Digital Content Contributor Partnerships, Collaborating with GPO, and Building the National Collection – Transcript of Audio

Please stand by for realtime captions.

Good afternoon and welcome to the FDL P webinar my name is Ashley and I am an outreach coordinator at GPO. And we have Sarah Sessler. I also have my colleague here today serving as our technical support. I'm going to hand the microphone over to David who is our first presenter today.

All right. Thank you, Ashley and thank you, Kelly for working with us today. We are going to be talking about digital partnerships. Collaborating with GPO. Thank you for joining us today. A little bit about the national collection. The national collection of U.S. publication is the geographically dispersed collection of the corpus of federal government information. It is what is accessible to the public. The development, the maintenance of that collection and preservation of the national collection is really critical to providing the free ready and permanent public access that our community expects. The ability to use that is not only critical for now but also into the future. We know that digital imaging is an essential tool for creating and providing digital circuits for greater access and of course long-term preservation. If you have special collections those digital surrogates can take a lot of wear and tear off the original items so that you can help preserve those original documents this work requires a lot of coordination and collaboration. Since the libraries have the collections and these collections the physical documents they have built for their own users they have selected those over the years and a lot of local decisions that are in place to digitally image those collections also occur at the local level for the libraries own digital development. And follow those digital collection plans. What the partnership essentially is is a community digital preservation initiative. It combines the local knowledge that they have about their local collections and their ability to digitally image them. It combines their knowledge with our knowledge as well as our digital repository. This collaborative partnership ensures that discovery, access, and long-term preservation occurs for that digital content. We can't talk about digital imaging without talking about the national collection of digital content and our digital repository which is iso certified. We also have a national digital collection of public information. This is found in GPO's secure digital repository. They maintain certification as a trustworthy digital repository under the national standard which is the highest global standard for ensuring preservation of digital content. The collaboration the collection development combined with our expertise in digital imaging and the knowledge of the digital repository is what ensures that the digital content contributor partnerships are a success. I am going to pass the ball to Jessica.

All right, hi everyone this is Jessica. For the next few slides I'm going to talk about the actual process of what becoming a content contributor looks like beginning to end. To start you might already be wondering if this type of partnership is available to you or if it is a good fit for the work and activities of your particular library and institution or you are just considering a digital project and the content or some of the content falls into scope of government info if either of those statements fit a project I your institution the next best step is to take a look at the preservation and partnership webpages and review some of the guidance materials that we have made publicly available. They provide guidelines on how content files are preferably digitally imaged named an organized for efficient's and these guidelines can be helpful for local digital projects of any kind really but they also provide insight into the level of effort required. We will go into how these files might be arranged in more detail in just a bit. After you have reviewed these materials and have a bit of a sense of what the level of effort would be for you to prepare packages or to perform the image in both the next best step is to simply reach out to GPO. It is

