Stenographic Transcript
Before the

COMMITTEE ON
RULES AND ADMINISTRATION

UNITED STATES SENATE

THE USE OF ARTIFICIAL INTELLIGENCE AT THE LIBRARY OF CONGRESS, GOVERNMENT PUBLISHING OFFICE, AND SMITHSONIAN INSTITUTION

Wednesday, January 24, 2024

Washington, D.C.

ALDERSON COURT REPORTING
1029 VERMONT AVE, NW
10TH FLOOR
WASHINGTON, DC 20005
(202) 289-2260
www.aldersonreporting.com

THE USE OF ARTIFICIAL INTELLIGENCE AT THE LIBRARY OF CONGRESS, GOVERNMENT PUBLISHING OFFICE, AND SMITHSONIAN INSTITUTION

Wednesday, January 24, 2024

U.S. Senate
Committee on Rules and Administration
Washington, D.C.

The committee met, pursuant to notice, at 2:15 p.m., in Room 301, Russell Senate Office building, Hon. Amy Klobuchar, chairman of the committee, presiding.

Present: Senators Klobuchar [presiding], Padilla, Ossoff, Fischer, Capito, Hagerty, and Britt.

OPENING STATEMENT OF SENATOR KLOBUCHAR

The Chairman. Calling this hearing to order. Good afternoon. I would like to thank Ranking Member Fischer, as always, and our colleagues for helping us put together this hearing. With us today is Dr. Carla Hayden, who is the Librarian of Congress. Thank you. We always enjoy having you here.

Also here is Hugh Halpern, the Director of the Government Publishing Office. Back again, thank you. As well as Meroe Park, thank you for being here, the Deputy Secretary and Chief Operating Officer of the Smithsonian. We understand that the Secretary is ill. Just -- he will recover but wasn't able to make it, so we really appreciate it.

Today, we are going to talk about the very important topic of artificial intelligence and the agencies that play such a critical role in serving the American people. AI has the potential, as we know, to lead to incredible innovation by supercharging scientific research, improving access to information, and increasing productivity.

But like any emerging technology, AI comes with significant risks and our laws need to be as sophisticated, as the potential threats are there to our own democracy. Understanding these risks and benefits has been a major bipartisan focus of the Senate, with Senators Schumer,

Rounds, Young, and Heinrich leading a series of nine forums since the fall, and a number of us on various committees working on proposals so that we are going to be ready when this hits.

And I think, as you all saw from the reports of the robocalls in New Hampshire with the fake voice of the President or other things that have also occurred to candidates on the Republican side of the aisle, this is not going to be one side or another.

It is something that we as a Congress have to deal with to put some guardrails. And that includes, of course, the work that goes on in these agencies. At our hearing in September, all the witnesses agree that AI poses risks to our elections, and we heard testimony on why we must work to put guardrails in place.

That is why I am leading a bipartisan bill with Senators Hawley, Coons, Collins, and also joined with Senators Bennet and Ricketts, to prohibit fraudulent AI generated content in our elections within the framework of the Constitution, for instance, allowing for satire and the like, and why we need to take other steps like disclaimers on ads that use AI so that the citizens of this country can actually believe that it is their own candidate or their opposing candidate who is speaking. Another example of a legislation going on.

Senator Thune and I have joined together in introducing a bill to put in place commonsense safeguards for the highest risk, non-defense applications of AI, and improve transparency. I see that Senator Capito is here. She has also joined us on this important bill that is mostly coming out of the Commerce committee.

So, let's get to your stuff, AI and how it affects the work of the three agencies before us, the Library of Congress, Government Publishing Office, and the Smithsonian. While it is important that our three witnesses today speak to measures they are taking to safeguard against potential harms, they are also, I think it is important to note, using AI technology in their work to protect our country's greatest treasures, advance scientific research, and improve public access to information.

For example, the Library of Congress is testing emerging AI technology to expand how researchers can better use the resources they already house in their collections, which make up the largest library in the world, we note with much humbleness, such as a new AI tool that lets users instantly search through 1.56 million images from digitized historical newspapers to assist in archival research.

GPO is working to harness the efficiencies offered by AI to modernize how it makes information from all three

branches of Government more usable for the public since much of its work has expanded to digital publishing, hence their new name, printing to publishing.

And as part of its work producing Government documents, GPO is using AI to ensure quality control of items such as the material you use to print passports. It printed 22 million of them, as we learned at the last hearing last year.

