

Please stand by for realtime captions.

>> Good morning everyone. Where a little past 10:30 and we will get started on the next section, bridging the divide preserving government information in all formats. As you can see by the slide on the screen, this is not -- this is a big program with a lot of beakers. Each one of these folks whose names you see on the slide our partners. Either federal partners or depository library partners who have agreed to partner with GPO to preserve US government information under one of the partner categories. If you go to the new preservation website, on FDLP .gov, you will see the definition of preservation and one of the things that that definition talks about is strategic initiatives. Our partner programs is one of our strategic initiatives of working with our comes to twins, or depository libraries and federal partners to preserve government information and one of the partner categories. Since we've got so many's acres and a lot may be tight for time I will quickly go through each of the categories that you will hear this morning. Under preservation storage, Susanna our partnership coordinator, talking with her a lot about some of the ideas you have about preservation's towards and signing up for preservation stored partnerships. Where going to have Thomas Mills who is the director of Notre Dame law library, Jana Ronan the regional government documents and Winston Claire who is the head of government documents at the University of Hawaii. And a couple of these folks have also agreed to speak to us this morning remotely even though they are not actually with us here in the room. We will talk about digital content contributors, a little bit about the guidance documentation, and hear hear from Chelsea Dunsmore the director of digital services at the University of Florida library. And then finally we will talk about the importance of archiving federal web content, Dori Bauer who is our senior archives nationalist will talk to us but the work we are doing collaborating at the federal level and that GPO and hear from [ Indiscernible ] who is the lead information technologies restless and the web archiving team at the Library of Congress. That's what we have here this morning. We will let each of the speakers talk in succession and then we will questions until the variant. Up next is Notre Dame law school library Thomas Mills.

>> Good morning. Thank you very much. Thank you for the opportunity to week with you. The story of our venture into being a preservation stored actually started before I arrived here at Notre Dame. I started as director in July. Before that, last year our government documents library and left and we reached out to GPO for help. Some training help, and another -- other issues that had arisen and it was decided that someone from GPO would come here to help us with training, and while the person was here, they took a tour of our collection, and noticed that our CFR collection was nearly complete or is she said at the time, the most complete collection that she had been. She told us about the preservation stored program and asked if we would participate. And that was right before I arrived. So when I arrived here, one of the first things that was presented to me was whether or not we would be participating in the preservation stored program, and my immediate response was yes. It was an obvious answer to me to say asked, but then I had to stop and say okay, I have to explain this to our Dean, my new boss, as to why we would be maintaining paper CFR. Our Dean is all about going electronic. And one of my first decisions was to keep some paper. And she had a point. CFR is readily available freely, Internet GPO website. Back to 1996. Legal information Institute at Cornell has it up, Westlaw, if you have a subscription so it is available electronically. So I had decided what what I say to her when she asked me? The first is because of the collection that we do have. It is nearly complete or the most complete one that the person from GPO had seen. And GPO had said that they would help us fill any of the few gaps that we have, so there we can complete the collection. And we have the space. First of all we are not tight for space here at the library. So there was no problem as far as that. In the future as we face some space issues, there is a lot of other material in our library that we can withdraw or remove to storage and if we commit to keep the CFR obviously that will be something we maintain on-site. And be designated as such. This also would allow other libraries to not have to maintain their paper CFR if they know that we are maintaining hours here, and have committed to do so. And this I think feeds into the

importance of collaborative projects amongst libraries. Which is something that I came from Cornell law library here to Notre Dame, that we were very much into at Cornell and saw the importance of collaborative amongst libraries and I felt this is something here at Notre Dame that we can contribute as we start to reach out and collaborate more with libraries, both in our region and across the country. The idea of having the collection, maintaining it, the idea that we could fill the few holes that we have with GPO help, the importance of the collaborative project, but also I think an important point for me was maintaining paper obviously -- these paper record even though electronic is available, but the importance of preserving the paper in this day and age of not knowing the stability or what would be taken down from the Internet, the idea that we do have it here in paper, and it is freely available, also played into my decision-making as to whether or not -- to keep it. I also think that playing into the decision of whether or not to keep it here at Notre Dame was also the idea of -- like I said for the future, continuing to collect it, not have other people not having to maintain the paper or going [ Indiscernible ] in their collection for CFR. Within the GPO program. And these are actually all of the points that I've been raised with my Dean and explaining to her that I decided that we would become a preservation's words, and and commit to keeping the CFR, and it seemed to be a win/win situation for both us here at Notre Dame, to become the preservation stored for something that where it had a complete collection of, and for the whole system and the collaboration amongst librarians and libraries to maintain the record, the public record in paper. So I said in the beginning, it was an easy decision for me and I think for most librarians it is an obvious decision to say yes to collaborate and importance of maintaining CFR and other government documents. I think the tricky part was explaining it to my Dean and educating her so that she also would agree with the decision that I had had. One of the first decisions that I had when I arrived here at Notre Dame.

