Government Documents in the News – Audio transcript

[Please standby for realtime captions]

Hello everyone and welcome to documents in the news. Before we get started, I have some housekeeping reminders. First, use the chat box for questions. We will make sure to further adjust them in the program. We are recording the conference and you will receive links after the event. Please join me in welcoming our presenters. I will now handed over to Andrew, to start the session.

Thank you, Melissa. Welcome everybody to documents in the news, and exploration of government information use in newspapers. Welcome everyone , GPO for putting together another successful conference. I am a librarian for political science policy studies and government information at Penn State University libraries. Sandy works across the unit along the rest of the staff. Documents in the news -- news articles cite a reference of government information. We have all read an article in the newspaper. I can remember when we first started reading the newspaper -- finding an article that set it some piece of government information. It could be a court case. We have all seen them before. This was an attempt to take a closer look at that, through the lens of the library database, to see what we can find. Sandy is going to drop the address in for the guide, where we host documents in the news and in the chat. I would encourage everyone to take a look at that, as she posts it from within the chat. If you want to, there are two different sections of documents in the news. There are federal documents in the news and then state documents in the news. The link will be towards our U.S. documents. There is also a tab for Pennsylvania documents in the news. I see she has done that. Thank you, Sandy. It is a project to curate both government and state information in news sources. It was never intended to be comprehensive. It is a snapshot of news we find in the database. It is an avenue for all of us in the unit to foster government information for ourselves and the ones we serve at Penn State. It was also initially a potential instruction for. It has been used on a limited basis for instruction at this point. It has been a wonderful professional development opportunity. We will be better prepared to handle questions at our service desk. It is also a great learning opportunity for all of us. It has sparked my interest in how government information is being used in media. I hope this may inspire others as well. That is something is a community we could be looking at, how all of this information is used a little bit more carefully. To understand a little bit better and serve our communities better. Now I will hand the next slide over.

Hello. Thank you, Andrew. Thank you for that introduction about me. Thank you all of you, for attending this session. I work at the government information as a library manager. I have been with them for quite some time. Microphones and government information is a new unit that began in 2018. How did that -- they are directly under social sciences. Going back even further in time -- government documents were in one department with maps. There have been a lot of changes over the years. Currently microphones and government information -- government documents unit has two faculties and six part-time staff, including work-study students. The full-time staff and most of the part-time staff work at the records desk. I would say most of the staff -- some students do different tasks. They are not scheduled at the desk. So, what they do is -- the shelving is done by some of the students that really -- that is their job. Sorry. Our microphone collection includes over 96,000 reels. One 44,000 piece. [Inaudible - static] we provide research services and government collection provides this material. [Indiscernible - muffled]

We want to have this work done and integrated. Staff work two or three desk hours in a row. They will have time to start thinking about curating. [Indiscernible - background noise] The university library had a desire to keep working during the shutdown. As we move forward during the pandemic -- we pivoted and actually spend more time working on the project. [Indiscernible - background noise] .

We had remote training sessions for curators, as we moved. Those training sessions involved discussions about multiple database options that we have, that we have available to us at Penn State. Parts of training sessions we have been working on. There were lots of discussions we had going to understand everyone's perspective. And what people were seeing from their point of view. One of the interesting things -- when to move on. If you have identified a document and you are trying to find coverage in a newspaper -- if you are not having any luck, when do you stop. When you start to think about another strategy. You have to move on to another document. We have instances where documents should be covered and we just can't find them. So, those strategies -- them strategies -- some specific strategies would be to embed any more specific web address links into a search. Into a newspaper database. We also discovered date lags when a document is produced. It might take a couple of days or a week or more, for a newspaper coverage to take place between federal and state documents. One of the strategies we settled in on -- basically, someone would go to an agency or a producers page -- look in the instance of the courts. You can look at the recent rulings from the courts. Executive orders from the White House. Or, news releases from department agencies. We had great luck with that strategy, somewhere between three and seven days after those hit the -- after those are produced or announced. Actually being covered. Again -- I mentioned both documents. We have guidelines. Samuel will talk more about those guidelines. We encourage independence. We want them to have fun exploring document use. We wanted them to be driven by their interests. And so that was one of the things we always came back to. And we had to develop a workflow, to get the articles to get the articles from curators, to the guide. One of the things that I will mention -- we actually have a format -- if you haven't looked at the guide -- Sandy put the guide in the chat. There is a format. It is specifically done with a format that was established with the idea to identify the document. Site the newspaper and database. An example of that would be -- something is taken up by the house with an HR number. A newspaper article has set it that. We give the direct link to that particular piece of legislation and we give the citation to the newspaper article and the database exam. Then we provide a link to house.gov or Congress.gov. The overarching producer of that document. One of the reasons we did not provide links to the newspaper articles -- we wanted to avoid -- we wanted this to be a tool for anybody beyond Penn State. Those links to our proprietary newspaper source -- we didn't want to give into that. With that in mind, Sandy will talk about workflow.

