word processing software that we use to type up the transcription. Express Scribe did have a transcription notepad, but we chose to use Microsoft Word since that tool offers grammar and misspellings highlights and it was much more efficient to file share.

Excel is how we tracked a ton of different information which included project statistics, progress, hours transcribed and hours to transcribe, notable business names and resident names.

OneDrive is a cloud storage software that we used to share all of the interview files including the audio files, transcription files, and eventually images of the interviewees

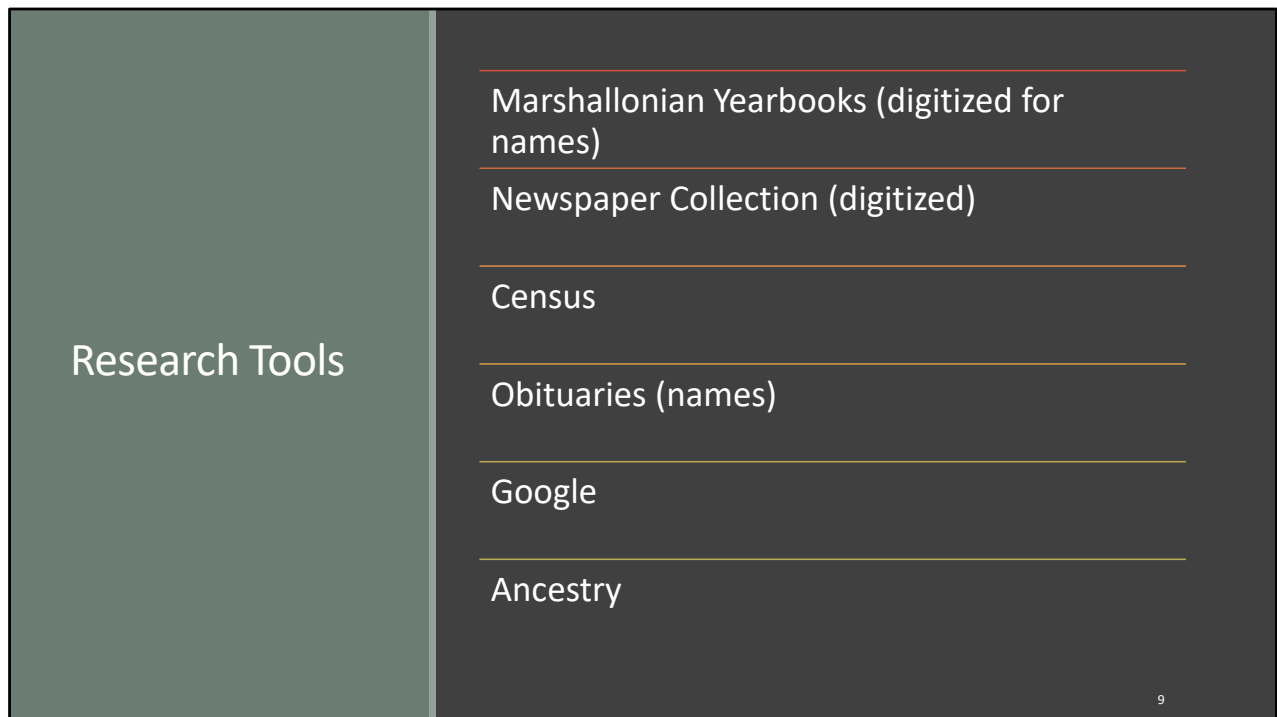
when applicable. This service was critical for us to easily share hundreds of folders in an efficient and safe way.

Adobe Bridge is a digital content organization software. Our primary use of Adobe Bridge is to apply metadata to each image individually. While not a necessary step it is a very helpful step to embed metadata into images for any future users who might download the image files.

Adobe Photoshop was used for extremely minor adjustments to some of the image files. Erin and I both have a background in digitizing special collections and it is very important to both of us to maintain the original integrity of photos. I want to make it very clear that any image modification that happened was minor and to enhance the viewers ability to see and understand the images.

VLC Media Player is a tool we use to convert mp3 files to mp4 files. This is a necessary step because the Illinois Digital Archive (aka IDA) requires mp4 files when uploading. There will be more on IDA later in this presentation.

Also, headphones are really important because they can be the difference between understanding audio and not understanding audio sometimes. We use gigantic, over-ear headphones and that was a game-changer.



(Heidi)

Now that we have covered some of the transcription tools, I'd like to cover some of the research tools. As mentioned previously we did all of our own research on names, locations, and events.

Marshall already has a digital archive that has their local newspapers and yearbooks, links to all of the resources we discuss will be found at the end of the presentation. We used these resources heavily when researching names, businesses, and notable events to confirm spelling and details of stories. All of these resources are text searchable so it was easy to narrow our search down by an approximate date and then text search for the name/location/event.

When we couldn't find what we were looking for by browsing the yearbook and newspaper collection we would utilize census documents, obituaries, and searching on Google. These resources proved to be very time consuming to sift through and sometimes wouldn't lead to anything relevant. I wish there was a software that had census documents, obituaries, newspapers, and all sorts of wonderful genealogical data... Oh wait there is, Ancestry!

One of my favorite resources was Ancestry library edition. I won't get too far into the nitty

gritty of Ancestry but this ended up being a critical tool for both Erin and I to research names, places, events, family trees, etc. Because of COVID 19, Ancestry library edition was opened for patron to use outside of the library space which made it possible for us to use this tool while working remotely.

Organization of Materials

Learning about the collection and organizing all the interviews/materials

Tracking through Excel documentation

Processing and preparing for transcription

10

(Heidi)

Erin received the transcript and audio files from Marshall Public Library. Early on Erin needed to tackle organizing the files which was challenging because of inconsistent file naming. Excel was the program chosen to do this. She made an Excel workbook with several tabs to organize information. Most of the interviewees had signed an authorization form, but some had not. She split all the interviews between those two main groups. There were also other oral history collection projects, that Erin mentioned earlier, Marshall was involved with that we made tabs for also. Once all the interviews were accounted for, we were able to identify any duplicates or missing files from our inventory. Our Excel document contained a lot of very important information besides the name of our interviewee we were able to record the interviewer, the number of minutes the interview was, date of interview, and additional content like topics so we could generate subject headings later. It also helped us keep track of what interviews were done, what were in progress, and what interviews we still had to transcribe. The Excel document was probably the brains of our project.

Transcription Process

YouTube

- Gather transcriptions to edit (CC/Transcription, Copy/Paste to Word)

Express Scribe

- Hot keys, playback speed, rewind/stop/play/fast forward; interfaces with MS Word

MS Word templates

- Consistency in header

11

(Heidi)

Note: Strive for perfection, settle for excellence.

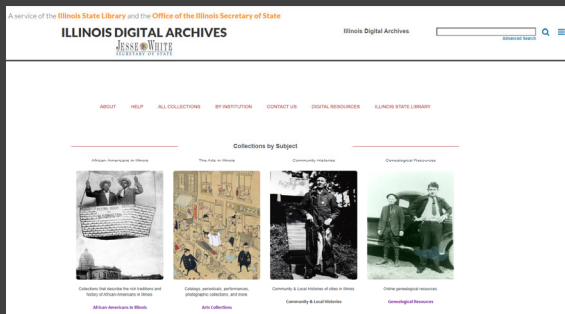
Marshall Public Library had uploaded most of the interviews to YouTube for storage and organization purposes and they had leveraged their volunteers to help with the transcription process before Erin and I became involved in the project. So, here is what our workflow looked like:

-If the interview had already been transcribed by a volunteer we would open the transcription file, load the interview recording into Express Scribe, and then play back the audio making corrections and adjustments as needed.

-In most cases, the interview was not transcribed so we would first visit YouTube to see if the interview had been uploaded. YouTube sometimes will generate captions for videos and sometimes not, it's not very clear on why they do or don't for some video files. We would grab the YouTube captions when they were available, copy and paste that into our word document, and then make corrections and adjustments as needed.

-Sometimes the transcription would not be available through YouTube, so we were transcribing from scratch:

-- We would create a new word document with the appropriate header and file name.
--We would play back the interview using Express Scribe and transcribe while also making important notes in our excel file to track subjects/topics and length of interview.
Research was always involved in every interview



Illinois Digital Archives (IDA)

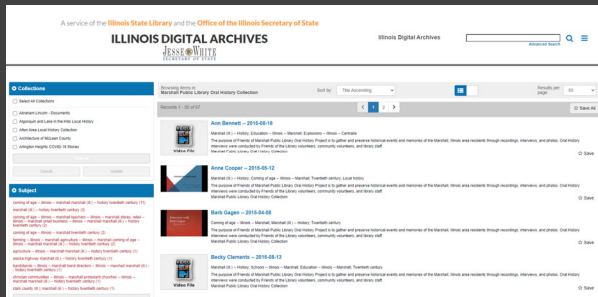
- Meetings
 - MPL: Alyson Thompson and Jamie Poorman
 - Illinois State Library: Sandra Fritz and Andrew Bullen
- Templates
- File-naming
- Subject headings - Local and/or LCSH
- Photos - Photoshop/Bridge
 - jpeg vs. jpg
- Upload process

12

Erin

Prior to actually uploading to IDA, we had to take steps to prepare. Those steps included meeting with Alyson and Jamie to lay out their vision for the site, our Illinois State Library/Illinois Digital Archives contacts, Sandra Fritz and Andrew Bullen, to make sure that we could do what Alyson and Jamie wanted, and then meeting again with Alyson and Jamie to discuss the process.

We had to create metadata templates for each interview, finish up any file-naming inconsistencies, create and apply subject headings, edit photos and append metadata to them, and go over the process to actually upload these materials to IDA.



Illinois Digital Archives (IDA)

13

Erin

This is what the site looked like as we finally started to get interviews and their related materials uploaded.

Preservation & You!

Archival Master Image	Production Master Image	Access Image
<ul style="list-style-type: none"> • Represents as closely as possible the information contained in the original • Uncompressed, or lossless compression • Unedited and “use-neutral” • Serves as long term, sustainable resource • Can serve as surrogate for the original • Highest quality that can be produced by the digitizing organization • Large file size • Stored in the TIFF file format 	<ul style="list-style-type: none"> • Produced from archival master • Uncompressed, or lossless compression • May be edited for technical corrections to the file, such as adjusting tone or stitching together multiple images of a large item (e.g. maps) into one file • Used to create derivative files for access • Large file size • Stored in the TIFF file format 	<ul style="list-style-type: none"> • Used in place of master image for general web access • Generally fits within viewing area of average monitor • Reasonable file size for fast download time; does not require a fast network connection • Acceptable quality for general research • Compressed for speed of access • Usually stored in JPEG or JPEG2000 file format

Heidi

Preservation and You! I am going to give you brief guidelines on file preservation. Always start out with a master file of high quality, these typically are quite large. For example, if your project involves scanning images, I advise scanning the images using a TIFF file format. This file will be your master file and you will try and interact with it as little as possible. Plan for the worst but expect the best – store your master files in three locations such as an external hard drive, cloud storage, and on a server so you have a back up WHEN a file gets corrupted not IF. Once your master files are stored appropriately you can then duplicate them to create Production Master Images. Production Master Images are digital surrogates of your master images. You will work with production master images as if they were the original files. With these files you can create Access Image files which are typically smaller file types like JPEG. Because they are smaller file types, they are easier to share and take up much less space than a large TIFF file.

Challenges We Faced

Audio Files
Accents, vocal issues
Jargon/Vocabulary
Personal/business name spellings
Sensitive content/Trigger warnings
Addressing difficult topics, subject matter

- Difficult interviews content-wise
- Mental wear and tear on transcribers

15

(Heidi)

A handful of the audio files were of poor quality, so we couldn't hear them well. We also had issues understanding some accents clearly, and some folks had harder to understand voices, whether they were gravelly or very soft-spoken. We had issues with jargon because while we're both from the Midwest, not all the regional jargon is clear to us.