>> Good morning everybody. Can you hear me? I am Jana Ronan and I'm currently serving as the interim regional government documents library in here at the University of Florida in Gainesville. I'm here to talk about our path to becoming a preservation toward the let me share a little bit about US activities with government documents before get into that. The University of Florida is the regional federal depository library for Florida, Puerto Rico, and the US Virgin Islands. We have approximately 38 collective libraries in the system. This marks 100 years that we have been with the program. We are also an active participant in the Association of Southeastern research libraries collaborative federal depository program with over 35 centers of excellence. US curate 21 USDA USDA agencies within that program, via CSR as Tom was talking about, and the Federal Register. There is also the Panama Canal commission including all other federal documents. Maps about Panama and the canal. For years the US has been committed to making federal documents more accessible. We started with cataloging over 300,000 documents that were stored in our off-site storage facility back in 2008. In 2014 we made the bold move of removing the remainder of the federal documents collection to off-site towards and as we did that we made the commitment to catalog the entire collection. And you can read more about this process and a presentation entitled management of federal documents at the University of Florida that the USD and of libraries Judy Russell and associate Dean of discovery digital services and collection Ben Walker made at the annual library association in 2016. Thanks for pasting that you are the chat message for me. Presently would have approximately 555,000 volumes catalogued. Turning to the topic of preservation stewards, these stewards play a critical row in GPO achieving it's mission of keeping America informed. These days when people are walking around with their eyes glued to mobile devices and all kinds of other technology, it's important to situate yourself at the library where your users are. Cataloging digitizing government documents is important sources out there where users can Google or easily access the content from library searching tools. US was already doing many of the things that a preservation stored as stipulated to do any memorandum of agreement. We been verifying as we go through and catalogued that they are in good condition, restoring the items, controlled access environment permanently. Florida academic repels her which is managed by the University of Florida

libraries. And we are providing conservation to the collection. It makes it sound to collaborate this national program as we have the infrastructure expert staff and established systems to create preservation copies of records. We've also digitized collections to contribute to the effort. Such as the centers of excellence that we are completing. Also committed to preserve Congressional hearing. I will say one concerned that our team had as we explore the possibilities of becoming a preservation stored were some of the requirements that were laid out in the template of the memoranda of agreement. We were wondering about the flexibility of being able to edit some of that. This team I am talking about included Dean Russell who had served previously as a superintendent of documents. Chelsea Dinsmore the director of the library production services who will be speaking later, associate Dean F Walker, the interim department chair of cataloging services, the manager of government documents cataloging and myself. When we reviewed the template it was unclear if there was -- we were able to [ Indiscernible ] where our practice is slightly very. Then our [ Indiscernible ] attended the virtual spring depository consummating where others talked about their experiences and it's there we learned other institutions have had the ability to also [ Indiscernible ] enjoining and we could potentially adapted to accommodate certain local practices. That was really the point at which we decided we really wanted to go with being a stored. Perhaps the largest taking point, US -- at this time. And previously it seemed like membership was expected of a preservation stored. However this wasn't the case and we are able to adapt the M away to reflect this. Cataloging practice was another concern but we were able to incorporate our high-level cataloging into the agreement and [ Indiscernible ] 583 field. Chelsea is going to -- will talk more about this in her presentation later today. We are contributing content to the program and very proud to be a part of the preservation stored program. Thank you for your time.