Hello. I am Sandra. I work at the university library full-time since 2006. Let's talk a little bit about the workflow. The workflow started with general brought instructions. We needed to tighten up those instructions. So the staff knew what to do and where to look. They wanted to focus on what their interest is. First we identified where they could find government documents online. Then we evaluated the Penn State library databases. We selected the ones containing newspapers. We also suggested searching online news sites, to find relevant government database documents. After those were established, staff ran the newspaper articles and placed them by month in a shared word document. Now that we have all this information, after a few discussions, we offer to place the information in a current guide and creating tab sports. This gave us a separate space for both documents. So, we transfer the content over from the word document, making sure it is in the correct format. No spelling errors. Basic editing. The guide has its own limitations. We had to work within their parameters. The layout and color of the text was preset and could not be changed. Only so much content could be placed in one of

their boxes. We adapted and broke it down and decided to archive the content in a PDF. Every time a new box was created, it needed to be republished. Back to Andrew on contents. [Inaudible - static]

We try to find that coverage. The executive branch is covered extensively. That is probably no surprise. I think the popularity of executive orders in the last 15 or 20 years is evident. State content really follows similar patterns to the federal content for the most part. A deep analysis of this content is limited, the way we curate it. I have some things in work to look at larger corpus is of newspaper data. And thinking about government information coverage. I will talk about that in some subsequent slides. Now I will hand it back over to Sandy. She will talk about duration.

I think these are the changes and challenges, right?

Yes. She will talk about changes and challenges. that is okay. Just like anything -- this program started out -- we started out in November 2019. With the discussion and slowly starting the training and data collection. But then, there was a challenge. We had the holiday season and staffing issues. Everything was kind of slow in the new year. 2020 we broke out and tried focusing on this project again. Then we were hit by the pandemic in 2020. And we found all of that writing. That was a difficult challenge. As a manager, I needed to know if the staff can work remotely. Do they have a computer at home? Do they have Internet accessibility? Can they communicate? It is a challenging time. Things were not going smooth for a while. Then some part-time staff reached out -- they wanted to retire. There was another challenge. We were trying to higher and train additional staff. While working on hiring -- I had to get my full-time staff to devote this project -- this allowed the project to move forward. Communication was essential. I ended up needing -- I met with the staff individually, more often to ask questions and to make sure they were making progress. With all of their problems -- after a few meetings I assigned a specific database. Soul, they were not duplicating searches and getting frustrated because they were not finding or, they are finding the same documents. There comes another challenge. Due to the budget cut, they needed to access -- by far was the favorite database to use.

That contributed to an article found especially -- the state documents. And required the staff to expand on their searching skills, for other databases. Now to Sandy, on duration experience.

