

Chronicling America: Using Historical Newspapers – Transcript of audio

This will be helpful for people for their loans. We have a couple different image formats on the site. We have a PDF for quick, full-page downloads. We also have that are created for the preservation process. These are not accessible through the website. Then we have JP 2000, these are the images for the site. These are download files. They are as zooming in. We have also XML file and these are for optical character recognition text. Or, OCR. OCR is what a machine reads when it's looking at the page. If you look at it, it reads, the Boston Journal has a graphic and racy corresponded in Alaska, it is however reading -- It got racy correspondence. It cannot read the content as well as humans can at this point. The thing to note about this, when you're searching the OCR, you're not searching what you can read, you are searching with the machine read. That is really a helpful thing to think about when you are doing your searching. I also wanted to make another note. We are standing for microfilm. Not the original product. If it was a little out of focus that day, it translates into the official image. As you can see from the bottom image, if the column is too tight, you won't get really good focus that way. They also tend to have tiny text compared to other formats like, books. That is another factor that makes for more difficult OCR. Another thing to think about, quality their abilities. It varies because the national digital newspaper program has been running from 2005 up until now. So OCR engines have really per move dramatically over the last 20 years. Even if two images look identical, you say, hey, why is OCR so much better for this page and worse for this page? The first page was digitized in 2005. The second one was digitized in 2022. That might be the difference that you are seeing for the OCR. So, I will do a live demo a little bit later but I want to show you what it looks like. This is the product. This is sort of the basic search by state, date, throw in a keyword or two. On the advanced search, you can look up, state newspapers. You can do FrontPage only. Let me get the stamp. Here, if you are a teacher and you are only the front page news coverage, you can click FrontPage is only in your search for we also have a language learner below that. If you want to search German or other languages, we have the third. We also have some options for what your searching. It could be a keyword, all the words, with a phrase, or all of the words within each other. If you're thinking about doing a phrase search, I recommend using five words. With the OCR, it might be that the engine is looking at the text and it says, oh, that does not say -- You might want to try the flexibility within five words. That is going to possibly get you some better search results. Now that we have gone through the basic and advanced search, there is another tab that is all digitized newspapers. This is from a particular state. A particular language. These are all the titles from North Carolina. It has information about the title. How many issues we have. Sometimes we have one, sometimes we have 1400 issues. Sometimes, way more than that. And the date range -- If you click on a specific title, that will take you to title information on the left. We have a calendar view. The calendar view is really helpful. This is a way to browse the content you have at a browse. If you are looking for someone who was married on a Monday or a Tuesday, you can see it here. The paper only publishes on Fridays. Don't only search one day or two days or necessarily three days of a particular life event, you might have to go further out to make sure you're catching the publications scheduled in this newspaper. It can be very helpful, let's say there is a stack of newspapers sitting under a leaky faucet. Three months of the newspaper was damaged from water damage. It's not recoverable and did not get filmed. If you look at the calendar, you might say that there is no digitized content. It could be me not getting the right search terms. That is helpful information that you can figure out on the calendar view. Now that we started drifting and talking to genealogy concerns, there are a few things you can do to improve your search for the newspapers. First, you will want to have any important dates or locations before you get started. Birth, marriage, death, where people were living, like a state would be good. Particularly if you have a common last name. If your last name is, Smith, you might be driving yourself crazy looking at all of the content in America, without a particular state. Modern newspapers are different than older