>> Aloma. I am [ Indiscernible ] the head of the government documents at the University of Hawaii library. We are a regional depository serving Hawaii, Guam, and the Federated States of Micronesia. I would like to share our experience in being a preservation stored. In the picture you see everybody in our department with our former university librarian on the right hand side. When I first heard about the preservation stored program I was very enthusiastic about it but I really wondered how we could participate him up because as many of you probably know, we suffered a disastrous flood in 2004 that destroyed 95% of our depository collection. We didn't have a complete set of anything except I remembered we have united dates reports, it was kept in a separate section of the library so we had a complete run of it. I thought maybe we can do this with United States report. And then we can actually support the preservation toward program by contributing to it and not just be one of the libraries that says what can we throw away. So I contacted the -- first I contacted my library director and got her approval. She was a former depository library and so she was very supportive and then I contacted the government publishing office to find out what the process was. And learned that we had to go through a series of steps. We had to inventory our holding, assess the collection, determine what the condition of the volumes was. We had to make sure that they were were all catalogued and linked to the records and property stamps. And finally, I will talk a little bit more about this in a minute, we had we had to determine which volumes were issued I GPO and which were considered officially government issued. I knew there were some reprints in our set of United dates reports and for anybody who doesn't know what United dates reports is it's the official report of the US Supreme Court. So what did we learn from doing that? A lot of the volumes were not in very good condition as I discovered. You can see a typical example of a book where the pages had become detached from the volume and it had some cellophane tape in it, many of the volumes had insect damage, kept kept in a building that wasn't climate controlled. And then there were a number of the early volumes that were privately published. They weren't issued by the government. So that I contacted Suzanne and asked her are these really part of the depository collection if they were privately published? They did a little bit of research and found that indeed, some of the volumes, the pre-1976 volumes were privately issued and they were not government-funded and so they are not in scope of the FDLP. It was only the 1876 to present that were

actually in scope for the to be covered by the memorandum of agreement. Even the all of the volumes from volume one to the most recent one, only 480 of the volumes were actually covered under the M away. Of those volumes 81 of them were reprints or they were not in good or fine condition and that is the condition they have to be in under the terms of the M away. So I had to create a needs list of those 81 volumes and still seeking most of them. I have not really gotten that much was once but I'm hoping that once the FDLP exchange goes online, I will get be able to get some more of those volumes and if you all know of any libraries that are getting rid of United dates reports, I am all ears. Let me know. One of the other things I learned was that many of our volumes had been acquired much later, maybe in the 1930s, some library and had gone back and acquired many of the older volumes because they had the property stamps of other libraries like Northwestern University Law library. We somehow got a lot of volumes from them and from many other libraries including federal libraries like the Department of Interior. We are our volumes in open stacks but they are not -- they are not circulating, and I also learned that most of the volumes had never been depository stamped so they didn't have any property stamps at all except the ones that had property stamps from the other libraries. But they didn't have our property stamps, they weren't depository stamped, they were Kellogg. This is a good thing because it resulted in this whole set of volumes being catalogued and properties and so forth, which is of course what we were supposed to have done with them decades ago anyway.

>> That is the very short story of something that took a really long time to do. Thank you.

>> Chelsea, if you are muted, please unmute your phone on the conference line.

>> While we're waiting to sort out that technical difficulty I will say that if you look on the preservation with [webpage@FDLP.gov](mailto:webpage@FDLP.gov) at [FDLP.gov](http://FDLP.gov) on the preservation tab you will see new guidance documentation on preparing actual content packages for [ Indiscernible ]. We had a document up previously and we got good feedback from you all about how to digitize for us. If you wanted to digitize and have content and follow the packaging guidelines.

>> This is Chelsea, I I am back on the phone. I apologize for that. I ship to phone and somehow it hung up. Thank you, good morning and thank you for inviting me to join you today. The University of Florida library became a regional depository library in 1907. And retention of federal depository holdings has been our responsibility sense that became a formal requirement of the program. The US signed a preservation stored memorandum of agreement in June of this year with GPO, and as as Janet mentioned earlier, signing up as a preservation stored was -- as we take our commitment to retaining federal documents very seriously. US was an early participant in the Association of southeastern region library since April. Not to be confused with a CRL. Collaborative plans for managing [ Indiscernible ] collections in the Southeast around 2010. Was exploring ways to enhance cooperative training outreach service and collection analysis and development activities to improve access to federal government information for the citizens of the region. The plan calls for the creation of centers of excellence or seal we, at depository libraries around the Southeast. Each library would select agencies to focus represent active collecting agencies on and commit to keeping those materials in perpetuity ensuring comprehensive collection and retention of FDLC content across the region while reducing the overall collecting efforts at each institution. The program envisions the creation of comprehensive collections of US government information from each federal government agency. Built upon the foundation of existing holdings at existing regionals, these collections become more complete with assistance from -- anytime [ Indiscernible ] states. The program now includes 11 regional depository libraries, and 30 selected depository libraries. Nine of these libraries participating members. The program visioned at least two centers of excellence for each agency to ensure a minimum level of redundancy within the Southeast for both quick level of delivery and presentation. Centers of excellence are asked to actively replace damaged or lost pieces, seek to fill holes their collection as necessary, provide active preservation and permanent retention for the collection, catalogued the pieces that they own and identify them as an archival copy as well is participating in GPO program to set holdings on [ Indiscernible ] for regionals free