Hello. Throughout this project training, training was ongoing. There were always questions that popped up. Some examples. The government documents mentioned in the newspaper article, two different agencies. Which one should I use? The answer was both. The courts are not in session. What should I do? The answer is to bring your search to government agencies and see what they have posted. That did create some problems. For some, they were created by courts or general assembly. It does take some digging to provide that content. Some strategies changed according to databases. Somewhere easier to search, than others. And is best to browse through the headlines on a daily basis. Time delays have been encountered from when the government was referenced on their site. There may be a delay of a week or more. When we did lose subscription to the world news, it is a database that was easily searchable. That contributed to a lower number of articles. Staff searching skills increased as they spend time on the project. They came up with different ways. I heard something on the news. I will check the agency and see if I can find a newspaper article. Now back to Andrea on moving forward.

Thank you, Sandy. So, a little bit more about where we are heading with the project.

Ashoo Kumar mentioned accessing the worldwide database. They will want to assist us with a decision on the project. I think we have four candidates now. They are getting started on interviewing those

candidates. I think that physician will run through May, up until the end of April. We discussed Rick hurting curators. The news and media library would have been interesting. That group is particularly interested in using that as an instruction tool with journalism students. Especially in the communication area.

I have not utilized this particularly sourced information on any of my instruction sessions. But I hope to actually have some of those sessions in the spring. I will work it into those. Some other things we have talked about -- we talked about subject or agency department rotations. We would look at a particular three or four agencies every month. We want them to dig a little bit deeper into the coverage of the agencies and learn more about how much information they can produce. We also talked about having curators specialize in particular areas. We kind of did that based on their interests. But, we never really said to anyone -- you need to cover this, this and this. That is something we did talk about. Although we did that with state documents. They found -- they enjoyed the state documents work. We thought it was important to cover state documents as well. I mentioned earlier -- the analysis of this content is a little bit limited due to sample size. I think there are some parts -- they reinforce what we already know. Someone in the chat asked about this. That was one of the predictable things. They were very well covered. They were in session and making rulings. That is something we probably already knew before the start. I plan to expand this analysis to include some technologies and advanced searching of larger data. The database has an API that will require some pretty advanced search strings to dive into what kind of content is being set it in that ETA base. That database is also not just a general newspaper database. News services syndicated services there. There are some things to think about. I already started working with some folks at LEXIS-NEXIS to develop some extensive search terms, to attempt to extract a little bit more data on newspaper coverage of government documents. At least within the LEXIS-NEXIS database. And I also hope to evolve other news sources outside of the library databases. I wouldn't necessarily say that was the most important project. There is some other technologies I would like to explore. To see if I cannot carry this a little bit further. Rather than just looking within our proprietary newspaper databases. To get maybe a better measure of news. Within the news. And to ring that up. Feedback and thoughts -- your feedback is welcome. We love feedback. Our emails are on the front and back end of this presentation. Please fill free to put anything in the chats. Feedback and thoughts wise. Ideas for instructional use. I would encourage anyone and everyone to share this page with your colleagues. Promote it for instructions. If you feel that it is in line. Just a special thanks to everyone who continues to work on this project. A special thanks to Ashoo Kumar and Sandy, for making this happen. With that, we will take some questions, thoughts or discussion.

Hello everyone. If you have any questions or thoughts, just go ahead and type them in the chat. We will relay these questions as they come in. Gabriel would like to know if you have a certain string methodology.

We don't. That is something I have been working on. Related to the news content. I attempted to come up with some search strings -- early in training I was looking at different strategies. Nothing seemed to be producing consistent results. So -- what we did, we really settled into looking at agencies. Looking at the White House site. Looking at the agency websites. Looking at Congress. Looking at what is being produced. Following that production into the newspapers -- to find that article. Sometimes these articles don't explicitly say -- they don't always give a websites. They might mention a particular reports. That was on us to spend a little bit more time looking. Often times those are easy to find. Even if they don't cite them explicitly in the newspaper article. We can still use that article. It clearly cites a government documents. It is easily available. So -- the search string methodology is something that -- I would say up to this point, the answer is no. I really thought that there would be some more patterns from a certain

perspective. We really settled into the other routes. But I have -- I have been working on a bank of search terms, to take a closer look at that.

