We the People's Collection — The National Collection of U.S. Government Public Information - Transcript

Please stand by for realtime captions.

Hello, everyone. By my clock I have to a clock p.m. Eastern. I want to welcome you to the FDLP Academy webinar , We the People's Collection - The National Collection of U.S. Government Public public information. Ashley Dahlen will be providing tech support and our first support presenter is Cindy Etkin a senior program planning specialist in the office of Superintendent of documents and she began working at the publishing officer as a depository library inspector and held other positions in library services and content management during her almost 24 years at GPO and her current job focus on policy, strategic planning for the public information program with the superintendent of documents and directly supporting the mission of GPO of keeping American form and the public's right to freely access government information on the second presenter today is David Walls . David is the preservation library for the publishing officer he brings more than 25 years of experience in digital and the task of developing and managing strategic initiatives for the preservation of federal government information. He is also a member of the advisory board to the Library of Congress and shares the preservation working group. With that, I'll hand the microphone over to my colleague, Cindy.

Thank you, everyone for attending this afternoon's session and David and I are glad to be here to talk to about the national collection of U.S. government public information and we are calling it we the people's collection and this is a collection that is made up of, as you know, government publications which are developed and produced with taxpayer dollars, so you have all paid for that and it is your collection.

I want to start out with our mission and vision before I talk specifically about the national collection and our mission is to provide readily discoverable and free public access to federal government information now and for future generations. And to provide government information when and where it is needed. Think of that and think about what is needed in order for us to achieve that vision and we have course know that the scope of the content is government information and we know that when, anytime and that will include when the library is open, when the library is closed, where it is needed. Are you sitting in your kitchen at your table and having your morning coffee and want to look up something, or are you going to the library to find some information? We want to be there at your point of need. So when you think about what that means, we need to have our comprehensive collection and it needs to be tangible as well as digital.

And so we have the national collection and it is a geographically dispersed collection of the body of federal government public information accessible to the public at no cost. That is our brief the definition, which is a little bit different than the scope. So the scope is any government information that was meant to be publicly disseminated or created by the work of federal employees or paid for with federal funds.

And we have been hearing some things about what people think the collection is, so we want to tell you what the collection is not. It is not a brand-new collection that GPO is building. It is not a new dark archive or a duplicate library housed at GPO. It is not in one location and it is not replacing any existing collections. So, in fact, the tangible portion is what you all have in your own library that has been sent to you through the depository library program. So tangible components of the national collection and you all play a part, very big part. So why do we need a national collection when we have over 1100

depository libraries providing this information? The short answer is we want to ensure future access. We know that depository collections of both tangible and digital government content are valuable for ensuring an informed citizenry, but we also know that libraries change and sometimes there focuses change and they have new constituencies that they need to serve and some constituencies change. And so what you collect and select from the program changes as well although it is not necessarily a comprehensive collection that all of our libraries have. The public has a right to access all of this information. And the government has an obligation to disseminated and provide the broad public access to this information. And the government has an obligation to preserve its information, regardless of format. A lot of people think that NARA is preserving information but if you look at their website they are only retaining permanently anywhere from 1% to 3% that is sent to them so they retain materials in accordance with retention schedules. So it is not necessarily permanent public access that you get through NARA. So what GPO wants to do with his national collection is to ensure future generations have access to the national assets of that information. And a lot of people think that there is a lot of duplication in our libraries and that many of them are the same and we don't have to worry about things about duplicity and redundancy. And it is very prevalent and in some cases they are right. But we also have so many libraries in our program and so many different times in which these libraries became part of our program and that coupled with a changing scope of what has been distributed over the years through the depository library program and also, heaven forbid, library has had some kind of disaster weather it is a natural disaster of flood or some other kind of disaster in which they have lost parts of their collection, we see how it can start looking at the rudder picture of all of the collections in our program and they are not the same. They're not the same. There are many of the same titles in those, but overall the collections are not the same. So the first depository library was designated in 1814 and by the way the American antiquarian Society is a library in our program today and we have more than 130 libraries that have been designated as depository libraries sometime in the 1800s we have a good number of libraries and old collections and a lot of old materials. We also know that between January 2018 and December 2019 the number of digital only depository libraries increased from 22 to 62. So that is a huge jump in that amount of time and we learned from the 2019 survey that 65% of depository libraries planned to do weeding of their tangible collection and whether that is extensive or selective, nevertheless it is the desire to remove and discard materials from their collections and we know that this was the 2019 biennial survey and looking forward to 2020 we all know what an off year that was not sure if those 65% actually got to do what they planned, but nevertheless, they are planning for the discarding of materials. Which at that rate 65%, that is a huge number, so we are looking at the material, tangible materials dwindling in our depository libraries so it is a little bit of a paradigm shift and the library services and content management business unit were the program and catalog and indexing is is taking a more active role. So the map on the right represents the depository libraries and serving the local needs of your users or their government information and the map above that represents GPO and you see four colored regions and those are the Northeast, South Midwest and West regions that we consider our preservation regions. So we are looking to preserve the content and you all are making sure that the user needs are being met, we have to do a lot of communicating and working together, but we are making sure that material will be there for future generations. We are looking at a broader umbrella over what we have been doing for years and years and years. And this takes a more active role on our part, and we are starting to do a lot of different things that we have not done in the past and we will get to some of those as we get through the presentation today. I think, this particular slide, and some of you may have seen it, it really represents that you are still a very vital point and part of our program in getting what the information needs where you are. We are taking a more active role in the preservation and of ensuring the future.