One of the struggles that both Erin and I faced was understanding/spelling/recording local names of people/places/events. We were lucky enough to find a resource of for many historic landmarks in Marshall which really did help us with the spelling. Once we confirmed the spelling was correct we tracked it in our excel sheet. The librarians at Marshall Public Library were both extremely knowledgeable and assisted us with the spelling of hard to understand or unique names of individuals and businesses.

Because we were making decisions on behalf of another library, we did face some challenges in our decision making when it came to inserting content warnings for triggering content and outdated terminology. We both found it challenging to bring up difficult topics

with sensitivity for members of their community. In our opinion, it made sense to add content warnings but as discussed before this isn't our collection and we wanted to leave the final decision up to MPL. What we decided was to bring the instances to the librarians attention – they have decided to trust our opinions and those warnings have been added.

Another challenge we faced was in some cases the interviewee would say “can you cut that out” or “leave this next part out” etc. Jamie did a lot of audio editing and clipping of things like this, but as a human being, she did miss a few. When we came across interviews like this, we would flag these interviews and Erin used the software Audacity to clip and trim them herself.

More Challenges

Planning process for transcription of project – This was a new project type for both of us

Finding the right tools – Erin had transcribed one or two interviews for a previous job in Express Scribe

16

Erin

The planning process, which can be fun at first, was my very first one in regards to oral history transcriptions. I started out by contacting some people I knew had worked with oral histories, but one had used rev.com (which I'd deemed too pricey) and the other hadn't actually transcribed, but had digitized taped oral histories.

In a previous job, I had transcribed two short interviews using Express Scribe and knew that it interfaced well with Microsoft Word, so I decided to go that route.

Then I had to plan out what to do with the rest of the project, which is when I decided to create the guidelines.

More Challenges

Creating guidelines/workflow

- research, decisions
- adhering to IDA requirements
- Consistency:
 - File-naming - some work done by volunteers resulting variances in title formats/etc.
 - Duplicate audio files/entries
 - Extra audio files
 - Missing audio files
- Difficulties in working with an outside organization
 - Some aspects may make sense to one org, not the other
 - Transcriber may not understand the importance of some subjects/topics to mention in the subject headings
 - Some things that are common knowledge in Marshall are not common knowledge to Erin/Heidi

17

Erin

There were a lot of decisions to make when creating the workflows and guidelines, and the addition of IDA's requirements could either help or hinder that process, depending on if what we wanted to do was possible.

There were consistency issues in file-naming due to multiple people working on the project before it got to us, there were duplicate or missing audio files in some cases. These consistency issues were another reason it was important to establish guidelines and workflows.

There are also difficulties in working with an organization that is physically separated from yours. Heidi and I created the majority of the transcripts, and cleaned up those done by volunteers, but since we weren't from Marshall we did not have a lot of the local knowledge required for name spellings of people and businesses, and we didn't always know what was important to Marshall's history, so deciding on important subject headings could be a little more complicated. We'd had Jamie, the head librarian, editing the transcripts for us before we considered them finalized, but when summer hits they are extremely busy with summer reading programs. Luckily, Heidi and I had started a file of names and how they are spelled (verified through Ancestry, obituaries, etc.) so that we

could edit as we go along and take some burden off of Jamie. I sent the file to her for review, she said that we had everything down correctly, so I finalized the edits as the project wrapped up.

More Challenges

Editing process

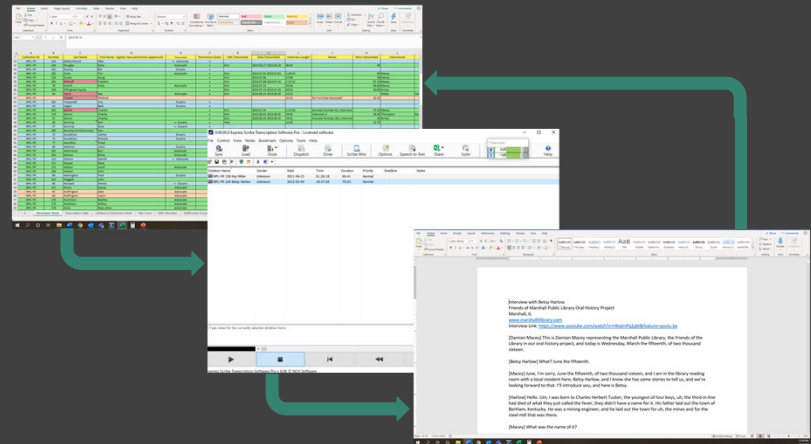
- Timelines/current projects differ, meaning one of us may be able to devote a lot of time to the project whereas the others may not at the time, and vice versa

18

Erin

Being separate from the organization that we were doing the project for was also difficult as it also applied to questions and meetings – it's hard to get people from multiple locations to a single meeting time, so there was a lot of self-reliance and meetings were infrequent. It was definitely not an unwillingness to meet or chat, but more a difficulty with coordinating times, since none of us could really pop into our colleagues' offices to see if they were free.

The Process



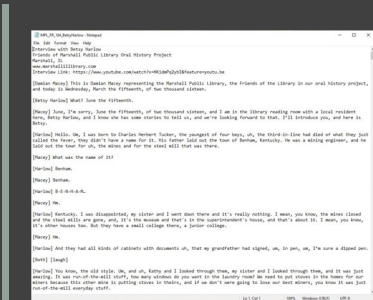
19

(Erin)

Now we're going to fly through a step-by-step of our process. I do want to note that I have created a guide for each process, and it's attached in our handouts on the presentation page. If there are questions about the processes, you are free to ask them after the presentation or contact Heidi or myself later on.

The top left Excel file is a list of the interviews that Heidi and I organized over time. You will probably notice that there is a lot of color-coding; I would highlight interviews in different colors for different stages, such as not started, in progress, and completed, I had a column to note the date or dates that it was started and completed, and I color-coded cells based on if I could get a transcript off of YouTube, if there was anything special about the audio file, if it had been transcribed by a volunteer and only needed to be edited by us, etc. I would also apply different colors in cells to note anything of importance for that interview, such as a content warning note, audio issues, etc.

Once I noted which interview I was going to work on in the Excel file, I would start working on transcribing it in Express Scribe and Microsoft Word. Once I was finished with the transcription, I went back and noted it in that master Excel file for tracking to ensure that we didn't have any duplicate transcriptions, and then moved on to the next one.

[illegible]

The Process

Once the interview was transcribed, it was time to start work on filling in the Metadata template and gathering any appropriate related materials including a .txt file of the transcription and any photos, etc.

Field Number	Description of data to be included
1	Title
2	Creator
3	Library of Congress Subject Headings
4	Local Subject Headings
5	Description
6	Date Original
7	Identifier
8	Coverage - Geographic
9	Coverage - Temporal
10	Type
11	Collection Publisher
12	Rights Management
13	Collection Name
14	Contributing Institution
15	Full Text Transcription
16	File Size
17	Bitrate/Frame
18	Language
19	Contributor
20	Content Warning
21	Digital Format

The Process - Metadata Templates

21

The metadata template fields include fields required by IDA (which is based on Simple Dublin Core) and then the addition of any fields that we thought were pertinent. I again applied color coding to the cells here because a lot of the fields remain the same across the board, such as Description, the Rights Statement, the Collection Name, etc., and some change, including the Title, Creator, Subject Headings, etc. This helped me remember which I needed to look at and change as I edited the templates and then again later for the upload process.

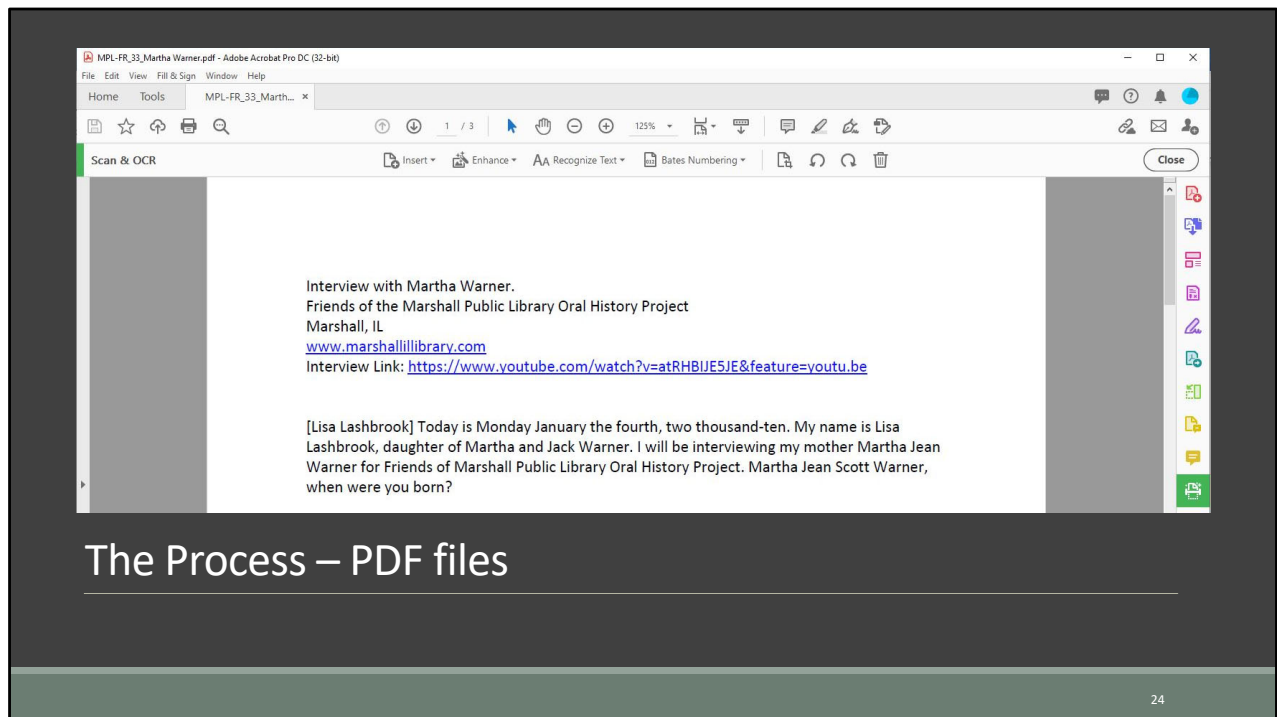
We used Library of Congress Subject Headings for the majority of our headings and were sure to subdivide those headings geographically when possible and appropriate. We also have a field for “Local Subject Headings” since there are so many local groups, organizations, and businesses that simply don’t show up in LCSH, and we also spelled out full names (so for women we would include maiden names when we could find them, and for names like “Jim” we would spell it out to “James” to increase findability).

The Process - Filetypes

22

For each interview, we would upload a .txt file of the transcript, .jpg files if there were accompanying photos, an OCR-ed .pdf file of the interview transcript, and an .mp4 of the interview.