newspapers. Something you can reliably think about, there are sections for obituaries. The further you go back, the less that becomes the case. There are other things to prep your mind for -- You have to think about why your ancestor might be in the newspaper. Newspapers wrote about certain things. They wrote about sensational accidents, crimes, gossips. Maybe they were involved in politics or community organizations. Before you get started searching for your ancestors, you might want to think -- You find what you find, it might be good or bad, but they will likely cover the bad. Keep that in mind. Newspapers are often like local notices. They could be called, newspaper news. It's usually on the third or fourth page of a four page newspaper. I would love to give you a name, but they have different names. I can't do that. They are never just published on page 3. Unfortunately, I can't give you a quick and easy search. These columns are really great for all sorts of information, but they can be tricky to find. I want to show this particular column, because there are some things that are helpful for genealogists to think about. First of all, we have initials. The first one here, J. M. Weeks. There are qualifiers like, Ms. or Mrs., where the first name is removed. I'm not sure how to pronounce this, I've only seen the abbreviation, but those are some of the things that you would want to search for. Mrs. Nelson, other than, Bruce Nelson. This example, we have Dr. and Judge. If you know anything about your ancestors occupation, that can be helpful for searching. We can see that sometimes it replaces names. This person, Judge Thomas B. Brooks. There are also titles like, Captain, Colonel, Lieutenant, master, Mayor. The more you know about your ancestors and what they were doing, the easier it will be to find them. As I mentioned, married women referred to by husband's names. Mrs. Pierce is here. That is something you want to keep in mind. Sometimes the children are not mentioned at all. A little child of Mr. Womack was kicked by a horse last Tuesday, but we are glad to state is doing well. You really have to sort of think about, Womack is an unusual name. If Yorty have information about your ancestors, you could say that they had a 16-year-old and a 4-year-old. So I figured out who this is probably about. It does require a little bit of logic and not necessarily searching by first and last name. I also encourage you to check for common misspellings. The publishers can be a little bit off. This is probably my great grandfather? You can see her, this is Sir -- You might want to try to check it with just one F. This particular article pops up in the last year. Something about growing -- This is something that has not been there in the past year. If they are good about adding content, like they are, it might be worth checking back. I mentioned earlier, you're sometimes limited to particular place and time. If you have an unusual name like mine, just throw your name into the advanced search and -- This is from Ohio in 1913. There was a newspaper article about a family member. It is in German. The font that you see is not fun to read. This is, New Jersey. Egg Harbor city, down here. If you don't want to take the time to figure out what the characters are, and you don't speak German, if you click on the text, use the OCR text page. This will take you to a translation tool, you will get a rough idea of what was said in the newspaper. That can be really helpful. You know if it is something worth pursuing, or if that is not your ancestor you can move on. I would not recommend that as a way to get around. I will point out that little text button. You will click on that to get the OCR text. Another great thing we have is a clicking tool. Once you find something you're looking for, click on the tool, it will make a copy of the article and include on the page all of the information about that particular article. If you are someone that really likes to print your findings, it's a good thing to do. Because you write something on the back of a newspaper clipping that you printed out, it gets smudged or faded, nobody knows where it's from, 15 years later, this is a good way to avoid that problem. I recommend that if you are a printer kind of person, use this tool, it might save you some trouble for years in the future. I would be remiss if I did not mention how many newspapers are still on digitized. We have a U.S. newspaper directory/directory of U.S. newspapers in American libraries. It's the same data, just a different format. This is a really valuable resource to figure out beyond what's been digitized, what other newspapers are there? What are we willing to pursue? You can see from this example, there are only two known issues that exist in libraries. This would probably be good for genealogy. There are not many articles that survived in libraries, this is something that would be good for you to know before