just fine if you don't have all of the details completely worked out. As long as we are in a good spot to have some initial conversations about what the content is and how it is been digitized or will be digitized. Establishing the partnership is absolutely a collaborative process and preparing content is also a cooperative effort. GPO acknowledges the different institutions require in managing the organization and the content. There's not really a one-size-fits-all approach and we don't think that would always work best either. It's best if we have several conversations throughout the establishment of the partnership and then come to an agreement which activities are responsible to which parties and what the expectations look like for timeframe. At this point you might be wondering what the actual details are in order for it to be into the repository. Regardless if there is a single publication are several volumes that includes the submission of usually at least a PDF some JPEG files in a MARC XML file all of that information can be found in our guidance documentation found online at minimum a high quality scan is required and that constitutes the version of the publication a JPEG 2000 file is preferred for each page of the PDF or if PDF documents must be regenerated in the future this is preferred, a smaller amount of storage. If there is no full MARC records for generating the bibliographic information this slide provides a visual representation of what all of that looks like once it is in the final package structure and we would refer to a directory as a package there are many other ways that GPO can provide direct assistance specific for you GPO might be able to use those filenames as long as they are consistent and unique. We can also provide advice on determination of your publications to determine what exactly would be. Content files might also come to us and other formats in addition to PDF we do provide some recommendations for file naming in our guiding staffing this slide goes through some of those recommendations this isn't necessarily super prescriptive this is an ideal guideline this is just something that you could use as a reference as a recommendation for how to go about it something like that so that it is human readable as well which packages for which documents. These are the ideal way and the use MARC.xml files would also be this specific fields the first day of the year the 245 field must also include subfield information if it's not a monograph and likewise the subfields A, B, and C should also reflect the individual title optionally, there can also be a 710 field utilized to the partnership are also two include their names. Those are optional. This slide shows an on-screen visual of a package, its components and how it presents on-screen this is an example from the University of Florida and it shows content that relates to the Panama Canal. We do create a partnership metadata field that displays info that field also makes it possible for us to search across the entirety of our collection specifically for all of the content this slide shows the breakdown of the content in the package one for each page of the document. This template XML allows us to provide additional metadata that is not exactly marks specific. They are not necessary. There is a PDF document that was submitted and also the thumbnail file. Ultimately we hope that all of this information makes the process a bit more transparent and provides enough information for your library to discover what content it could be contributing to our repository even if it is just a single title. Boise State University has been a long time content contributor throughout the pandemic, still to this day I'm really happy to allow Sara to talk about their experience as a partner .

Sorry about that, thank you Jessica and thank you all for joining us today, my name is Sara Sessler and I work at Albertson's Library at Boise State University as a bit of a background the library was given a building of its own Boise college became a four-year institution and our designation as a federal depository began in 1966 evidently the 60s were pre-productive time and yes while we are best known for our football program and the blue turf we are a research focused University Nancy Donahoe spearheaded our effort to become a digital content partner after hearing about the program from the newsletter. A large reason why we were interested in joining is because it had been moved not once, not twice, but three times in the last decade making its way up the building from the first floor to the fourth as the library lost more and more space in the building to colleges to adapt to the times we had

to take a hard look at our collection. That was a problem that had to be addressed in joining the program there began talks of GPO joining the program April 2019 one step of the process was an informal zoom meeting due to the fact that our library was understaffed we were and still are able to submit documents without metadata we were free to go ahead with our test submission and it was time for a personal crash course. There is unfortunately a difference between scanning a piece of arts or a busted up DVD cover utilizing the basic functions I still had that much fun to learn. New software which were new to me which said the paper to put up in the -- thank you for you guys out there who posted live guides. Gathering of any information that can be used for us became sort of a treasure we were also able to consults in the library for the knowledge on top of barring their equipment and for that I will always be thankful. While we were entitled to using Adobe's latest and greatest Acrobat the version of Acrobat that we had at the time would only save the files in J PF instead of JP two. We thought we were fine shame on us for not double checking. To rectify this mistake we worked with library computing information systems to come up with a solution. The squished red cat turned out to be a lifesaver not only could we save the scans we can also use it to make the thumbnails required for submissions. The more software you add to a process the steeper the learning curve this move was ultimately for the best. Here is a look at our very first submission. The document that we sent 10 was a spiral-bound document that was published in 2003 and it was an ideal candidate submission now if you're wondering what in the heck a Mormon cricket is don't worry I had no idea what it was either until this document inspired me to look it up. It is a huge katydid that is very destructive in large numbers hence the need for their grasshopper friends the file format out of the way we resubmitted the document when it was accepted they signed an agreement this is not the end of the journey though as we still had much to learn and work through it has been a constant challenge for modern documents, born digital and have nice clean typefaces older documents can be an entirely different story. We have a bunch of hand written documents that we have been working to digitize reviewing and correcting documents as time constraint when it comes to submissions. It is something that just cannot be helped. Teamwork does in fact make the dream work communications were able to recently secure licenses and is generously letting us share it with them so thank you guys for that. And here is the table of contents for the workflow document that we used here someday I hope to have a complete, completely fancy version done. That can be one of my goals for this year students volunteers and new workers are not hit to the many acronyms of library work and I tend to forget some of them sometimes too simple is best in this case today Albertson's Library has 31 documents, this is a much lower number than we anticipated when joining the program there are plenty of circumstances over the past for years GPO has never tried to dust over this number we have filled GPO, it's truly a partner in this sense this partnership not only benefits our students and other stakeholders it also benefits the public it makes these documents readily available I hope I can give you better insight with the government publishing office.