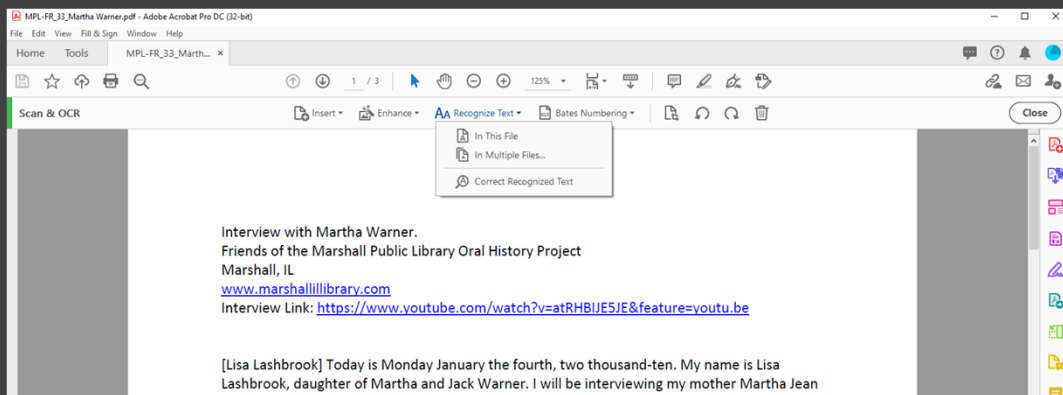
PDF Files



I create the PDF files from Word documents by just saving the Word document as a PDF (File > Save as Adobe PDF)

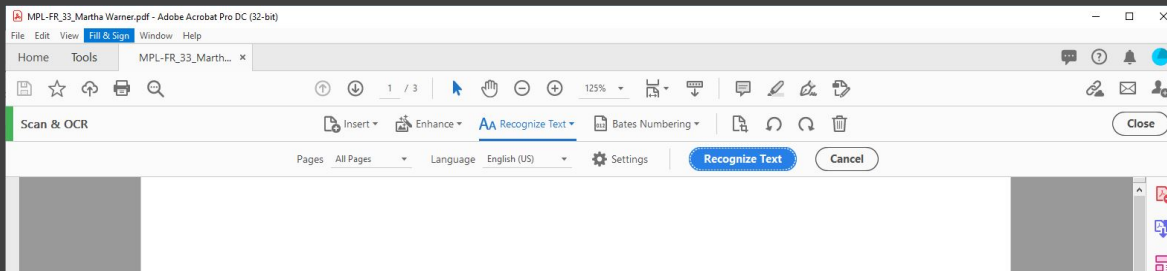
I use Adobe Acrobat Pro DC for most of my work because it has little features that I really like, including the ability to add OCR, or “Optical character recognition” to my documents. This makes the text searchable. Thankfully, since these are derived from Word documents, Adobe Acrobat Pro DC has no issue applying this, and I have a sneaking suspicion that it does this automatically when I save my Word documents into the program. However, just to be safe, the process I use is the following:

With the file open in Acrobat Pro DC, there are two ways to get to the OCR option. You can either click on the “Tools” tab at the top of the page or click on the green printer icon on the right. Both will get you to the OCR option.



The Process – PDF Files

A new toolbar will appear above your document, and you'll click on "Recognize Text" and choose "In this file."



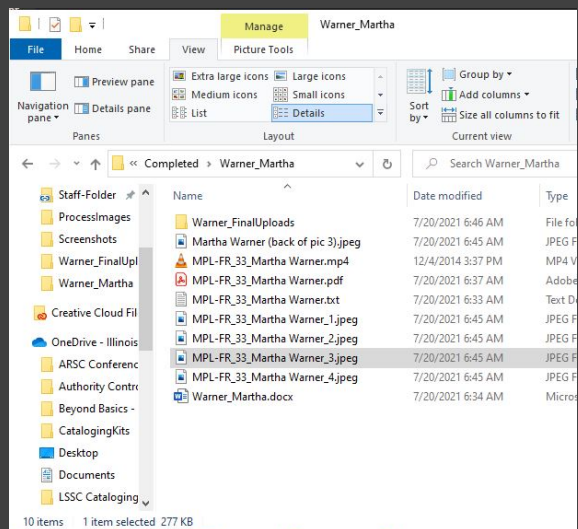
The Process – PDF Files

26

Yet another new toolbar will appear, and you will click on the blue “Recognize Text” button.

The program will run through the process, you’ll see a progress bar, and once that is done the OCR is applied. Remember to save the document, and then you’ll be ready to go!

JPG Files

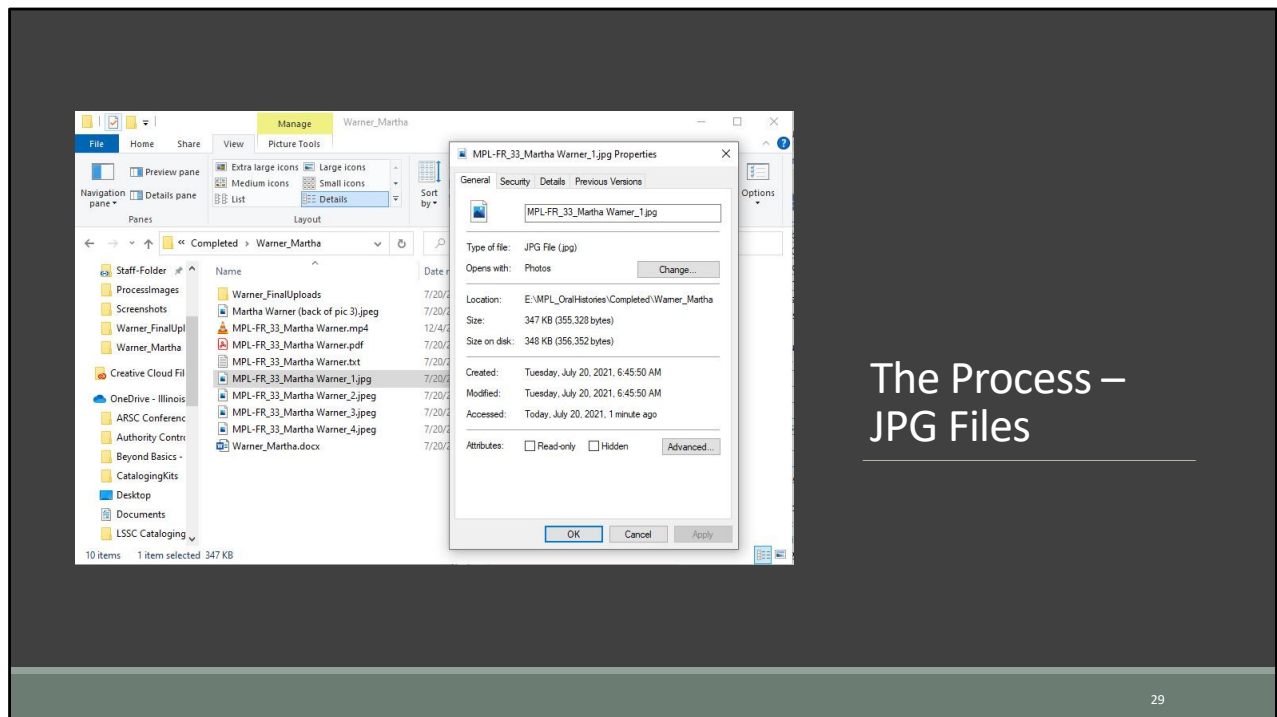


The Process – JPG Files

Before I realized that the file extension for JPGs made a difference, I had them all saved as JPEGs (the four-letter extension instead of the three-letter one). I asked OCLC about this, and I was told that “It's because Project client is a Windows app and used to only support 3-character file extensions.”

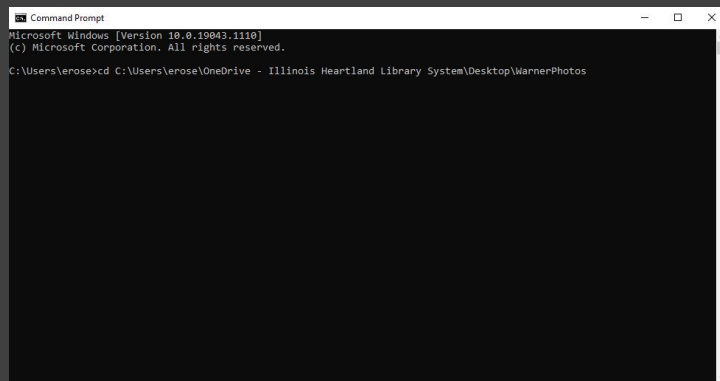
Because I had a LOT of photo files to edit, I was able to find a script to run that would change all of the extensions. I do want to note that I had to put the photos in a folder on my desktop for the script to work, but it works! The place that I got the script from is linked in the guidelines document:

[https://fileinfo.com/help/windows batch rename file extensions](https://fileinfo.com/help/windows_batch_rename_file_extensions)



The Process – JPG Files

I know how to change the extension by hand for each individual photo, but that would have been too time-consuming.

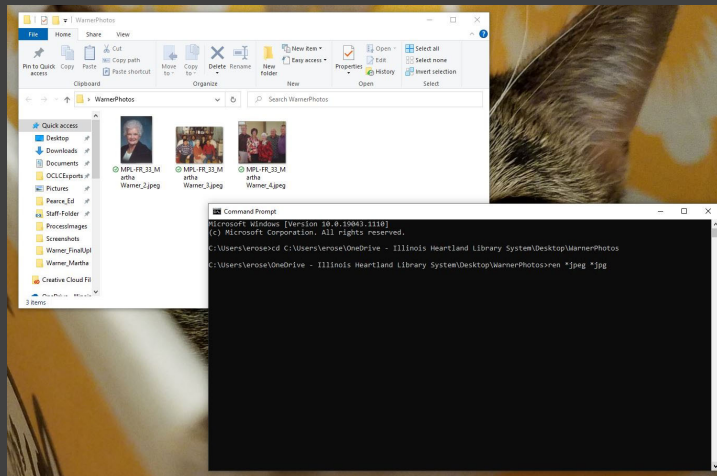


```
Command Prompt
Microsoft Windows [Version 10.0.19043.1110]
(c) Microsoft Corporation. All rights reserved.
C:\Users\erose>cd C:\Users\erose\OneDrive - Illinois Heartland Library System\Desktop\WarnerPhotos
```

The Process – JPG Files

30

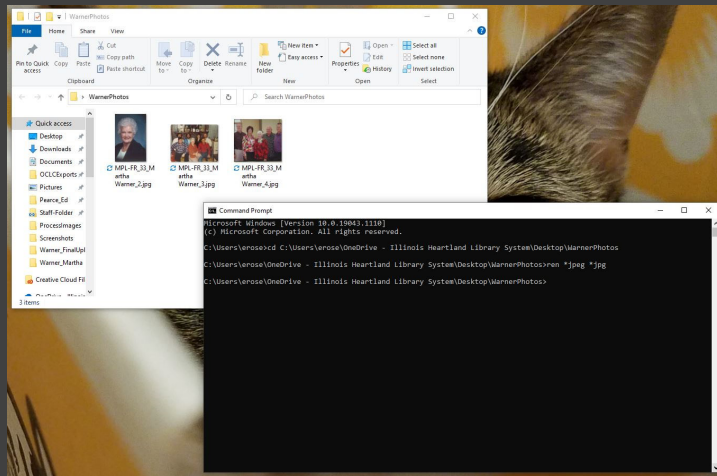
The script runs through Command Line, and I would enter CD, a space, and copy and paste the filepath directory into this part. This command points to the folder that we want to process the changes in.