Thank you. Do we have any more questions for Andrew, Sandy or Ashoo Kumar? I don't see anything else coming in. I will give everyone just a few more moments, in case somebody is typing.

If anyone wants to reach out to us at any point in time, please feel free to do so. I know -- this could easily be replicated for your particular state. Depending on how you go about this. I would be more than happy to talk with you after the conference. With anyone, if they are interested.

Thank you. One more question just came in from Marina. How has the guide been received by the student community and faculty?

I have not -- I have not had any feedback from the news and communications members on how on classroom use -- and how successful it has been for their teaching. We were down a position in that unit. I am not sure. I haven't had any direct feedback on their use. -- I have not personally introduced this to any students. The news and media library and I know is going to give this to journalism students. But I have not had any feedback. As far as faculty -- I don't think I had any faculty feedback. That might come in those particular sessions. Especially the journalism students. Were, somebody in a school for communications.

Okay. Another question from Robert. He is curious to know if you have considered or tried searching on newspaper media.

Robert, I guess what you mean by non-newspaper media. Like blogs or --? I wonder if you can't elaborate on that. I am thinking about -- I am thinking about the next generation of this. Looking at a larger piece of data. Thinking about LEXIS-NEXIS, there is something in there called target news service. They are actually government information services. They have them on a whole host of other news media blogs. I think the trick there is to separate -- source them separately. Will treat them separately. The government information news source services -- I forget what the name of it is. That is what they do. That is going to be -- it will be evident that they are back there citing that. This particular project was looking at more traditional newspaper media.

I think as far as blogs or radio or streaming video -- I think that is something down the road we can look at. But, that would require some technological advancements to do that automatically. We would have to figure out a way to do that.

Teresa asked, if it is okay to use your lip guide on her website.

Absolutely. Feel free to link to this. Feel free to use this for teaching. We really kind of designed it for teaching. PDFs were available. If you are looking at the guide -- we have the dates. We have the actual document linked. We have the title of the article and the author of the newspaper and the database. That is not linked. I specifically thought it would be great to teach students, how to find that article based on the title. Or, the newspaper or the offer, or all three. And then the last link we have included -- the overarching producing information. The latest entry in documents and news under federal. October 19th, 2021. Short term energy outlets. It is a great documents. We got a call from Allentown. They set it that and provided a link to the administration. They used the actual document. Not only do students have the actual documents -- they have the exercise and finding that document. Then, they also have

the link to the overarching producer of the contents. Which gets them thinking more broadly. That is kind of how -- that is the idea I had.

Gabriel asks, do you have a RSS for the news?

That is not something we have. It is something we consider. Not at this point and time.

Thank you. She wants to confirm that the main audience has been media study in communications.

I think that is accurate. With the exception of also one of the goals -- I gave up a brief introduction to this project. I wanted to share this more in depth project with Sandy here. One of my main goals moving forward is to promote the idea of government news and how professionals -- there is lots of room to study these. There are monologues on news. That is as much of the audience as anyone. So -- we are excited to hear feedback and thoughts. We welcome anything to that nature.

He wants to know how many hours on average you spend on a project.

That is a great question. Sandy, Ashoo Kumar? Any thoughts?

I am trying to connect. That is a great question. During the pandemic we had more hours available. Everyone was working from home. Especially when the project was new. It took lots of time to explain everything to them. I will say at that time, maybe 30 hours a week with several different staff. But it can take more than that. Once you start searching -- it is not something -- it can take -- if I have to request for the budget -- I will go with 40 hours a week.

Thank you. Are there any more questions? I'm seeing a lot of kudos. You have a lot of people excited about doing something similar in their library. Just going to give a couple more moments here for questions.

It doesn't look like anything new is coming in. I would like to thank Andrew, Ashoo Kumar and Sandy for a great program. Up next we will have the D.O.C. business meeting. The wrap up session will begin at 4:45 Eastern time. If you want to join that program, you have to close out of this program first to enter that meeting program. For now, we will take a short break and pick up again at 3:30 Eastern. Thank you.

Thank you, look Melissa. Thank you, everybody.