So what is in the national collection? It is both tangible and digital. As I said the material in your depository library makes up the tangible content and we don't have a library at GPO and we are not building a library at GPO and the material is in your collections. Some of you may have shared housing arrangements so it may go beyond your library. Preservation stewards have made a commitment to take preservation conservation actions on certain titles within your collection and also part of the national collection both tangible and digital art unreported publications and those that are in the scope of our program but for some reason never came through GPO. We are unaware of them. But they are out there and you may all have some of those in your collection and you may have received them as gifts from people cleaning out there offices or as faculty officer who served on a federal committee somewhere or all different ways that you may have some of these publications in your depository collection but we didn't receive them from GPO. Like I said, our program goes back to 1813 the prior to that, if you are in an older depository or older library designation, you may have received publications or agencies who are sending out publications to whoever wanted them and there was no pattern to it and if they had publications they would just send them out and they would like to send them to libraries or historical associations so you may have some of those. And the digital content, there is a lot of it out there as you all know. The majority and I should say the majority but a lot of the digital content in the national collection is in govinfo and harvested things by GPO and it is in partner content and on agency websites and unreported publications. It is everywhere.

So who manages the national collection? There are a lot of different people who have touch points to the national collection, whether it is the catalogers or those doing metadata or graphic records to the collection management and development librarians who identifying content for in just into govinfo and then the digital preservation library and technical service librarians as well. Almost everybody has touch points. And the outreach library and also plays a very important role as they go and visit librarians and come back with information that they have learned from their visits but in depository libraries you also are managing the collection and it is the coordinators and other library staff, you probably have some of the same people that we have identified and the superintendent documents organization in your library. So it is a collective effort for all of us to manage the national collection.

And this is taking on different roles. You are looking at managing your tangible collections and providing access to the digital content and that is GPO and also managing additional content and we are providing guidance to you all for your tangible collections and access. But we are also providing tools and bibliographic content you need and the metadata you need for catalogs and may be pearls you may use and we are all managing the content and the access to it but with different roles.

So where is the national collection content accessible? It is a whole lot of different places and of course from federal depository libraries and from websites and federal agency websites and their libraries and the national libraries and their OPACs and the primary tool that we like to think of as the main point to access the national collection is the catalog of U.S. government publications. And also govinfo has all digital and our web archive and partner archives and I want to point out Congress.gov is also where you can find information from the national collection and they actually have files from GPO to feed into their system of Congress.gov's if you're looking at some links and there that they link off of hearings or bills, you can see our authentication symbol and you can tell they have taken those files from govinfo. We send materials to the digital public library of America, DPLA, so all of the records that we create our there as well. And getting to the national collection and content for many many many places and for those of you who are fortunate enough to subscribe to various commercial vendor services, you know they are available there as well and we don't consider those part of the national collection because it's not accessible. Not accessible to everyone.