of charge. So as part of our seal we program, US undertook to digitize our holdings as well as to collect and retain the print edition. This row had several benefits including reducing loss or damage and improving overall access be on the patients who could visit our library.

>> Lest you think this is an excuse to advertise the CLE program, what this meant is that when we selected the Panama Canal as a vessel use case for the CLE program, we found ourselves in an excellent position to then participate in the preservation stored program. The Panama Canal project was a special use case. Developed around a topic rather than his civic agency. So this meant we had a set of materials that was quite varied. The collection spans the gamut of maps, reports, hearings, cultural research, and even a few pieces of China though we are not sure about the government status on those. Our goal had always been to share these images once we digitize them with GPO Internet archives and other trusted repositories as a means of furthering their extensibility as well as for preservation site. This meant that we an extra time I quality control ensuring that our scans were accurate replications of the original. This put us in a possibly unique position when the opportunity arose to become preservation stewards. We had a set of inventoried and condition assessed materials that were completely catalogued and scanned to standards and carefully QC that we could provide to the GPO system with very little delay. When it came to contributing our content we found it to be a risible process. Our system already had titles and individual folders and we had [ Indiscernible ] of our images and we use XML to hold our metadata. We did have to create PDFs in some cases but that supported our broader efforts to extend access. So it was time will spent. The set of files for each document was packed into a folder and we used FTP an electronic folder of course and we used FTP [ Indiscernible ]. The first round of documents actually went on a hard drive but the more recent batch went electronically and we anticipate that being the process moving forward. And we sent to batches so far totaling about 800 items. In exchange for sending our files to GPO we gained the added reassurance of having another copy held at an external repository. GP also provided back copies of the PDF files modified to include a stamp of offenses the. Authenticity.

>> You can see a picture of the stamp your. It's actually very small on the page. Sometimes you really have to look for it but I find it cute. Putting the files also places the digital images in a more accessible place for FDLP librarians and users who are familiar with the system. They don't have to go to another portal to find these files. GPO also reviewed the catalog records which provides additional confidence in the content and finally we authenticated -- can be used as surrogates for the material providing another layer for researchers. We have now turned our attention to the the other agency collections that we have on hand to begin prioritizing that which will contribute to the project. Thank you for your time.

>> I am Dori Bauer and I am an archive specialist at the government publishing office and one of my responsibilities is that I manage our FDLP web archive. I thought I would put people on the spot in the beginning here and just ask, I'm curious because I have presented a few times before this conference, just in general how many people are familiar with our FDLP web archive? I never know what sort of level of introduction I should give. I do know that we have a few new attendees this year. But naturally, a repetitive theme around our work is that since the 1990s much of the content that is being disseminated is now in electronic format and no longer in Prince. And in 1998, that eight, that was when GPO first began manually harvesting electronic government publications and using [ Indiscernible ] for permanent access. Since this time naturally agencies are starting to more and more disseminate their information through their own website, and they are not going through the FDLP or informing the superintendent of documents as stated in title 44 about when they are posting information to their website or taking it down. Done this causes a problem for us in that how do we continue our mission of keeping America informed? Especially considering the amount of content that is now being posted directly on to agency websites, and the seminal nature of it being posted and taken down without warning. It was in 2011 that we got a good response to this, we would start doing website level web archiving. And we became an archive partner that. And it was important to us of course to get the traditional publication, PDF and information like that but this website developing so much, so much

interactive content, video, what what is really a publication and what we often discuss, so it was important for us to get the look and feel and all of the interactive nature of a website and I just have an example here of one of the websites that we have been doing from the beginning, US Holocaust Memorial Museum and this was one of our first calls in July of 2012. And you can just see over time here is May 2014, and this is just their main webpage. And our last call we did was March 2017, but the website is one that changes regularly. And they are constantly adding new content and you can just see that the website is constantly changing and updating and adding new information and new ways of navigating to the website. It's important for us to get the whole piece here.