The Process – JPG Files

31

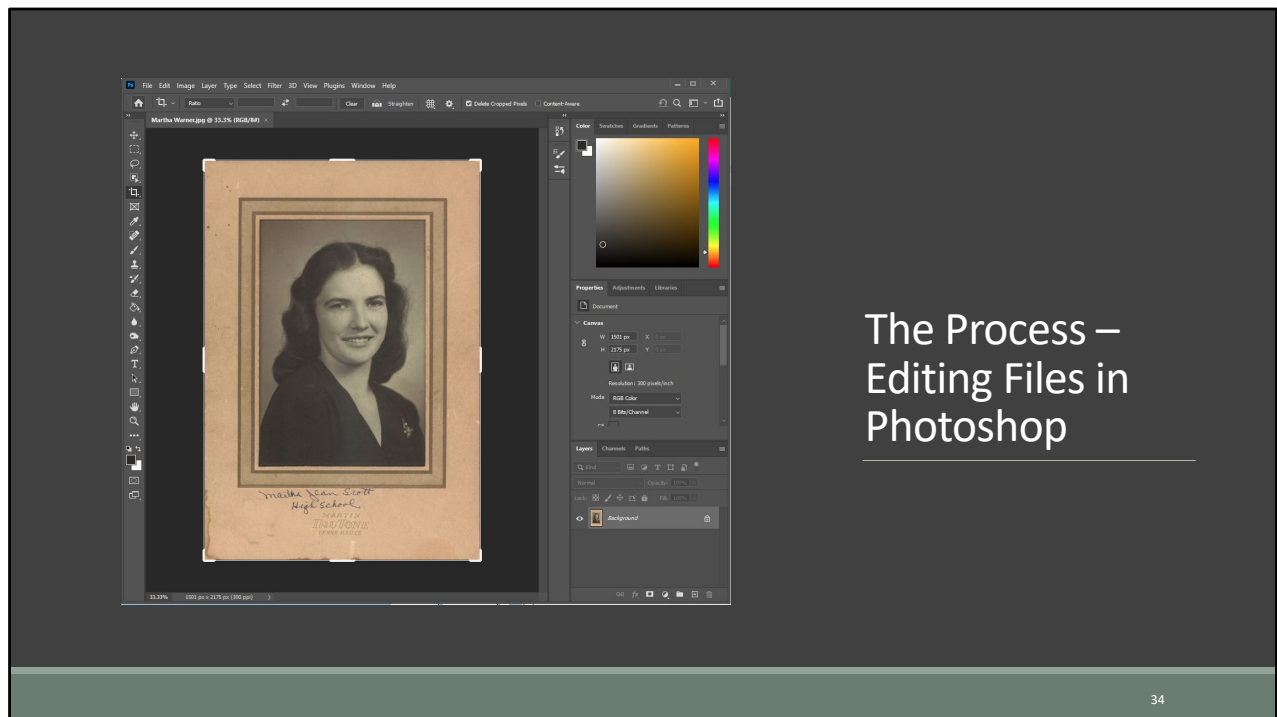
Next, I typed in the second line of code required and hit enter...



The Process – JPG Files

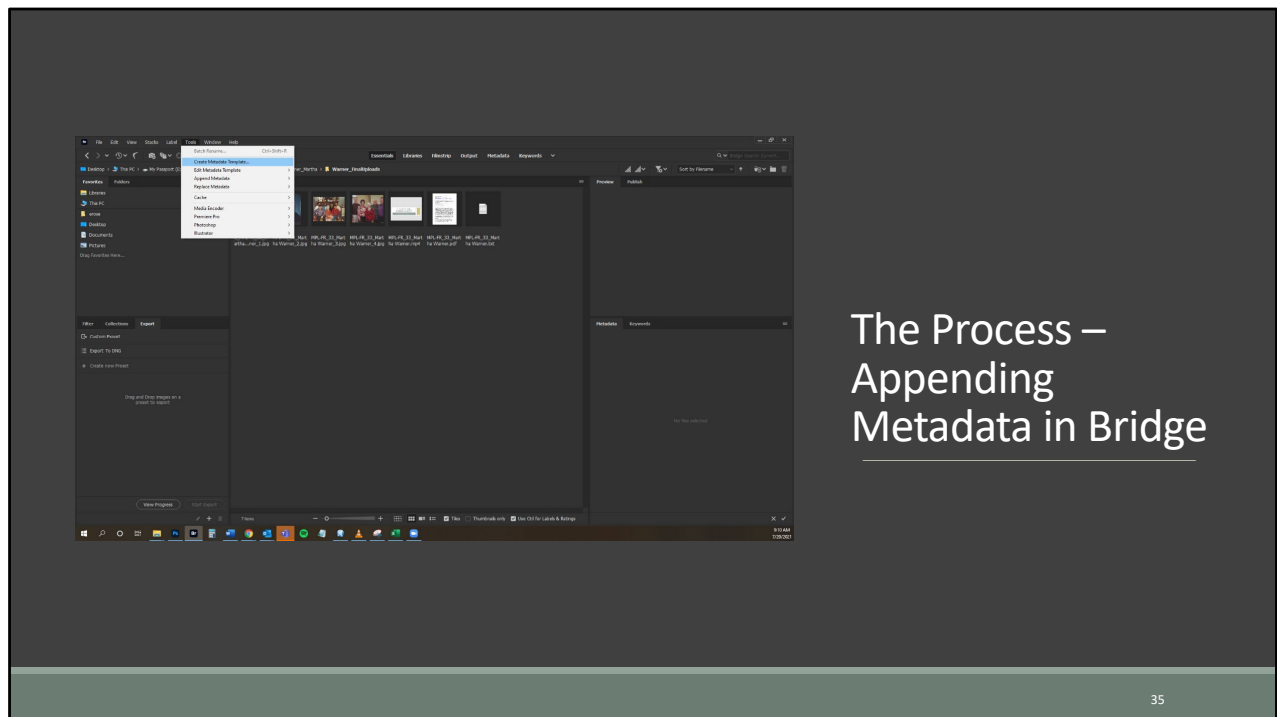
And once I ran that script, you can see that all of my photos are now the 3-letter extension for JPGs. Hooray!

Adobe Photoshop and Bridge



The Process – Editing Files in Photoshop

I use Adobe Photoshop to crop the photos because it's just what I'm comfortable with; for instance, the majority of these photos were digitized on a flatbed scanner and had the white background that was just unnecessary to keep. However, with this particular example, while I wanted to crop out that white background, I wanted to keep the full item because it has important information (if you look closely, you can see "Martha Jean Scott High School" written below the photograph). If you have a preferred program to use for this, feel free to use it. Just make sure that you save the file as a .jpg after you complete edits.



The Process – Appending Metadata in Bridge

You can create or import metadata templates into Bridge. For this particular project, I created a Template that has metadata that I want to append to all photos across the board. You can do this by going into Adobe Bridge, navigating to the Tools dropdown, and clicking on “Create Metadata Template.”

✕

Edit Metadata Template

✕

Template Name:

Choose the metadata to include in the template

☐ Use Unseen

☐ Intellectual Genre

☐ IPTC Scene Code

☐ Sublocation

☒ City

☒ State/Province

☒ Country

☐ ISO Country Code

☒ Title

☐ Job Identifier

☐ Instructions

☐ Credit Line

☐ Source

☐ Copyright Notice

☐ Copyright Status

☒ Rights Usage Terms

Marshall

Illinois

United States

Marshall Public Library Oral History Collection

Unknown

☐ IPTC Extension

☐ Person Shown

ⓘ

Only checked properties will be added/changed to this template.

Clear All Values

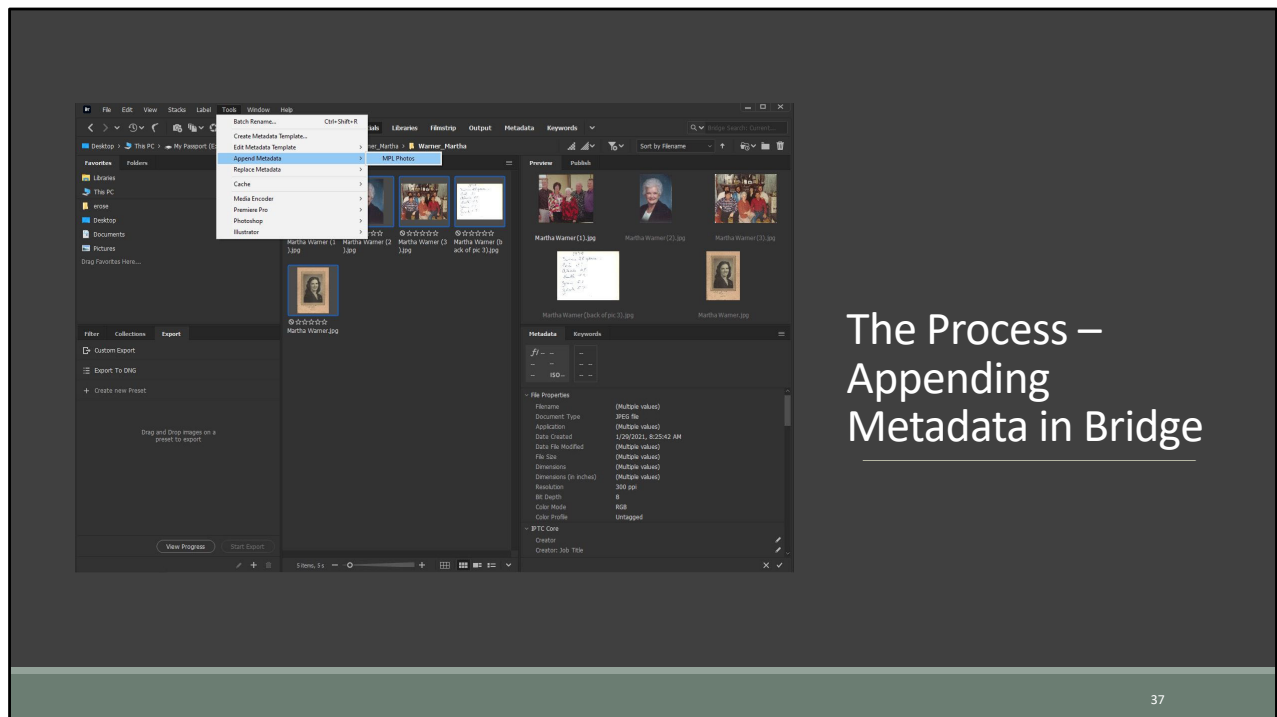
Properties selected: 5

Cancel

Save

The Process –
Appending
Metadata in Bridge

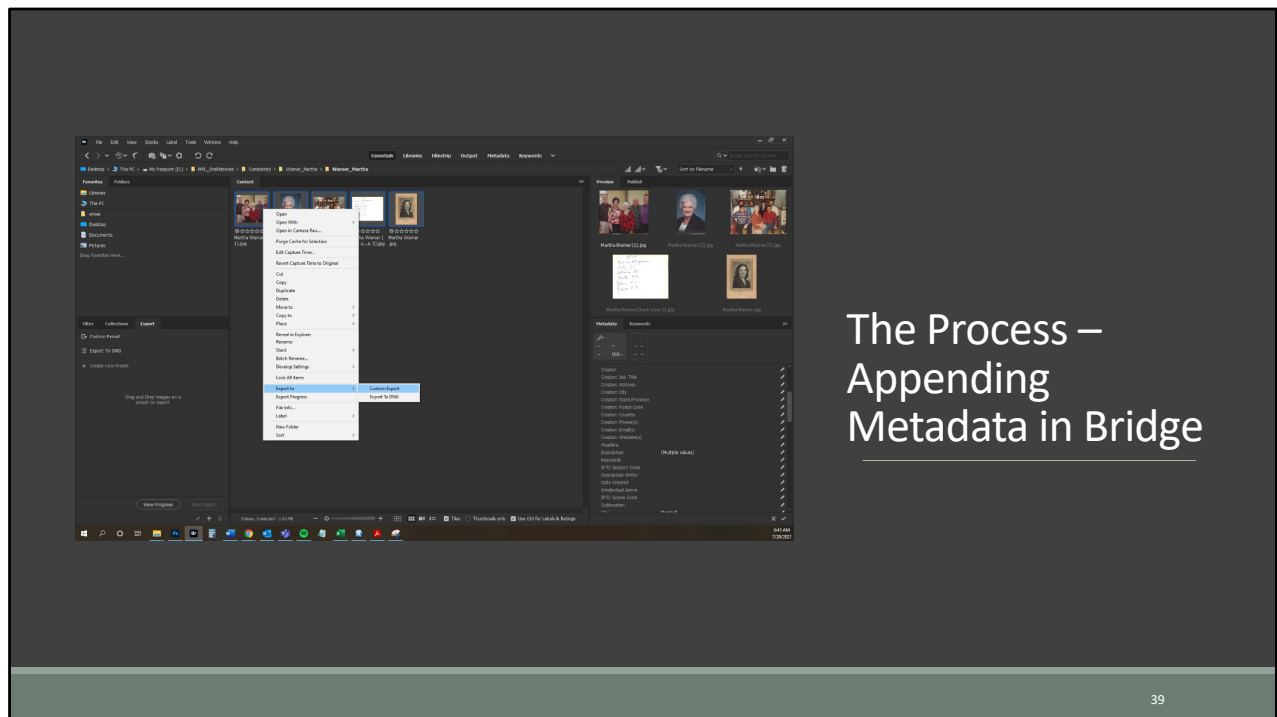
Next, type in the data that you want to be in the template and when you're done click "Save."