Finally, at the Smithsonian, which is the world's largest museum education and research complex, researchers are exploring how to use AI to do things like tackle some of the most challenging problems in astrophysics, classify species of fish in the Amazon, and make collections more accessible, accurately identifying the contributions of women, I like this one, in historical texts, in which they were often identified in writing by only their husbands' names. That will be interesting what you discover with AI.

We must continue working to stay ahead of the curve, and I am committed to working in a bipartisan way with Senator Fischer so that our country can benefit, and your agencies can benefit from the best of AI, while protecting against any threats. Thank you again, and I will turn it over to Senator Fischer.

STATEMENT OF SENATOR FISCHER

Senator Fischer. Thank you, Chairman Klobuchar, for calling this hearing on the use of artificial intelligence at the Library of Congress, the Government Publishing Office, and the Smithsonian. I also want to thank our three witnesses who are here with us today.

I know we are all wishing Secretary Ben a speedy recovery. I believe this is the first time the committee has heard from the Library of Congress, the Government Publishing Office, and the Smithsonian at the same time, and Director Halpern, it is nice to see you back with us so soon.

The Library of Congress and the Smithsonian both safeguard a vast collection of our nation's treasures, welcome our constituents to their beautiful buildings to learn and explore, and service crucial resources to the scientific and the academic communities.

The Government Publishing Office also performs vital functions for the American people as it produces, distributes, preserves, and publishes documents for all three branches of our Government.

This is the Rules committee's second hearing on the use of AI. As we discussed at the first hearing, there is no question that AI is transformative, and it is poised to evolve rapidly. While AI brings the possibility of

creating efficiencies and competitive advantages across Government, it also creates risks.

Understanding and weighing the benefits and risks of AI are necessary first steps before widely adopting the use of AI. Today, I look forward to hearing from our witnesses about how our Legislative Branch agencies are exploring potential applications of AI.

There are many important questions to ask, such as are these existing issues or specific challenges at our Legislative Branch agencies that AI could solve? Are there existing services that our agencies provide that AI could improve? Can AI provide new avenues for these agencies so that you can fulfill your missions?

I also look forward to hearing about the guardrails our agencies are putting in place to mitigate those risks of AI. We need to understand the cost of AI and AI's impacts on agencies' budgets.

Just as importantly, agencies need to understand the implications of AI use, especially the risks related to privacy. Understanding Legislative Branch agencies' use of AI is an important oversight question for this committee. That said, we cannot lose sight of existing challenges facing the agencies that we oversee.

As the Library looks to hire a new director for the Congressional Research Service, and GPO seeks to maintain

high quality staffing standards in the face of your retiring workforce, it is important that this committee supports their efforts to improve hiring and retention efforts.

Furthermore, as the Smithsonian endeavors to establish two new museums authorized by statute, we must maintain rigorous oversight of their efforts to tackle a deferred maintenance backlog, a chronic issue for that agency.

With that said, I look forward to a productive discussion about our Legislative Branch agencies' use of AI, as well as other issues that you face. Thank you, Madam Chair.

The Chairman. Well, thank you so much, Senator Fischer. So, I will introduce our witnesses. Our first witness, as I noted, is Librarian of Congress, Dr. Carla Hayden.

Dr. Hayden was sworn in in September 2016. She has done a tremendous job leading the Library. She previously served as the CEO of Enoch Pratt Free Library in Baltimore, and she received her undergraduate degree from Roosevelt University, and master’s and Ph.D. from the University of Chicago, that I also attended.

Whenever I see Dr. Hayden, I always think about Barbara Mikulski, who was so proud that a Baltimore librarian got this job, and you have made her proud too.

So, thank you. Our second witness is Director of the Government Publishing Office, Hugh Halpern, who has been in his position since his unanimous confirmation, it doesn't happen all the time, by the Senate in December 2019.

He previously worked on Capitol Hill for more than 30 years, including for former Speaker Paul Ryan. And he received his bachelor's and master's degree from American University, and his law degree from George Mason. Thank you, and we really enjoyed your attendance at your last hearing.

Our final witness is Meroe Park, who became Deputy Secretary and Chief Operating Officer of the Smithsonian in December 2019. Previously, she worked for 27 years at the CIA. Okay, I didn't know that. When I read these bios, I learn new things all the time. You probably don't advertise that, but that is interesting. Including as executive director and chief operating officer, and she received her undergraduate degree from Georgetown.

Okay, now we are going to swear in the witnesses. Okay, if you raise your right hand, do you swear that the testimony you will give before the committee shall be the truth, the whole truth, and nothing but the truth, so help you God?