looking for this particular newspaper to only find out there are only two issues that exist. Let's start talking about chronicling America in the classroom. Newspapers and journalism usually refer to, "first draft of history". They are pretty good reading level for students. They are not -- They are designed for a general audience to read without too much knowledge to start with. They are also really good to find the local spin on national news. How did your particular area experience World War I? This example to the right, you can see that this is how suffrage would help people mobilize in Pennsylvania. You can see multiple perspectives here. You have women in favor of suffrage, and I don't know if you can read the sign? It is the Pennsylvania Association opposed to women's suffrage. You can often find different perspectives within the newspaper, as well as different newspaper titles. There are some complicated parts about using newspapers in the classroom. Often to do with language. We talked a little bit about suffrage, there are sometimes older terms, you will want to use suffrage for things happening in. There are also different spellings for words use. Airplane instead of aero plane. A more complicated, is the language. A lot of the language is a product of the time. It is quite biased and offensive. This is something that you want to make sure that your students have contacts for. That they may be searching for words that they would not necessarily use in a conversation today. Context. This is a really key concern. You really want to think about what was happening during that time and if it relates to the article. It helps to check newspapers that were written by and for being covered. If you're doing a search for an article about a particular trip, you want to find a newspaper to see how the wrote about that particular event. The story often involves. It is helpful to check the title essay. I will show you one shortly. The title essay's provide valuable context into the perspective of the publisher and the particular audience for the newspaper. One way to get around the two Guinness of newspapers, as the check out one of our topic guides. They are reflections that we pulled together of popular topics in history. You can see here that we have a couple different options here. I want to show you what one looks like. There will be a short summary, a timeline of important events that happen for that particular topic. The organization of the first mining district in Alaska in 1959. Verses, 1902. Your students are going to want to search within these two boundaries, that will probably be the best time. There are also good search terms. When you're looking for topics on the gold rush, your students might not immediately think about prospectors. The topics really pull it together so the students can get started even if they don't know the right categories. Now they know some of the words they will be looking for. There are also selective articles. This is a way for students to jumpstart right into the topic without having to know precisely what to look for. Title essays, I've been talking a little bit about these. The are fabulous. I highlighted some of the important parts of this essay. This particular newspaper was founded in Salt Lake City in 1895, by an African-American name, Liz Taylor. I was delighted to find out that this was also published in Salt Lake City when there were fewer than 2000, 1000 African-Americans in the state. This stood out as an animal I anomaly in the 1890s. The students are getting search results from newspapers in Salt Lake City on different topics. The perspective may be different than the perspective of another newspapers from the city. This is helpful to show multiple perspectives. Finding all that information for students, a lot of these essays require background in histories, background in newspaper histories. A lot of times you have to read lots and lots of issues to wrap your head around what a particular editor or publisher was concerned with. These can be really helpful and can save a lot of time. We have a bunch of other resources for teachers. These are from our outreach team for the library of Congress. If you want to get started with a newspaper or article in your classroom, this can help build some content around how students will work with that particular article in your classroom. They are really helpful and interesting. We also have some webinars. If your interested in national history Day, we have articles that were recorded this year. We have one about using newspapers were generally from 2020. There's a really good back one from 2015 that I did with Tom over, that was really fun to do. These are from people that use newspapers all the time in their work. They also have other resources like, lesson plans. This is all from our educational outreach team. It contains newspapers, as well as sources from the

library of Congress on a particular topic. They also have a fantastic blog and twitter account. We also have a blog in my division called -- We do a little bit of newspaper,, book, in that particular blog. We do have a twitter account about chronicling America. I want to pin it a little bit here. I want to mention the citizens history project. This particular one is history unfolded. It's out of the U.S. Holocaust Museum. It started in 2013. Researchers investigate what U.S. newspapers published about the Holocaust. You can see an article in chronicling America about anti-semitic violence in Berlin just before the start of World War II back in 1938. As part of the project, you can go to the website, the newspaper database, go to the microfilm at your local library, or you can participate in the project by participating in research print. We've helped a couple in person -- We also did two online this year. If you're interested in this particular project I will go to the Holocaust website at the top. This is probably the best way to be informed in all future -- If you're interested and want to pursue this, that is probably the best way to find out about what was happening. Pivoting a little bit more here. I want to talk about the data projects of chronicling America. We have two major ways to access chronicling America content in a programming way. We have application programming interface. API. And we have download options. The API, a lot of the information is on the page. One thing I want to say for an answer to a research question, add and format equals JSON one. That will turn the search results into. It can really help getting started to figure out what you can use. We also have list serves. If you have anything that we want to let the data users know, we will send it out to that user. We also have an email address. This is used for any questions you might have. If you run into a problem with limiting, this is a way to work with that. We can tell you more information there. Beyond that, we have a ESRI , and they help with things that you are thinking about as you're going along. That can be a helpful way to see what you are doing. If you're not getting what you. If you're trying to do a project on Utah newspapers through 1840, we may not have the newspaper from Utah from 1840. That is something to think about. For a large-scale data project, we have download options. This is an OPR database. Here you can download plaintext or the file, or if you want to use all of the data, you should use the data page rather than individually. We also have JPEG files that I was talking about earlier. You could grab them that way and it might be easier for you. We also have a WS, you could email to the email address. That is a good way to get the conversation started. I wanted to highlight some recent projects that people have done and it would be a navigator. This is from Ben Lee. He has a data page and a interact page. The interactive page, he identified all of the photographs. With the app you can find images just like you wanted. It's really interesting project and it's well documented. We will be talking only about a few projects for time reasons. That is why I am pulling out some of these particular projects. There are general articles about these projects and more information about how they did some of the other projects out there. There are also really good projects called viral text by Ryan Cordell and David Smith. It's about 19th century newspapers. They did some interesting stuff to work around how not great OCR is. When you're identifying the content, you really care about the OCR. You want to give the text. It's hard to match. To pull our worlds together in the classroom and the data project, there was an interesting data project recently at Arizona State University that looks at newspapers as objects. You can see that they have notebooks about some of the things they did with newspapers. Again, this is really well documented. You can see some of the -- I put some of their stuff up. It is about a data project that uses newspaper data. A lot of the stuff that Arizona did, there are similar lessons for this particular project for the digital library of the Caribbean. Giving a shout out to those student data projects. What are we talking about today? We have a page called/actress. I shortened this up to show you an example of the stuff we have on there. If you're interested on this kind of content this is a good way to find more of it. We have lots of different ways to connect. There is a and DNP page. If you have a research question, the asked the library and section can be very helpful. We have our blog, Twitter account, and if you have technical questions about the API and the text, this is a great place. Now I will head over to the demo site. I think Helen that you wanted to make an announcement?