Okay, it sounds like we are ready for questions. Several have come in. Let me read off pick which to use for the metadata?

That is a good question. Truthfully I think someone GPO might be able to drop something in the chat but I think typically it would be data publication of that is available.

Yeah, Suzanne says data publication.

From Trina, who decides what documents should be submitted? Do they have a list of items?

The collection development plan for access that is available on the website. You should be able to find it to the website possibly. We make that available. The collection development plans often it's the case of

some of these things are publications or titles that we would like to add to the info. We already know exist in some of our preservation steward collections so that we have a geographically dispersed preserved copy of those both digitally and on the shelf so some of those are of high priority if we can get them that would be great but truthfully if your institution just has items and they are in scope of the government and in the scope of the repository then we are more than happy to consider them even if they are not something high-priority in that collection plan.

All right, Jenny has a hypothetical question. She asks, a document, probably older. If it is co-authored by both the federal and state agency could a state digital repository share a file even if it is not a partner?

We do have some instances where a publication was effectively co-offered between an agency and a state agency of some kind or a state author of some kind but the item is being held by an academic institution in the institution it still could totally be within scope but we would look at it at a case-by-case basis. If it turns out that it was not produced with federal funds, it is unlikely that we would consider it as being able to being added to our collection.

I can see that my colleagues are definitely filling in the chat. To continue on with the questions, is there a minimum level of accessibility that they except for these documents?

At this time we do not have specific guidelines that relate to that. In the future we may need to if regulations were to change but at this point in time we ask that things are minimally, not minimally but we ask that things are as full as possible and if they cannot be or are not that is also something that we can work with we might have the resources to do that ourselves. Depending on the quantity of the content.

The last question that I see in the chat is from Sherry. It's great to hear about creativity with Boise State. I wonder if it would be possible to subsidize or provide licenses for professional software images.

I think that is great feedback. I don't think at this point in time is something that we have considered or looked into or have the availability to do.

All right, do we have any more questions for our presenters or any more comments? I'm giving everyone a chance to chat things out. In February so far we have five webinars scheduled. We have one on the first on the web space telescope we have one on the eighth in the federal rulemaking process we have on the ninth a webinar on the prices and wages and an overview and on the 16th we have a webinar on investing in the digital age resources for U.S. securities and then on the 23rd we have searches and statistics. Let me put out a link to our webinar survey really quick. If you can fill that out we would appreciate it. We are currently looking at gearing up for training in the next fiscal year here if you have any ideas or any other questions please fill out the survey and let us know what they are. We have another question from Jenny. There is some advanced software but not all libraries afforded. It is unlikely that we can provide that to everyone the frequency, it's a little bit different between receiving a large amount of files all at one time and then applying the process to those verses the expectation of doing a one-off file every single time. It's just the kind of thing that we would have to consider and discuss. GPO does currently have access to technologies. It is something that is open to consideration.

She can't remember if they have the new software. A

As far as I'm aware that would be not GPO I don't know for certain if they have it but GPO does not.

All right well I am not seeing any more questions in the chat. Inc. you very much to our three presenters for pulling this information together so with that stay tuned for future webinars. A take care everyone. Thank you.

Thank you everyone have a great day. Please do not try to OCR grafts. It will lead to insanity.