Dr. Hayden. I do.

Mr. Halpern. I do.

Ms. Park. I do.

The Chairman. Thank you, and you can be seated. We will now proceed to five minute testimonies. We will start with Dr. Hayden.

STATEMENT OF HON. CARLA HAYDEN, LIBRARIAN OF CONGRESS, LIBRARY OF CONGRESS, WASHINGTON, D.C.

Dr. Hayden. Thank you, Madam Chairwoman, Ranking Member Fischer, and members of the committee. Thank you for this invitation to appear before you to discuss a topic that is of great interest across the Government and within the public archives sector, artificial intelligence.

And I am pleased to have the opportunity to further engage with my Legislative Branch colleagues on this topic and to update the committee on the Library's activities exploring AI. At the Library, there are four areas of focus with AI, to expand access to our collections, to enhance services for users, to improve internal processes for increased efficiency, and to implement a slate of strong governance for the use of AI.

Becoming a more digitally enabled agency has been a key focus of my tenure as Librarian, and in 2019, the Library published a comprehensive digital strategy to guide the agency's use of technology in an increasingly digital world.

Building on this first major step, we have now fully integrated our digital strategy into our 2024-2028 strategic plan, and the plan embraces a central idea, that technology must be, as our CIO, Judith Conklin says, baked into all we do. Since 2018, the Library's Digital

Innovation Division, also affectionately known as LC Labs, has investigated AI and shared the results of its research experiments with the public.

Working with digital researchers we call innovators and residents, LC Labs has made its mark as a launchpad for innovating engaging uses of AI to expand access to our collections. A very popular example is citizen DJ, a music sampling application that allows users to remix and create music using free to use non rights restricted audio from the Library's collections.

Other exciting AI use cases include experimenting with machine learning and optical character recognition, or OCR, to help manage metadata and machine readable text for digitized documents.

For example, OCR has increased the discoverability of more than 20 million historic American newspaper pages through the Chronicling America Project. And building on that technology, users can also search those historic newspapers for photos using an application LC Labs rolled out in 2020 called Newspaper Navigator.

However, enhancing services to Congress and the public is also a major area of focus. In 2022, we released a Government, a congress.gov API, application programing interface, to make it easier for the public to access and use accurate, structured Congressional, Legislative data.

And additionally, several Library service units are successfully demonstrating AI's use in bolstering the Library's information services. Our digital innovators have been working with the Copyright Office to test approaches for extracting data from historical copyright records.

Now, this project combines human skills with AI capacity to make a written or handwritten analog record more accessible and easier to search online, and it is just one example of how humans in the loop can be the model for AI and successful integration of technology with human skill, because a very important principle is that well trained human beings will always be critical to the work of the Library of Congress.

Now, the Library's AI use cases have also been a testing opportunity to increase efficiency and staff productivity. For example, our Cataloging Division is currently experimenting with AI to help staff process more efficiently bibliographic information and catalog records.

There are many opportunities with AI, but there also must be, must be robust governance. The Library is in a strong position as our existing technology governance policies provide an adaptable foundation to use in guiding us with emerging technologies like AI. And further, the Library's approach to implementing AI closely aligns with

the practices of other Federal institutions and is informed by NIST's AI Risk Management Framework, which also aims to improve the trustworthiness of AI applications.

Responding to this fast developing area of technology calls for collaboration across the private sector and Government, which is why the Library participates in the General Services Administration AI Community of Practice and is a leading member of the International Artificial Intelligence for Libraries, Archives and Museums Secretariat.

To conclude, discovering the role AI has to play in enhancing services to Congress and other users remains an ongoing effort. And as the Library of Congress charts this course forward, we plan to draw on our history of technological innovations, and our rigorous development of standards, and the input of our stakeholders and partners to align possible uses for the public.

We appreciate this opportunity, and we hope that you will join with us in making sure the values of transparency, accountability, and efficacy remain as we explore this opportunity.

[The prepared statement of Dr. Hayden follows.]

The Chairman. Thank you. Mr. Halpern.

STATEMENT OF HON. HUGH NATHANIAL HALPERN, DIRECTOR, U.S. GOVERNMENT PUBLISHING OFFICE, WASHINGTON, D.C.

Mr. Halpern. Thank you. Good afternoon, Chairwoman Klobuchar, Ranking Member Fischer, Senator Capito. Always good to see you. I am pleased to appear before you today to share some of the potential uses of AI and related technologies at the Government Publishing Office.

GPO differs from the Library and the Smithsonian in that it is fundamentally a manufacturing operation. We publish, produce, and maintain materials for all three branches of Government.