We do have some strategies to develop the national collection. Both for tangible and digital content. And we have inserted ourselves into the disposition of materials process for federal depository libraries and relinquishing their designation and leaving the program. And when we look at collections, we are looking at at risk materials and we want to make sure, again we are looking at agencies that no longer exist and those materials are increasingly rare to find pics we want to make sure we can preserve those and we are looking at the condition of materials where physical and intellectual content is threatened and information may be lost in that way but they are the actual physical pieces that are still there. And we are trying to determine some kind of rarity and that is what we are looking for for tangible materials and at the same time we are providing services and preservation services and cataloging collections and David will talk more about those little bit later and from the digital content we do have a collection plan and we do have some priorities for content and the plan has more than just the immediate. And we are getting ready to launch some digitization pilots and David will talk a little bit more about that as well. And those bulleted items in the red box, partnerships and those applied to both tangible and digital content for the national collection and we can't do this alone. We have to do this with partnerships we have another type and number of different partnerships and cataloging is essential. If we don't know we have, then we have a whole lot of other unanswered questions about her collections so cataloging is important and helping you all to get your collections catalogued is also important to us. Again, David will talk more about those, but the other thing is the good steward of taxpayer dollars and this is paid for by the tax dollars. And we want to make sure that we are spending those dollars wisely and we don't want to duplicate unless there is a real need to and for instance we don't want to replicate pub med central me to want to replicate the NASA technical report server in govinfo.'s we do have a partnership and they will provide the access and we catalog materials and make them accessible as part of the national collection.

Talking about the strategic disposition of depository materials and I mentioned putting ourselves into the disposition process. And this depository and guidance document became effective earlier this month, and I want to repeat again that 65% of our libraries were planning to weed materials and the number of libraries self identifying themselves as digital only depositories has really increased. So we really need to think about the disposition of materials in this guidance document, however, only applies to the disposition of materials for libraries program and weeding and establishes a program in which we can absorb tangible publications into depository libraries throughout the FDL P so the material gets back into the network that comprises our shared geographically distributed national collection. And a reminder there will be a webinar about this tomorrow.

The FDLP exchange. I don't know how many of you are using is in many of our depositories are not but it is through this means that it offers an exchange and helps ensure the government documents are housed where they are needed. So provide a needed list and match them up and find homes for these and GPO monitors the FDLP exchange and we are in that level of access to the offers and that last level. We make sure the state and regionals for the libraries discarding materials obviously have first choice and then we come along later and also everybody else. But the whole intent is to get materials where they are needed.

We also have a needs list, and warehousing that and the URL is there so you may want to take a look at that if you're not using that FDLP exchange. And we will be discarding anything we don't need and I mentioned the collection plan for govinfo and here is a link if you want to take a look at that later and get it has are digitization priorities but also it has different collecting levels and what kind of material we are going to in just into govinfo if you're interested in a digitization project and you don't know what you

want to do you make it some ideas and take a look at our plan and see where we are headed and it may give you some ideas for your digitization project and maybe even share your files.