Yes. You have less than 20 minutes we have quite a few questions, so we may run out of time answering those questions. Your screen is going to change as Robin goes ahead. If you want to activate the chat, you are going to have to move your mouse to the blue bar at the top of the screen. Click on chat. Add your questions and send them to all panelists or all participants. Robin, back to you.

Great. Helen, do you want to read some of the questions?

They are by general topic. We will start with some -- Let's see if there are some false? First question -- How soon will you be adding more newspapers after 1922?

Probably this month or next month. We are constantly adding content to the chronicling America website. We added more this morning. It really varies from date to date, because it depends on who is submitting it. Today we uploaded content from Laramie Republican from 1896, two a long period of 1000 issues. Wyoming content. This is from the Virgin Islands. This is 1935. St. Croix. Often times we are adding new content.

Similar question. The question, you may have already answered this, why do digitize newspapers not available beyond 1963?

I am not a lawyer, but there are some changes in how the rules for copyright registration changed. It does not mean that something published after 1963 is necessarily in copyright. As the rules get more difficult to comply with -- I am sorry to not be more specific on that.

The next question is similar, you may have a similar response. Do microfilm producers have a copyright claim to microfilm that would prevent it from being published even if the content itself is in the public domain?

That is a interesting question. I do not necessarily know the answer to that one. I will waffle a little bit in terms of -- Often times for digitizing microfilm, you have to have access to the microphone. Some microfilm producers might not provide a copy to you to digitize in certain scenarios. I guess the thing is, less copyright and more, if the producer will make you a copy, in many cases.

We will switch gears about OCR. There are a handful of questions. Is there a way to report corrected OCR? Sometimes I have to look at the original, too.

We don't have that option on our site. A lot of our partner sites have -- Have an option to view OCR. Unfortunately, no, not chronicling America. I'm going the wrong way. The Library of Virginia does not.

Two questions came in, they are similar, so I merged them. Will you be returning to the early digitized papers to try to OCR them to take advantage of advances in OCR technology?

Did is a interesting question. We are looking at that now. We are looking at the very worst OCR that we are thinking of improving. We have over 20 million pages of content now. If we started this process, it would not be quick. This is something that we are actively investigating now.

New question topic -- A complete is the holdings information in the directory?

The directory is a really interesting thing. Let me pull up the district of Columbia, this is a smaller state. The directory is not perfect. This is an example of the directory having a record of it. If you click on that, and will take you out to world cat. It was show you other libraries. Sometimes these are electronic, so they are in one database. There might be 62 libraries that have a, but it's only one database. It looks like the case of this one. The directory is a good starting point. Let me back up for a second and talk about how records get into the directory. Catalog users across the country, catalog newspapers. Some of them push them to world cat. We pull down cataloging records and catalog them to create the directory. You will not find unique content that you won't find in WorldCat. If anyone has ever searched newspapers in WorldCat, there are a lot of records and it's hard to navigate them. That also makes updates to the directory tricky, because a cataloger who has content who can update the records, that is how updates are made to the directory. Every time we do a new poll from WorldCat, we get all fresh data. We don't make any changes to the directory, we would make a change to a cataloging record. That would go to WorldCat, then we pull it down. We don't want to make changes to the directory that we have to layer on data that's getting refreshed from WorldCat.