The Process – Appending Metadata in Bridge

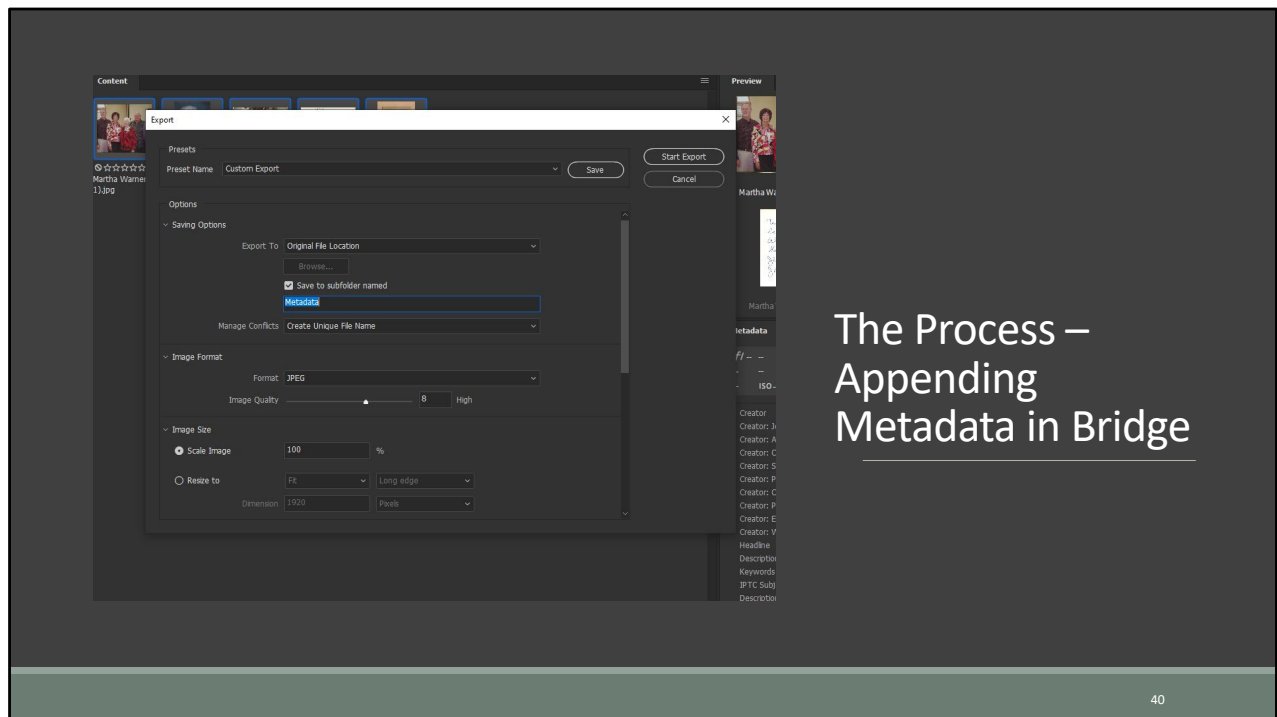
Now, navigate to your photos folder, select the photos that you want to append the template data to, then navigate to the Tools dropdown, click on Append Metadata, and choose the template that you want to use.





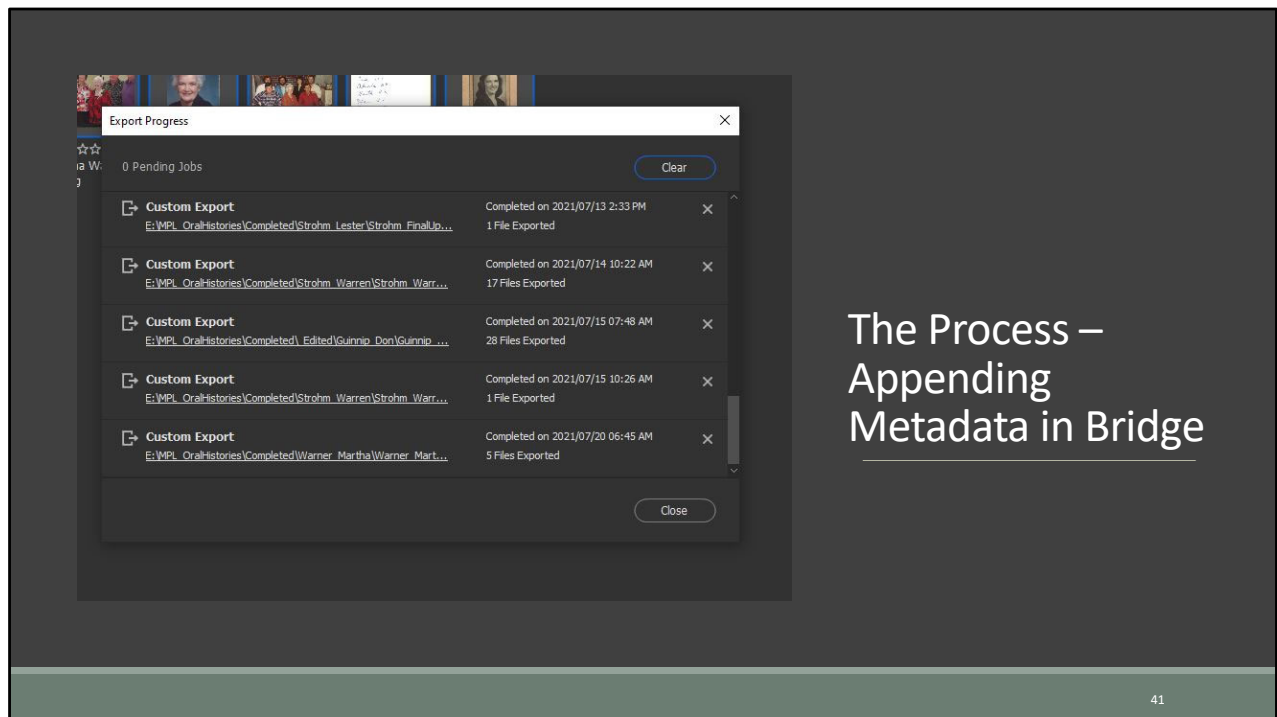
The Process – Appending Metadata in Bridge

Once you have appended all of your metadata to all of your photos, you will want to “Export” the photos with their new metadata. Do this by selecting all of the photos, right click to bring up the dropdown, click on “Export to” and choose “Custom Export.”



The Process – Appending Metadata in Bridge

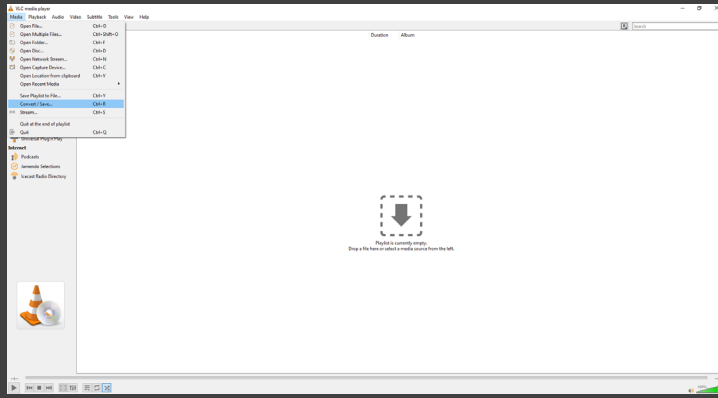
I always click into “Save to subfolder named” and create a title (I use metadata). I’ve tried to replace the original photos (the ones that don’t have metadata), but this doesn’t always work and instead all of the photos get mingled together. Exporting this way puts them into a folder nested within the folder that holds the originals. Once you’ve done this, you can click on “Start Export.”



This window will pop up to show you the progress (it goes fast, if you have a small number of photos, you might miss it!), and once it's Completed, you can close and you're done with Bridge.

At this point I go to File Explorer where I delete the photos that don't have metadata, then move the photos *with* the metadata from the Metadata folder into the old ones' place.

Audio Files

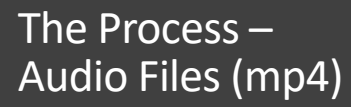


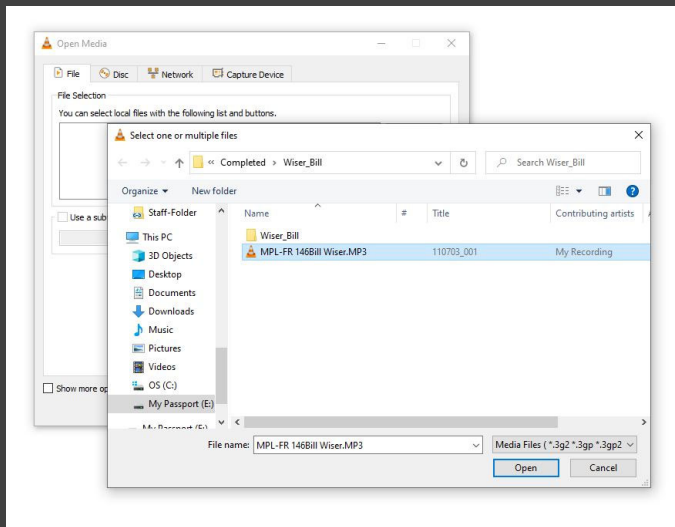
The Process – Audio Files (mp4)

43

I use VLC Media player to convert mp3 to mp4.

I start by navigating to and clicking “Convert/Save” under the Media dropdown.

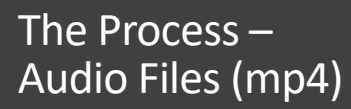


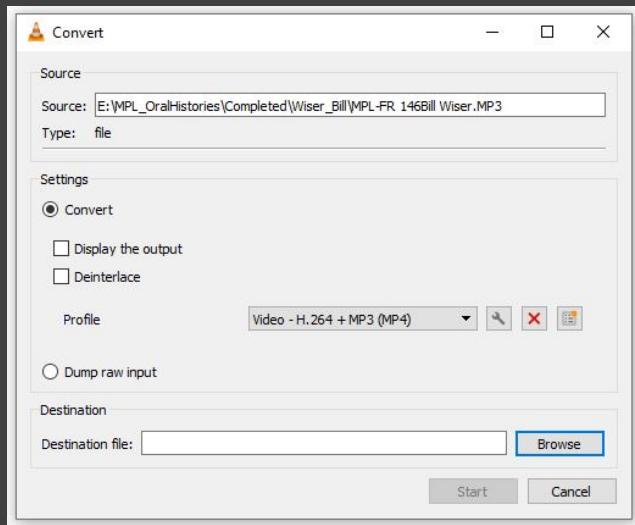


The Process – Audio Files (mp4)

45

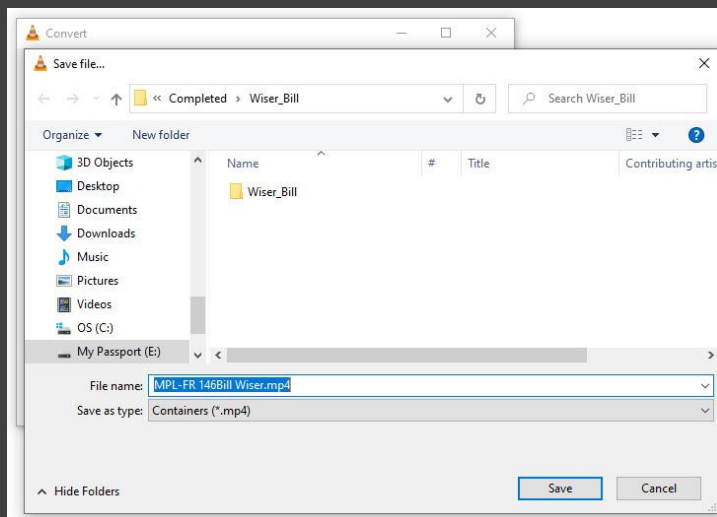
Select the file...