Govinfo has statutory authority in chapter 41 of the U.S. Code and it is a 16363 trustworthy digital repository that was certified in December 2018 and this means that we are able to demonstrate the capability to ensure access security usability and discoverability of its content for the long-term according to industry best practices in our continued to vacation is a key strategic initiative of GPO, and we are now in our recertification process that began earlier this month.'s we do have the certification from 2018 and we had two different surveillance audits that we had to go through which we did pass, and now we start the recertification process all over again. At present GPO is the only institution in the world that has this certification. Also, about govinfo we are a NARA affiliate archive and we have been since 2003 so the digital assets are safe and secure and it is a strategic initiative that they remain so. Will talk a little bit about the catalog of government publications and he mentioned it is the primary finding tool for the national collection and GPO began producing the monthly catalog in 1895 and it was in 2007 that the joint committee on printing or oversight committee allowed us to discontinue printing the paper version of the catalog and allowed us to replace it with the CGP. And all of its predecessor titles will be accessible from govinfo coming very soon. And another update here we have received funding for the next generation of the catalog of government publications, and we have been working on requirements for that so that is moving along quite well so we look forward to something coming along down the line, the next generation, but long-term goals for the catalog of government publications is that it becomes the comprehensive catalog in the national bibliographic graffiti of U.S. documents and that is that it will be the tool that will lead us to the content in the national collection. And we also ultimately would like it to be the union catalog for the holdings in our depository libraries. And we aren't there yet, we are working on it. I know that this is really difficult for you to see. But let me see if I can make it bigger. And now it is little bit blurry. I have pulled up the public papers of the Presidents record, and if you click here on holdings. You can get this record and you can click on whichever ones you want, and you will see that holdings for a preservation steward for the public papers of the President actually show up here in the CGP. So we're putting the holdings of all of our preservation steward's into the CGP as a first step . It will be quite some time before we get all the holdings of all the libraries but we are starting with those that have identified titles for them to preserve and retain. You can see here we have Indiana State Library, San Bernardino law library in the West and we have North Dakota State also in the Midwest and we have Queens Borough Public Library in New York in Northeast and the state law library of Mississippi in the south. So you can see the preservation steward's and we have all four of those preservation regions covered.

And with that said, we do have a lot of challenges. First of all, we have a lot of unknowns. We don't -- we say one a comprehensive collection but we are not sure what that is because we don't know and we can't identify yet all of the titles that it has produced since the beginning and in the formation of our country. So challenge number one. Challenge number two is there seems to be some disagreement about how many government agencies actually are and you'll see a table over there on the right of the number of agencies for different sources, and this was from an article back in 2017. So, you know, I don't know how they identify their different lists and I don't know if they were counting agencies that are no longer in business. And I can't explain the different numbers, but there seems to be obviously some disagreement in that. And we also don't know how much public information is published each year. We used to be able to know that when all of the printing came through GPO and obviously now it doesn't all come to GPO, and some of it is more digital and appears locally on websites. So trying to get a handle on that is a slippery slope. And we also don't know all of the holdings in federal depository libraries and they are difficult to inventory. And we don't know the conditions of those holdings. So

David will talk a little bit about some best practices for this but these are some major challenges that we are facing. And the challenges of the things continue as we look at the landscape in which we work and the environment and collections are evolving and we have some libraries [Indiscernible] but the FDLP itself is transforming and we have a new service model and in September you will learn more about a task force to study the feasibility of an all digital federal depository Library program and this is in response to a recommendation from the depository Library Council from their fall 2020 meeting. Resources can always be a challenge and we have unsure budgets and we have changing staffing levels, and we always have potential legislative changes, whether they are the ones we want or not. So this is the world in which we live in. We do have our challenges and what we are trying to do and sometimes they are often complicated by the environment in which we have to do those things. But we are marching forward. And with that, I will turn this over to David.

Thank you, Cindy. Good afternoon, everyone. I am David Walls , preservation library and at the GPO. And talking a little bit about the preservation perspective of the national collections. The national collection of U.S. government information spans the distance of thousands of miles across the cultural and political history of more than 50 states and four U.S. territories. And the designation dates for these libraries stands more than 200 years of collecting and selecting information products of the federal government for use by the American people. Each one of these libraries selected and collected content unique to their state or their geographic area. And needless to say the preservation challenge for this national collection is huge, to say the least. The photo on the top is the reading room of the American antiquarian Society in Massachusetts and it became a depository library in 1814. And the bottom photograph is the outside of the Lake Pleasant public library in New York. This became a depository library 206 years later.

There are incredible risks to providing permanent access and access for future generations, access for Americans far into the future. The risk to permanent access to the many geographically dispersed parts of this collection are called to develop strategies to mitigate these negative events that impact and our threat to permanent access. Physical decay because things simply get old and we know that paper does get brittle and binding fall apart and pages fall out of old books. We know there are disasters and floods and fire and water creates mold. In these disasters can occur through the failure of mechanical systems in buildings or from outside like storms and phenomena we have little control over. And as Cindy mentioned, we don't even know what is all in the national collection because comprehensive cataloging eludes this program in a way that it becomes very difficult to make two strategic initiatives to determine what we need to preserve and where that material is. We have obsolete media because at the time of the federal government put material on CD-ROM and computer diskettes, nobody ever imagined that these things would fall out of use and become obsolete. And one of the things library seem to be under a tremendous pressure to free up

---

-- Stack space or free up space for other uses and another collections are weeded and there is no problem with that in this long as it is done intelligently and systematically with regard to how many copies of an item may be left and what the rarity of that particular item is. But in my 25 years as a preservation library and, the largest threat to permanent access is human neglect. It's simply believing collections don't matter and all FDLP collections are duplicates so losing one is no big deal and a persistent belief that seems to be out there that everything is actually online and it is not.