>> When we started doing website level web archiving, what were were our priorities for collection development? It must be within scope of the FDLP, that it is generally not distributed through prints, that it is information being disseminated on agency websites and content that is not being catalogued. We do try to avoid duplication of effort with some content that we know that other institutions might be concentrating on. And then we worked closely with our acquisition staff to help us determine our needs in our early collection development. To date what we have done is we first started with doing the class of the classification scheme which is the commission committee and independent agencies. And when we completed that, we gathered a list through our acquisitions department of what we call nonstandard government websites. For example, choose my plate that love -- choose my plate .gov. We also did special collections in the first one we did was for resources for Native American resources on the web. We also get notified through the FDLP community about content that they might like to see that you all might like to see a part of this collection. And also I wanted to mention maintaining regular frequent the and this is something that is very resource and labor-intensive for us. We do not just set a crawler to automatically crawl a site every six months. We actually physically go in and re-examine the site, so it does take a lot of time to go through this, and then also doing all of the while we are maintaining the size of our account we archive it. We all know that the size of the government web is huge, and we had originally started doing that three terabyte a year, we are up about five terabytes a year for our account and that really is not a lie considering how much is out there on the government web. We will have to monitor what we are doing and how much base we are taking, and it causes us to have to make a lot of decisions about maximizing the use of our count. I mentioned that notification through the community. We get notified through a number of means. Document discovery, which is a way which agencies contact us. Through ask GPO we get information about websites to crawl. There is a reporting form and then also our staff are always monitoring listserv's. Moving forward from here, where working with our new collection development library in and we have been working on a new set of collection development parameters for adding new content and we will hopefully be doing that soon. We initially started doing a lot of smaller sites and we have been moving into doing some larger sites now. The special collections I mentioned we're doing more with that, thematic collections, we have been working more doing some healthcare related websites, and of course now the hot topic has been environmental and energy websites. Also, we do extensive monitoring using Google analytics of our use statistics. And monitoring who is looking at what website and for how long, and we see this as a means for understanding our community and what you all are interest did in doing and being in that can help us with our collection development needs as well.

>> I also include here a link to myself and my contractor did recently a webinar on Google analytics so it gets into more detail about what we are doing with Google analytics. And so I have provided the link there. To learn more about our Google analytics. The collection size now, we are at 12.2 TB with 99 over 99 -- over 99 million URLs crawled. 141 collections and 181 records in our CGP. Added here also a link to our project page and this has an extensive question and answer section. To tell you more about our collection. And also I wanted to mention that Ashley Dahlen who I see out of the corner of my eye one of our outreach librarians did a webinar as part of the FDLP Academy on using web archives in reference work and she did an excellent job, so I encourage everybody to go look at that. I have provided a link to

that webinar there as well. So there are two locations for access to these collections through the archive it page you can get to it and then through the CGP. Also wanted to mention that a year ago I was very excited that we have our own catalog no for the FDLP web archive in CGP's you can get to that right from the main page and then you can just search within the FDLP web archive or there is a link that you can get to all of our records from there. I also wanted to mention then and preservation of web resources and mentioning the huge size of the government web, and what can we do, that collaboration is very important. And again I mentioned with being contacted by the community, by the librarians and with concerns about certain web content, and with these nominations that came in, there was just always questions about what are other agencies doing and who is doing what? And again I have mentioned that we would like to prioritize what we are doing to maximize the use of our account and archive it, and not do websites that may be other agencies are working on or concentrating on. So at that time a couple years ago we began talks with the Library of Congress who organized our inaugural meeting of what we called the federal web archiving working group. And the initial group was the Library of Congress, GPO, and the national archive. And in our initial meeting we got management approval to move on with this group. Once we had our approval, why not? We started then initially our early meetings were detailed presentations letting each other no who is doing what and how they are doing it. The tools people are using and how they are tracking and information like that. We started doing outreach and I have included a link to our first blog that we had, Library of Congress digital preservation blog, and also sense this time we have been working towards increasing participation in the group, and we currently have participants from Smithsonian, national library of medicine, Health and Human Services and Department of Education. Some topics that we discuss, just to let you know some of the things we talk about, and, and some things that we have accomplished, we meet on a bimonthly basis at this point. And we are giving updates for to each other about what is going on with your program, problems that we have encountered over time, over the last two months, and development and keeping each other well-informed of what is going on. We had developed an internal wiki where we could post information about our programs and post list and things like that. We talk a lot about it ends up about developing contracts and RFI and RFQ. And of course some challenges that we have with that. That tends to sometimes take up a lot of our discussion. We have discussed processes for transfer of work files and shared with each other how we are doing this. And offering advice to each other on this. We talked about policy topics, and we have shared internal policy documents with each other and technical topics and the one that is always at the forefront is how do you archive social media? And it is constantly changing. We are constantly having new challenges with us. Can't really having to share what are the problems you are having with states both know and how are you meeting these challenges?