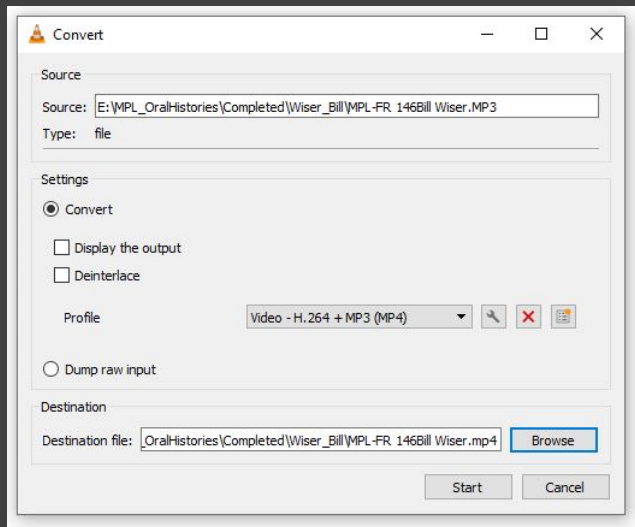
The Process – Audio Files (mp4)

A new window will pop up, and under profile you will want to choose “Video - H.264 + MP3 (MP4)” and then in the Destination section click “Browse” to choose where the file will go...



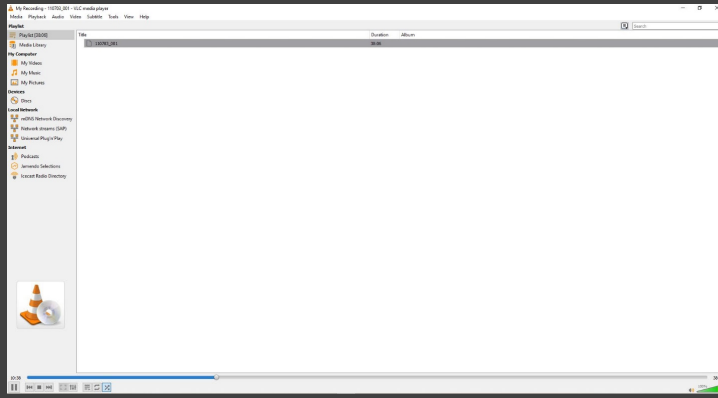
The Process – Audio Files (mp4)

Navigate to where you want to save the file, and under “File name:” you will change the extension of .MP3 to .mp4 (upper/lower case don’t matter, I just prefer lower case), and then click “Save”...



The Process – Audio Files (mp4)

Now you can click “Start” and the window will automatically close and the process will begin...



The process –
Audio files (mp4)

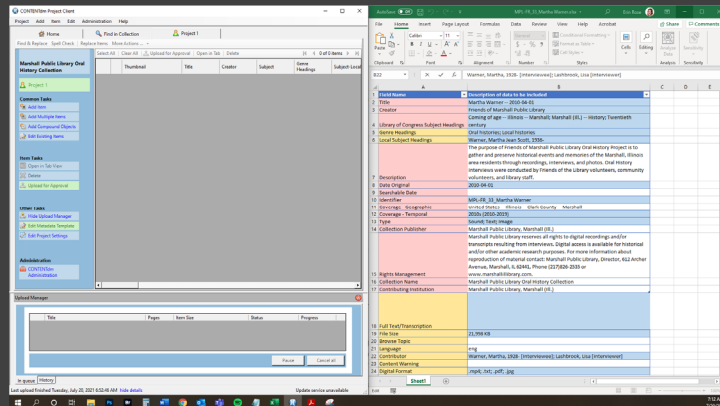
50

You can see that it's working by looking at the bottom of VLC Media Player where you will see the progress bar (blue) moving. Once it moves across the bottom all the way to the end of the bar, that means it will be done, and you will find the .mp4 file along with the .mp3 file in your folder.

Illinois Digital Archives - CONTENTdm

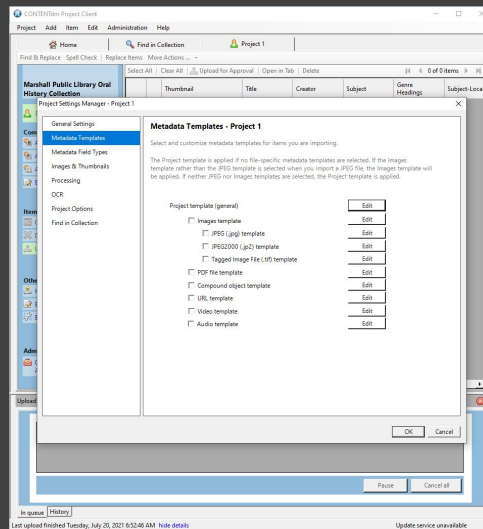
51

I put all of the items that I wanted to upload for a compound object into a single folder titled with the last name of the interviewee, an underscore, and FinalUploads at the end so that I had a clean filepath to just those files. The reasoning for this will make sense shortly...



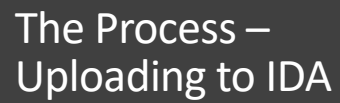
The Process – Uploading to IDA

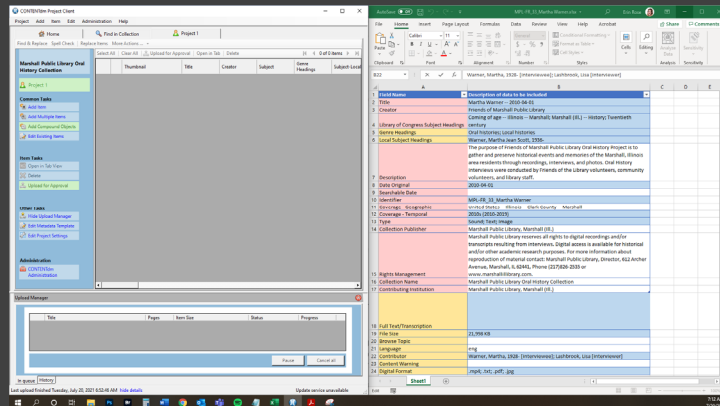
I open the Excel file template at the same time as the CONTENTdm one to copy and paste information over easily. To open the template in CONTENTdm, click on “Edit Metadata Template” on the left toolbar.



The Process – Uploading to IDA

Since we had Illinois Digital Archives create our project template custom to this project, I can click on the Edit button next to “Project template (general).”

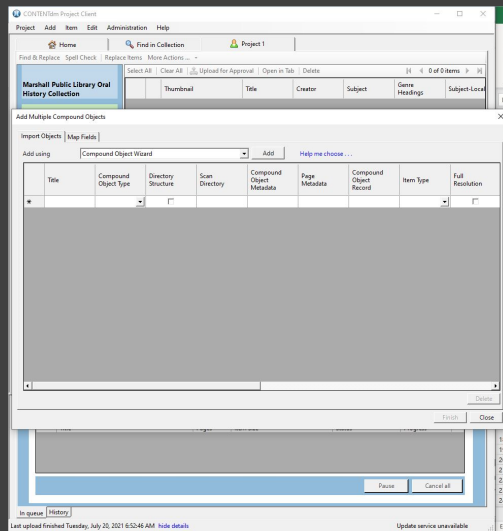




The Process – Uploading to IDA

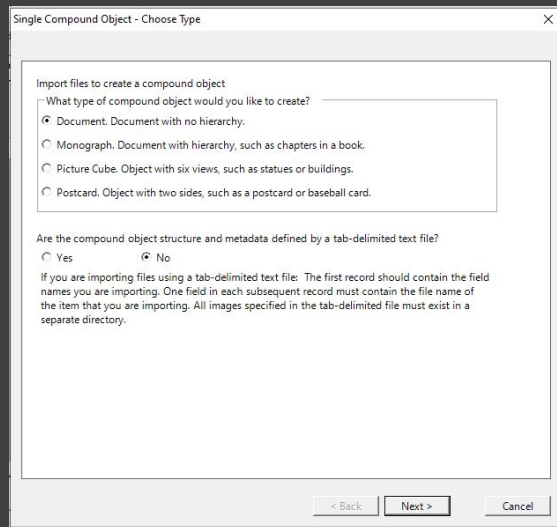
Next you will begin the Compound Object upload process. I am fairly confident there are ways to batch load this, but I was not taught them, so I did each individually.

To add the files as a compound object, click on “Add Compound Objects” on the left toolbar.



The Process – Uploading to IDA

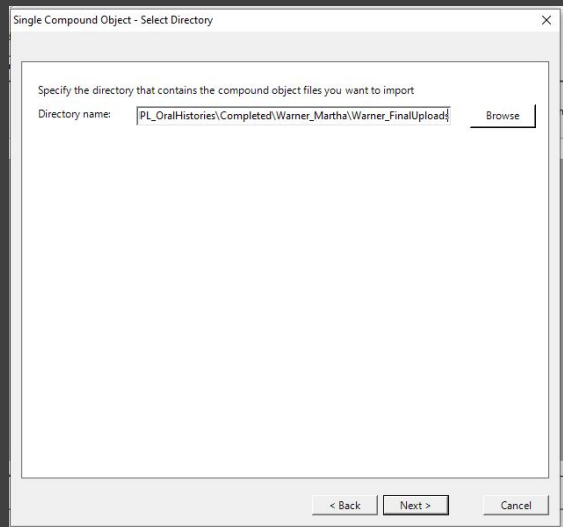
This window will open, click on the “Add” button...