Preservation practices exist and they are virtually global and I participate in conferences with people in Asia and Europe and we are all facing a lot of the same problems, and they exist because they are tried and true principles that have been developed over time, and these are the strategies and processes that

are employed to keep collections in the hands of users, and they hold true even for the national collection across the span of the distance among libraries. In the column on the left shows preservation best practices or principles and the strategies and work on the right and the GPO is doing to meet these principles and how we are working to keep the national collection accessible and usable. The basic principle of preservation is to know what you have and we are working on making the CGP more comprehensive and working on inventory projects with some libraries that have completely cataloged collections to bring unreported publications into the CGP and we want to the number of copies and we want to do some bibliographic assessments about the age and condition of copies. And to work with libraries on collection repair and we're forming a task force to look at how content on obsolete formats may either be migrated or we make some other decision about its ultimate disposition. And we do digital imaging of our own through a blanket purchase agreement, and we preserve and distribute that content and make it accessible on the world's only digital repository at govinfo. As we enter hurricane season coming up , there is a number of webinars that I have done in the past that are available through the archive to talk about how to plan for and prevent disasters from storms and things in libraries.

Part of our preservation strategy is to work closely with colleagues within GPO and to provide new collection service models that can add new information to the publications that make up the national collection and provide publications to libraries where they are needed and to be able to evaluate the specific risks that publications face based on age, rarity, and whether and accurate digital copy exists. And in addition to our conversations and collaborations with depository libraries, we are planning pilot projects a depository libraries to extend the reach for services and a better partner with libraries to preserve and provide access to the national collection. And this is a list of some of the current services and tools we either have in place or developing further. Currently, we have about 70 partnerships overall and 44 preservation steward partnerships. This is something that was one of the first things that I developed with folks at GPO to try to offer a collaborative approach to preservation and I was hearing there was a lot of preservation activity going on at a very local level where people were trying to preserve content and they may actually have a preservation program within their library where they may be digitally imaging some government content or they may be in a position that they even have a server of their own or their hosting digital content that's federal for the local university or their user community. So this was a way to kind of give a focus on local preservation efforts and to give them recognition and a more national context because the library content is not all the same.'s of people become preservation stewards, they are helping geographically preserve content if people are being contributors and they are providing digital content that they have digitized themselves for inclusion in govinfo for national access. There is also cataloging and metadata contributors that can work with GPO to work on cataloging publications. We try to keep this very broad will most of the partners are federal depository libraries and when we set this up we got permission to have a very wide perspective on things. So we are interested in all kinds of partnerships. It is better if you are thinking of some way you may partner collaborate with GPO that you reach out to us and that starts a conversation. We have had many conversations that things are in the development stage and in the planning stage. And sometimes people who contact us need to have time to go back and think about it because we have added a different perspective to their thoughts. And sometimes, a period of time may go by before anything really comes together, and that is fine.

We talked about pilot development. So the pilots would be, in addition the partnership's, a way to actually provide services to the libraries using federal funds to be able to accomplish work that met the libraries needs as well as our need to provide access and preserve the national collection. And in developing these things, it really wasn't the case of trying to on our end learn how to catalog it digitally and conduct inventories but more of a case of coming up with metrics we needed to be able to work at

a distance or to put GPO staff or contractors in the field working with the library and what that may actually cost and how some of those processes might be managed.'s we decided to do some pre-pilots as a part of a development phase. And we reached out to a number of libraries locally or groups locally, and the slide I have up at the moment is just sort of the stage of things at this point. We did some cataloging and assessing conditions and we have done some finishing up of things, doing some book repair for some special collections and raw materials for the Department of the Interior library. They reached out to us, and they are a preservation steward for one of their main titles and we were able to work with them. And just to save some federal tax dollars, we reached out to libraries for the pre-pilots that were local where it didn't require a lot of effort to travel to look at collections, and of course all of this is delayed by the pandemic, and we are still working to be a little bit further along on some of this then we would've been like a lot of things over the past year.