>> Where we want to go from here. More outreach is very important to us in increasing the participation in this group and the knowledge of this group outside of it. And who is doing what and making sure the public is aware of this. We also are looking to build more relationship with content creators. And along with this, creating guidance documentation. Whether it be on just web archiving guidance or also for content creators and how can you design a website that is archive from the? These are a couple guidance documents we are looking into doing. And then a web presence for the group and having a means for us to be able to distribute this information that we are gathering together or guidance documentation we might be writing. We would like to find a means for getting this out to the public. With that, I will introduce one of our partners, and we [ Indiscernible ] with the Library of Congress.

>> Good morning everybody. I thought I would only have two minutes left but apparently I get a lot of time so good. I won't use all that. I will talk a little bit about what the Library of Congress is doing around web archiving and the federal government web archiving and talk a little bit about our end of term project which I think we have spoken by that this conference but it would be good to give you an update

on where we are with that. We have a lot of content in this a little bit out of date. We ran our numbers and I think we're up to 1.3 terabytes or maybe for after the last fiscal year. So I like to point out that when YouTube came in around 2006, that's when we really started [ Indiscernible ]. You can see our skill is a bit bigger than the GPO numbers and we are collecting a lot of content, about 25 TB terabytes a month at this point. We are outside of the scale of Internet archives that we do our own in-house we contract out the crawling to Internet and do some in-house crawling. I won't go into detail about this but we can share a slight later. We have a small team that manages the program and a lot of people around developing collections and helping make it all happen. I lead the technical team that manages the general project overall. Our collections are pretty broad. We don't just cover government information but we do a lot of it. Some of our you can see her on the chart we do various events in the US like our US elections. A lot of somatic archives which are where our government and law collections are. We have a lot of international collections because we have offices overseas and other divisions of the library that are selecting government and other materials and other countries. We have a real mix of content that we are managing and different types of collections. In terms of the government collections we have we really focus on a couple main things. We're comprehensively collecting the legislative branch and we now have a public archive. Of the legislative branch and we also collect we are collecting the congressional websites monthly including committee site. We just recently launched a federal courts web archive so that as of now. And we do selective crawling of executive branch. We are not doing it all, we're not comprehensively archiving the government web but we have -- if there are questions about selection decisions she can jump up and help answer those. But we are we have a lot on our plate but we are not doing all of the government web. We do have a permissions -based approach for most of our web archiving. We collect those without permission. Everything else falls under this permissions -based approach where we are notifying sites and trying to get permission for various aspects of the work. And I wanted to point out on the bottom there there is a link to our collection development policies. This was recently updated. I think it was posted a few weeks ago. The latest update to the guidance that the recommending officers in the library use to help the web archive collections.

>> Our collections are publicly accessible. From the homepage you can search the web. Get to our web archives. About our third a third of our collections that we've done over time available publicly among. We've got a bit of a backlog and getting collections out. We made some progress this year and launching things like the legislative branch web archive, and the federal courts. We also did some comics and some of our other web cultures, web archives, and we are getting a little bit more attention. We've got some press on the LSE homepage which we were excited about to promote the use of our web archive. Hoping to get more collections out this year. We do a faceted search and then there is no URL search that is easily found. I will share that link in a moment. But each website is catalogued, and described. We are short on catalogers to do this work so we are doing more automated creation of records to get content out faster. But it is very labor-intensive as you know. This is not a government website. I wish it could be. Dinosaur comics is an example from web archive but you can see that it's got a catalog record there and then we've got our own version running to access the content itself. Where also a one-year embargo. If you're looking for anything, even government websites that we didn't seek permission for Wilma have it up to about a year ago. Here is an example of our website. We've got the banner at the top there. We do have a URL search. It is very bit. We're hoping in the next year or two elevators and refresh our public presence about the archives. Things that are not catalog are available through here. Even if we haven't made the collections available as a collection on our website you can get to the content. Where also working with others around the library and hope to focus on this in the coming months. And your about -- the lake of data that we have, over a petabyte of data. So much to dig through. So little resources to describe it so how do we make that available for researchers is really on our mind. I don't know if you have seen the labs which is an effort to buy our colleagues in the national digital initiative part of the library that is allowing people to access data sets and we really want to get in