The Process – Uploading to IDA

57

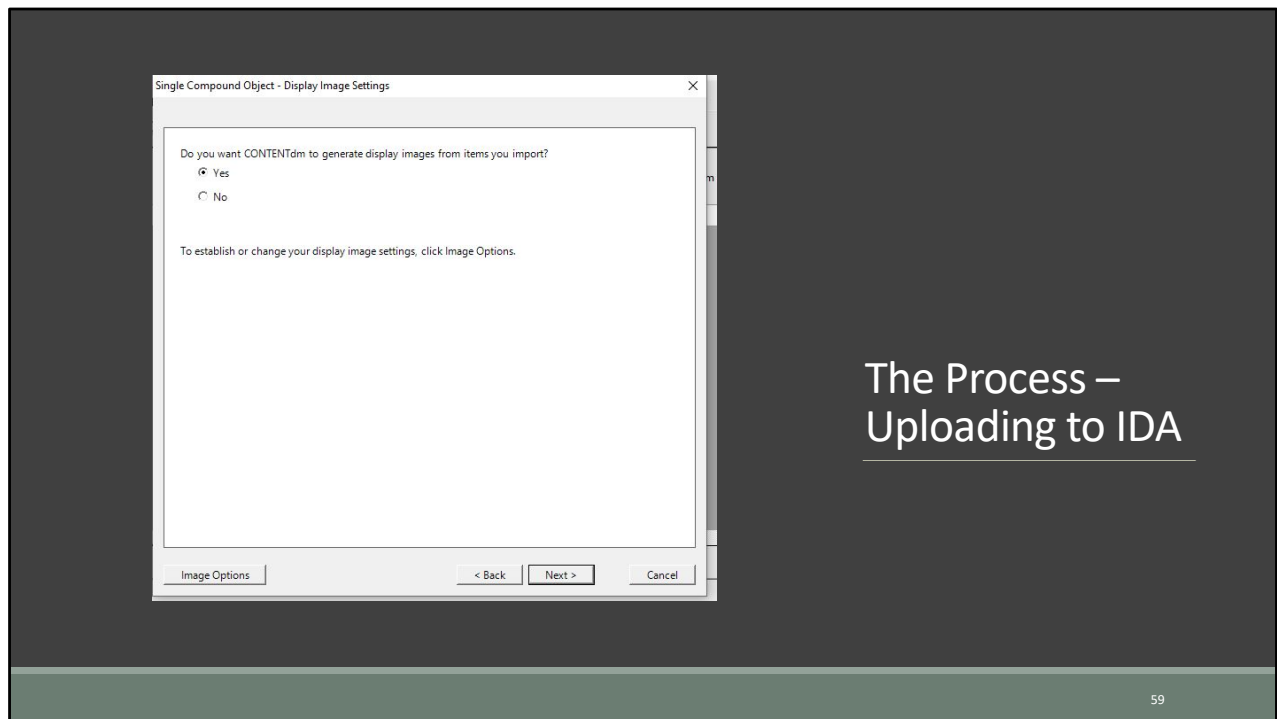
Make sure that “Document.” is chosen...



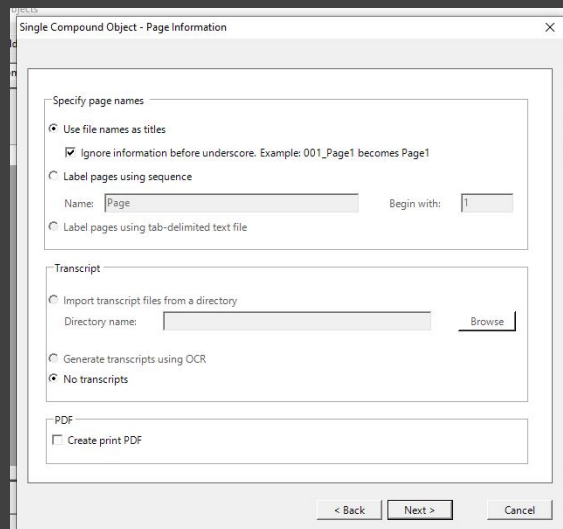
The Process – Uploading to IDA

58

Enter the filepath to what you want uploaded (it should be your lastname_FinalUploads folder); you can do this by either copying and pasting the filepath into the field, or by navigating to it through the “Browse” button...

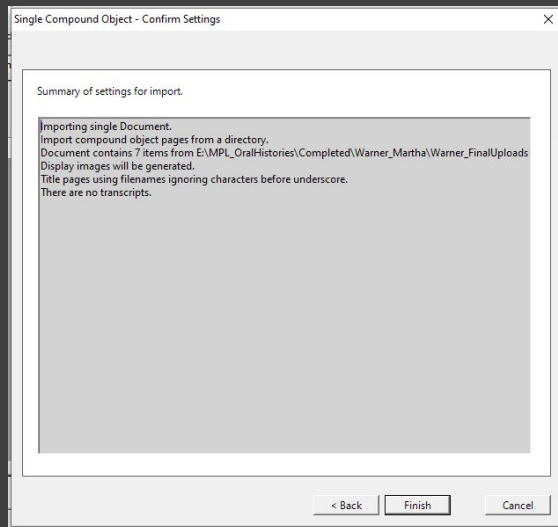


Leave “Yes” selected...



The Process – Uploading to IDA

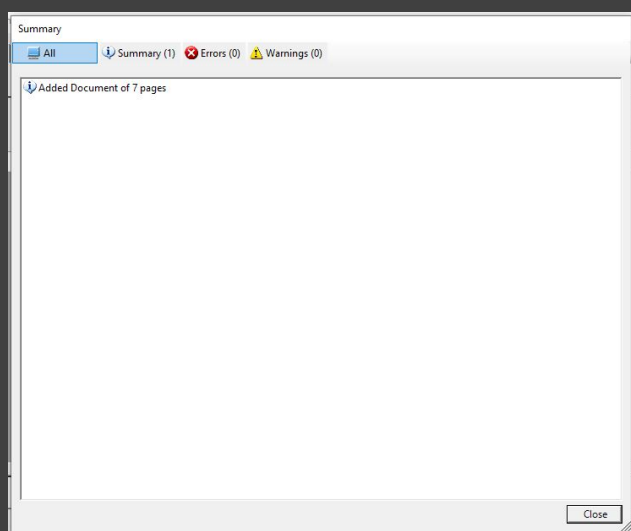
Leave “Use file names as titles” selected and click
“Next”...



The Process – Uploading to IDA

61

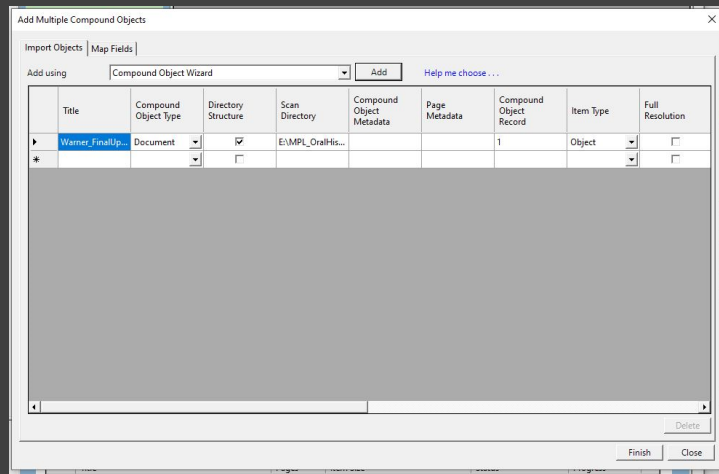
You will have a summary here, click the “Finish” button...



The Process – Uploading to IDA

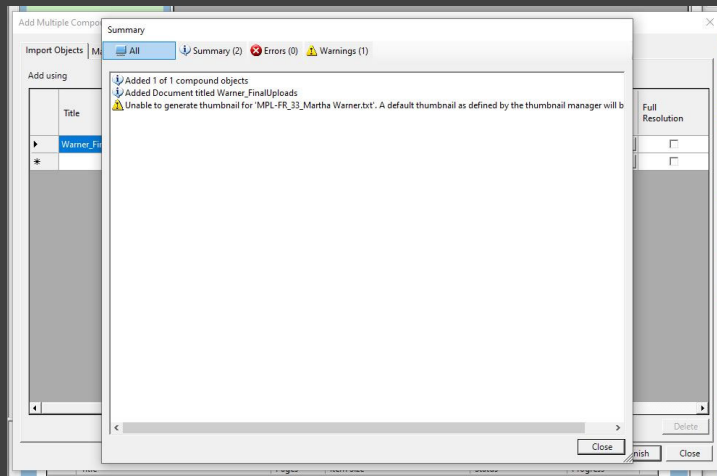
62

This window will pop up showing how many items have been uploaded for your Compound Object, it should match the number of files in the folder you uploaded. Click “Close”...



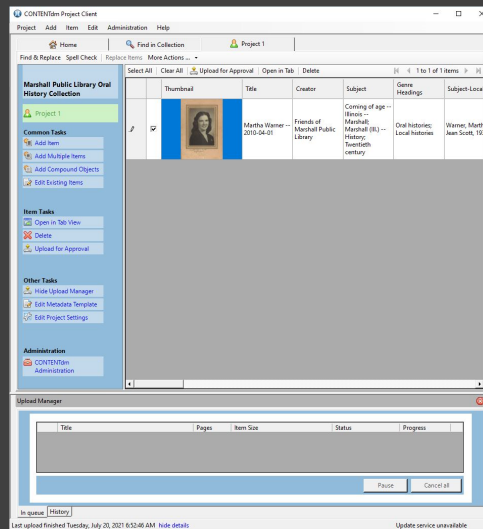
The Process – Uploading to IDA

This window will show up, you can scroll across the fields if you want with the bottom bar, and once satisfied, click the “Finish” button.



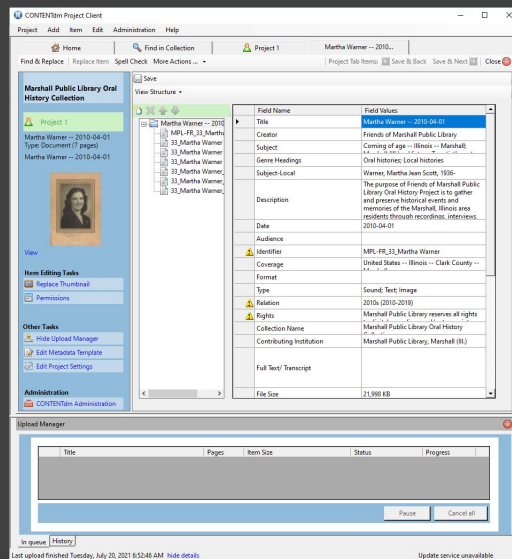
The Process – Uploading to IDA

You might get warnings - this should be okay, typically it refers to issues with generating thumbnails for files that aren't image files. If everything looks okay, click the "Close" button.



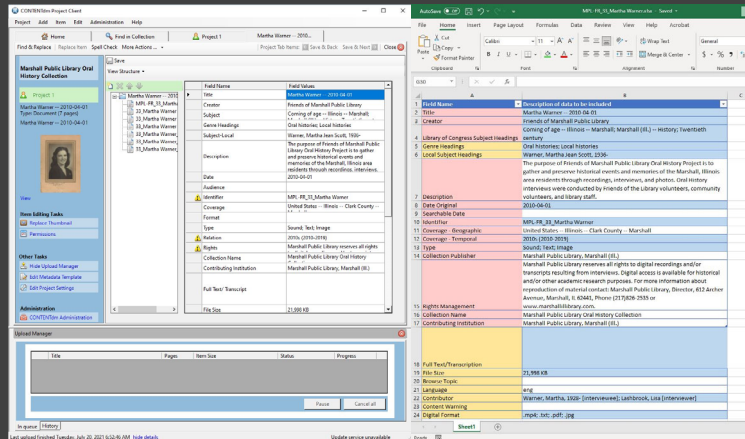
The Process – Uploading to IDA

You will see your compound object show up, it is recommended that you verify the information in your metadata template still looks accurate and that all of the items within your object look correct. You can open the template panel by double clicking somewhere on the compound object (usually I just double click on the Thumbnail).



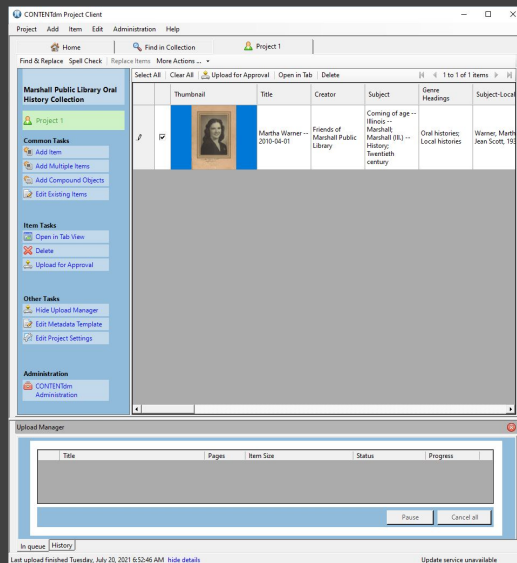
The Process – Uploading to IDA

If there are edits to be made, you'll have to edit the information for each item, as seen in the hierarchical "structure" list.