The next these questions are related to black newspapers. I will start with the first one. Is there a way to identify black newspapers, also is there a way to search a black only newspaper?

Yes. This is a wonky thing at this point. We are moving a lot of our contact and we will have a lovely search interface. Right now the way to search and only African-American newspaper is kind of wonky. I will pull it out. If you email us to the research library, I can send this information to you. But the University of Florida did a guide on how to -- Essentially you have to throw a lot of newspapers into the search. Let me see if I can grab this. There we go. You create a really long search, essentially. I guess I have to download it. This is from February of 2021. We have quite a few of those newspapers. You plug them into the URL of the site. Sorry, this is taking a moment. Like I said, if you send us this question, I can send you this very long search. Essentially, it will let you search all of the African-American news all at once. There is a much better way coming in the next version.

A couple questions were partially answered. What would be the alternative terms for black? For example, African-American, colored, is there a guide produced by chronicling America staff? As a general knowledge that you have gained through experience? Someone would in the chat link to one of your guides concerning the subject.

Yeah. So, we have -- Let me -- We put out if thesaurus, different terms that can be helpful to look for different ethnic groups. You can go here. There is obviously a note about harmful language here. A lot of the terms would be considered offensive. If you're trying to search for terms in newspapers that you need to use words like these. There is really helpful information to contextualize the terms in the thesaurus. It also shows examples of the words in context. We have a newspaper here called, colored American. And, similar. Those are some of the terms I would use. The first source is a good place to get some other information.

Can you please recommend why you used a clipping tool?

You get all of the information about the page. You can see here, there is a nice zoomed in version. Here, the state was involved in dairy men. Get all of the information from where that came from.

Where do we find, titled essays?

Okay, you can use this in a couple different ways. You click, about. There is a title essay. You can also get the title essays from the all digitized newspaper list. Wherever you see the word, that will take you to the title essay. In the next few questions are more specific. We have about three minutes left for Q&A at this time. Is there a plan for.

Right now we have state partners for the past couple of years in Massachusetts. The Boston Public Library. That state partner is starting to digitized Massachusetts newspapers. I don't think they are specifically digitizing the newspapers from one institution to another, they are picking titles that are important to the state history. Informally, for the early papers, they are re-digitized. If your interested in antiquarian newspapers that have been digitized or you don't know they have been digitized, email or ask a librarian service. We can let you know if they are out there already.

Okay. Someone asked, would you expect the lovely new interface?

That is a great question. We hope sooner rather than later, but we don't have a specific date at this point.

Let me look -- It was already answered in the chat, I will reiterate. Our states that have Artie done the two years, would they be able to apply again? Somebody replied, yes. They had several grants over the years.

Yes. You can reapply, most definitely.

Okay. I'm going to break apart this one question. Somebody came a little late to the webinar. I'm looking for a specific political cartoon. For example, during World War II, how would I start my search?

Work for two is -- They are the worst. Email or ask the librarian service, if it's by a particular cartoonist or a particular newspaper, that might make it a little easier if you know the cartoonist. They are harder to find, because they have handwritten text that OCR engines are bad at reading or they have no caption text or text on the image at all. Searching the OCR is meaningless. There is no text to search. We can certainly try to help you and we also have some very good cartoon people in our photograph division, we can check with them to see if they have leads.

Okay. One more question, how did you get to the list of the latest updates? That repeat that motion.

It is just events. You can do/evens. You can see all of the boring media stuff getting added and unedited. There is also a way -- You can see recent things that we have updated. You can say that I am scrolling down. It will say, we just updated content for this particular title. We have a lot of titles that are updating frequently. It can be hard to keep up.

Okay. If you would not mind Robin, stop to share your screen. I will start wrapping up this webinar. Robin, your talk was fantastic and informative. Thank you for spending time with us today. I am now sharing the chat a survey link, please take a moment to let us know what you thought about today's webinar. It has been recorded, and you will be notified shortly after of when it is able to be viewed. We have one more in the month of December. It is called, supporting the rights of the public to know. We will do our best to get the links for all of the things that the people were asking for in the chat. Thank you for your presentation, and Robin. Everyone, have a marvelous day. Thank you. New line. [Event Concluded]