But digital deposit is of course a working group and they're working on a number of good initiatives there. And for the pilot projects there was an announcement that went out I want to say a few weeks ago about how you could apply as a federal library for a pilot project. I tried to find the URL that announcement was on and maybe some of the folks providing support can to get up or put it in the chat if they can. What I would suggest though is if you got a great idea for a pilot or questioning just how you may fit in to the pilot process and of those projects, after looking at the information online, we would be happy to talk with you about your ideas and fears and goals or anything you want to achieve or learn about the pilots and we are always happy to reach out and collaborate and have a conversation about it. You can reach me and I would be happy to set up a meeting with folks to chat about the area or pilot you may be interested or applying for and thank you for putting the link in the chat.

So specifically for preserving and providing access to the national collection, how can depository libraries help? Well, you should know your collection. What is it that makes your library unique and special? What is the time for this designation to now, what it is collect and you have any rare or special collection items? You have unique material with local interest that may be no other library outside of your state may have? Do you have a lot of items that seem to be used heavily by your users but they seem to need repair. And you have very high use items that need a digital surrogate or a copy that is not either behind a pay wall or not on govinfo yet? You should know your building and location of some of the items in it and this can be valuable in times of storms and things like that and knowing what threats the building may provide to the collection of the time with leaky pipes and leaky roofs and windows and things. And we can talk about helping to develop disaster prevention and recovery plans. Share what you have digitally imaged. If you have digitally imaged content, it doesn't have to be a whole collection of something. We do have a few content contributors and that is material requested for digital library loan if it is a federal document. They do it according to the specifications that we provide, and they give a copy of that content. And the most important thing throughout all of this process is to discuss and share and collaborate and communicate with us.

Another way that you can help more specifically is due thoughtful weeding. There is no such thing as wholesale collections that just need to be gotten rid of because nobody really wants them. You can use the FDLP exchange and notify us and work with your regional in the process and you can create your own and review the collection development plan to see if there is any content we are really looking for. Talk with us if you want to be a partner or if you are weeding material and having materials you are looking to digitize and let us know if you identify unreported publications and try as hard as you can to catalog your depository collection. If you can't, and you have a lot of interesting material that is on catalog, contact us and see what we can do and overall stay steady and continue to do that excellent job

you are doing in stewarding and providing service to your depository collection, your folks and your area of the world and the digital content. Thank you very much.

Thank you, David and Cindy. We are now ready to take any questions and we have about 10 minutes left until the top of the hour. So does anyone have any questions? And while you may be typing, a friendly reminder that after the webinar is over, you will get a webinar certificate along with a link to the recording and slides. If you're asking a question, please be sure to send it to all participants. David, can you advance the slide for me?

Sure.

Thank you. We are having another webinar tomorrow Thursday, August 26 2:00 two 3:00 Eastern and that is a strategic disposition to support the national collection where we will cover the process for libraries leaving the program based on the new policy that has recently come out so if you're interested in signing up for that webinar, you can still sign up and you can put that in the chat box. And you can still sign up for that webinar for tomorrow. And we also have a satisfaction survey, so if you wouldn't mind filling out that webinar satisfaction survey, that helps us with programming here at GPO. Will send that out one more time.

All right. I am not seeing any questions. I wanted to thank David and Cindy for presenting today. I want to thank all of our attendees for joining us, and we hope to see you tomorrow for our webinar, strategic disposition to support the national collection and that will be at 2:00 p.m. Eastern time. Thank you, everyone.

Thank you.

Thank you to Jaime and Ashley for their support of the webinar and thank you for all of you who attended and again, you know or rework so if you think of something later that you want to ask or share, please get in touch with us. Thank you so much.

Absolutely. Thank you.

[Event concluded]