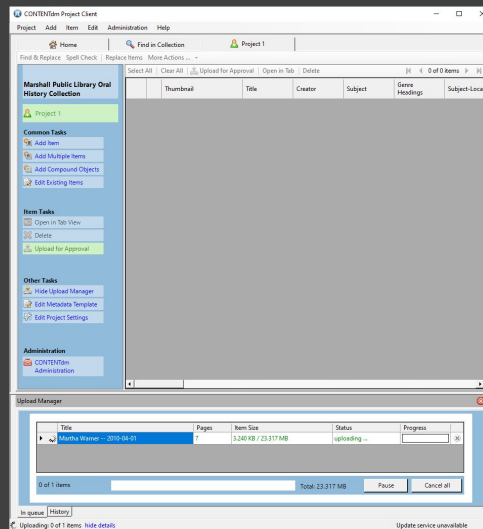
The Process – Uploading to IDA

Compare this metadata to the Excel file. If everything is fine, you can click on “Close” at the top right.



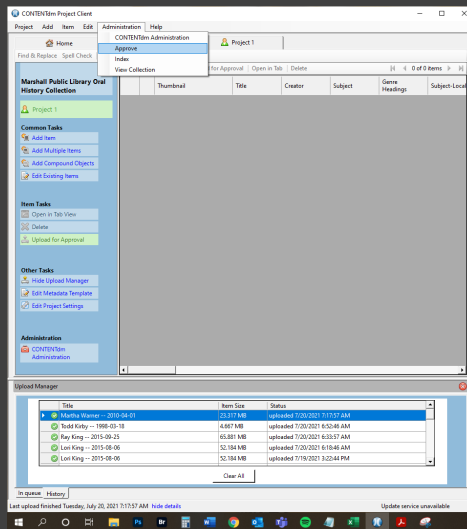
The Process – Uploading to IDA

Once verified, make sure that the checkbox next to your compound object is checked, and then click on “Upload for Approval” on the left toolbar.



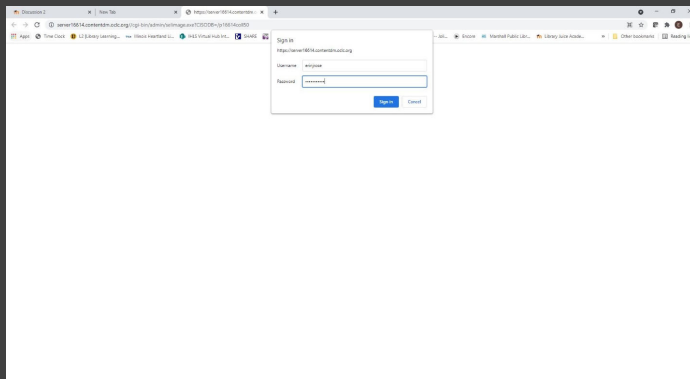
The Process – Uploading to IDA

While it uploads, you can view the progress by clicking on “view details” at the very bottom of the screen and then selecting the “In queue” tab.



The Process – Uploading to IDA

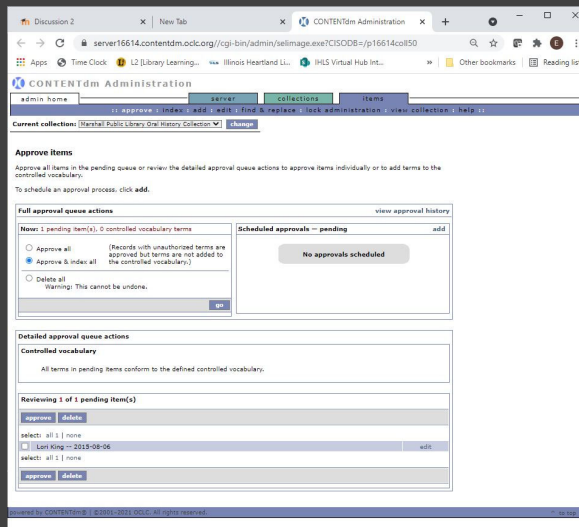
Once it's done uploading from the Project Client desktop program, it will show up in the "History" tab with a green checkmark. Next, you will navigate to the web interface. Do this by clicking on "Administration" dropdown and clicking on "Approve." This will automatically open the web interface...



The Process – Uploading to IDA

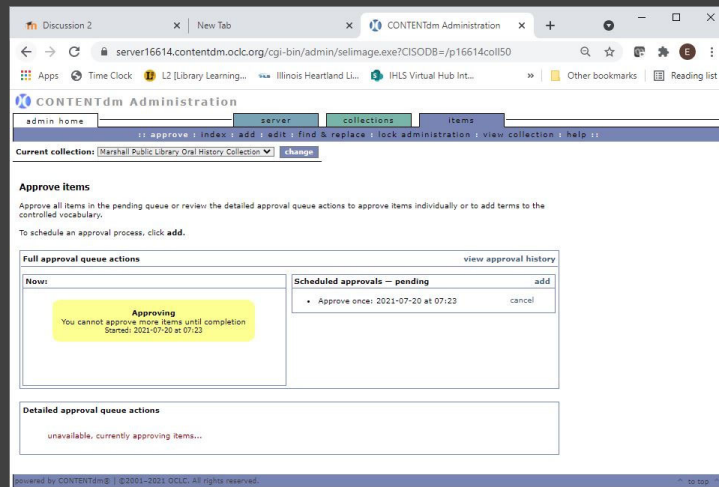
71

You will be asked to log in with your CONTENTdm
username and password...



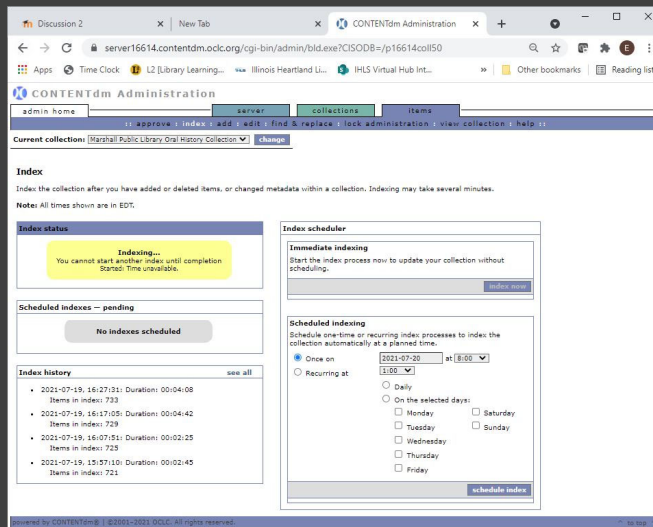
The Process – Uploading to IDA

Navigate to the purple “Items” tab. Click on the “approve” option in the purple toolbar, and you will see that there is an item pending for approval. Click on the radio button next to “Approve & index all” then click on the purple “go” button.



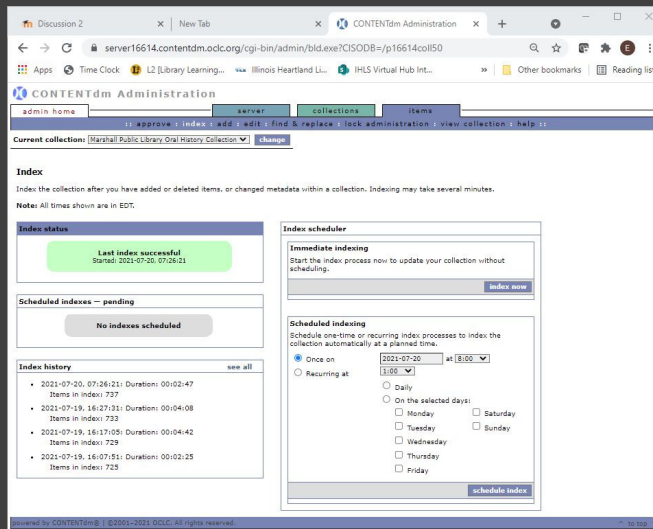
The Process – Uploading to IDA

You will see the yellow “Approving” rectangle, which is the first step of the Approve and index process. Once the Approval process is complete, since you chose “Approve and index” it will automatically move into the “Index” process. You can track the Index progress by clicking on “index” in the purple toolbar.



The Process – Uploading to IDA

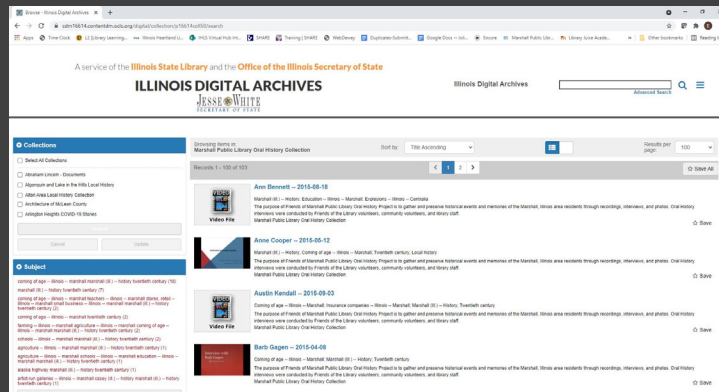
The box will be yellow and say “Indexing...” until the index process is complete...



The Process – Uploading to IDA

Once the index is complete, you will see a green rectangle stating that index was successful. You can now view the compound object on the website by clicking on “view collection” and it will open your collection page.

Final product

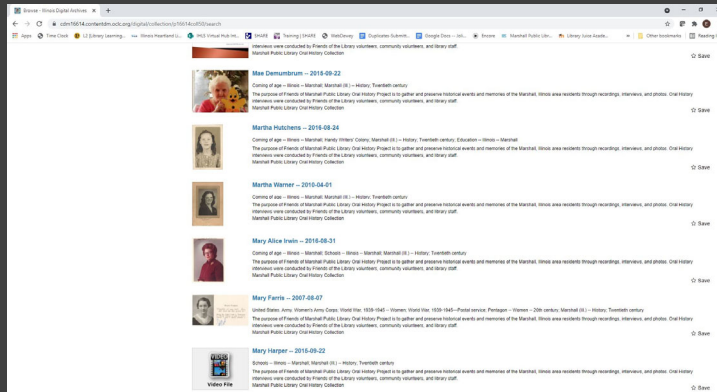


Final Product

77

Erin

Once you are on your collection page, navigate through to find your compound object...



Final Product

78

Erin

And click on it to open it up and make sure that everything looks right. You've done it!

Now we will pop out of the presentation and I will take you to the live page to see the object.

<https://cdm16614.contentdm.oclc.org/digital/collection/p16614coll50/id/752/rec/2>

Resources

Marshall Public Library YouTube Channel:

- https://www.youtube.com/channel/UCTcb5In_PdeTau_wJEG6FPQ

Marshall Public Library Digital Yearbooks:

- <http://marshall.advantage-preservation.com/>

Marshall Public Library Digital Newspapers:

- <http://marshall.advantage-preservation.com/search?pt=31509&bcn=1>

University of Rochester transcription guidelines:

- <https://www.cs.rochester.edu/research/cisd/resources/nonwords.html>

Batch rename file extensions

- https://fileinfo.com/help/windows_batch_rename_file_extensions

Temi

- <https://www.temi.com/>

Rev

- <https://www.rev.com/>

GMR

- <https://www.gmrtranscription.com/>

Link to MPL IDA Collection

- <https://cdm16614.contentdm.oclc.org/digital/collection/p16614coll50>

Q&A

Heidi Margold:
hmargold@illinoisheartland.org

Erin Rose:
erose@illinoisheartland.org

Cataloging Maintenance Center (CMC):
cmc@illinoisheartland.org



80

(Heidi)

Open up for the Q&A session and Emily will read any questions that have been asked and both Erin and Heidi will do their best to answer.