on that. I think we are next on the list in terms of getting doing some experimenting. We participated in a pilot. Where they use some of our crawl data and a lot of researchers using web archives or looking at how they can use it in bulk and how they do analysis and ways that aren't through public website and not clicking around the website like the [ Indiscernible ] we have now. We're interested in doing this and being able to provide collections of government data for instance to researchers who might be interest did in just a content. It is something on our minds. We hope to have some success with that. And since we are talking about collaborations I wanted to talk a little bit about the types of things we do. We are doing a lot at the Library of Congress but we are definitely not able to do it all and we rely very heavily on partners in the federal government web archiving group and the international Internet preservation Consortium. And then things like the NDS a survey that we participated in the documents who is doing what web archiving around the US. If you are doing web archiving the survey is open right now so we encourage you to take that. But I wanted to mention specifically the end of term web archive which is a collaborative effort. Have people in the room heard about end of term? This is a project to document the changes in the government web at the end of a presidential term. And we didn't just start it this year based on news about the changes that were happening. We began in 2008, repeated the effort in 2012 and then again last year and the varying -- the partners have varied over the years GPO Library of Congress, University of North Texas, Internet archive, people at Stanford joined this year. George Washington University. There are a number of us that have gathered together like a band reunion. Every couple of years to figure out what is the government web? Which is a big challenge in itself and then try to preserve it. And the amount of data is just increasing incredibly. Into thousand eight we didn't have that much we try to do our reach elements as a part of this and we collected 25 terabytes of content. And the increased in 2012 last year we the partners in entirety collected 250 terabytes of data. 100 of that was have to be content that was a project of the Internet archive that part of the effort is to distribute the preservation copies to all the various members, so the Library of Congress for that is going to have to transfer 150 terabytes from our partners to the library this year to make it part of our collection. It's a huge effort. We had about 11,000 nominations from the public and government documents experts and there were -- the data refuge project that was also doing a lot of preservation activities over the last couple of months around that time contributed number of URLs to the project. Just an amazing amount of collaboration, and interest, we got a lot of press this year and James and Internet archive and some of the other leaders in the project did a lot of talking to the press about web archiving which is great to see that it was in the news, and perhaps -- I was little surprised. In the last year we didn't get that much attention. It's a really great effort and Internet archive has made that data available on their website and we are working on up dating -- we have a portal. California digital library. I know I left them out. There hosting their front and to the collection and we will be updating met soon with the latest results. A lot of information. On our website. And I will hand it over to [ Indiscernible ]. Feel free to contact me and I can share other information.

>> We have heard bit about preservation's words, we have heard from digital content contributors and we heard about the FDLP web archives. And our collaboration with the Library of Congress. Any federal web archive working group. Do we have any questions for any of the speakers? Even the ones remotely this morning?

>> [ Indiscernible - low volume ]

>> Thank you.

>> I was wondering about the Cytex agency. The [ Indiscernible ] people with EPA and Defense Department and so on. Is Elsie working with those players in GPO?

>> I know that we have briefed [ Indiscernible ] and some of those agencies during the end of term project time. And we also reached out to things like the federal webmasters and other groups, some some of the data rescue events that were held I think brought in some of those agency partners.

>> GPO is also working with [ Indiscernible ]. The data cure should group and others, we participating with them as well. Any other questions? No other questions. All right. Hopefully we have given you lots to think about. If you have follow-up questions, the session is recorded and you can also contact any of the speakers directly. Suzanne and I will be around the rest of the week as you see us and have questions about our partnership, then please do ask. Thanks very much for your attention. [ Applause ] [ Event Concluded ]