Our 1,600 craftspeople and professionals produce virtually all of Congress's documents, along with numerous other publications, and manufacture secure credential products like the U.S. passport. We also provide digital information either through our own trusted digital repository, govinfo.gov, or by serving data to our partners like the Library where they use that data on sites like congress.gov.

No matter what you call it, artificial intelligence, machine learning, or a large language model, GPO's operations are just as susceptible to disruption as any commercial firms, and that is not necessarily a bad thing.

My written testimony describes GPO's policy approach to this new generation of tools, so my statement today will

focus on three potential applications for AI and related technologies in our day to day operations.

First, we believe these tools can improve our quality assurance process by automatically recognizing defects that a human inspector might miss. We already use a rudimentary form of this technology in the production of the current version of the U.S. passport. GPO uses equipment that optically scans the pages that will become the identity page in a personalized passport.

This equipment looks at each strip of three pages for variances that exceed the specifications for the material and rejects those pages that don't conform to the standard. AI technology has the potential to further refine this review, allowing the machines to learn what may constitute a natural variation that is within the specification and what is not.

This has the potential to reduce defect rates, lower waste, and free up our quality assurance team to focus on solving bigger quality problems as they arise. Second, we see tremendous potential for supplementing our proofreading team.

Proofreaders are very difficult to hire, and we need to free them from making routine, repetitive corrections and allow them to focus on more subtle issues that really require a human being to interpret. One example is

capitalization.

GPO style says that we capitalize the letter S in the word State when we are referring to a political subdivision of the United States. Currently, we use computer scripts to perform global search and replace functions on documents to correct our most common errors, including that one. But those scripts are really blunt instruments.

For instance, they can't tell the difference between the State of Minnesota and a New York state of mind. AI holds the promise of tools that understand context and know when text refers to one kind of state or the other.

That will cut down on the need for our proofreaders to review and rereview, and then correct material that has already been run through our automated tools, and free them up to focus on more difficult contextual issues.

My final example comes from GPO's public information mission. We have had great success in making Congressionally mandated reports publicly available since the end of last year, with more than 180 of those reports available on govinfo as of yesterday.

Most of those reports come to us as PDFs. While that is a good format to show how the printed document looks, it is not always the best format for viewing on a phone, a tablet, or even for folks with disabilities.

While GPO would like to get these reports in more

flexible formats like XML, the agencies' authoring reports aren't always equipped to supply them that way. AI technologies hold the promise of allowing us to extract the information from a PDF, understand the document structure, and produce an alternative view that works on different kinds of devices, all without manual, time consuming work from our team.

These are just three examples where we see applications in GPO's operations, and we are considering many pilots in the future. All of these are intended to act as a force multiplier for our team, allowing our folks to be more productive, and deliver a higher value for our customers and the taxpayer.

Madam Chairwoman, Ranking Member Fischer, thank you for the opportunity to testify before the committee today. I look forward to any questions the committee may have.

[The prepared statement of Mr. Halpern follows.]

The Chairman. Thank you very much. Last but not least, Ms. Park, thank you.

STATEMENT OF MEROE PARK, DEPUTY SECRETARY AND CHIEF OPERATING OFFICER, SMITHSONIAN INSTITUTION, WASHINGTON, D.C.

Ms. Park. Thank you. And thank you for the opportunity to present today. I will pass along your well wishes to the Secretary. Early in our time as an institution, we collected meteorological data from volunteers nationwide via telegraph.

Fast forward to a few years ago, when we helped create the first image of a black hole at the center of the Milky Way. In short, we have always embraced revolutionary technology, from the locomotive to the pocket watch, from the telegraph to the internet, and artificial intelligence is no different.

Throughout the institution, scientific uses of machine learning are actually quite common, applying it to the identification of exoplanets, tracing the evolutionary history of pollen, and conducting field research in conservation. In other words, artificial intelligence has been around the Smithsonian for almost a decade. What has changed is the widespread availability of generative AI.

These widely available applications of AI hold great promise, from making us more efficient and effective as an organization, to improving the visitor experience and extending our reach across the nation. We are learning

from our current use cases and are exploring new ones as this technology continues to evolve.

For example, the National Museum of African American History and Culture's Freedmen's Bureau Transcription Project, which is the Smithsonian's largest crowdsourcing initiative, is transcribing genealogical records of the formerly enslaved. In a recent trial run, AI tools reliably transcribed nearly two thirds of the test content, showing how AI has the ability to dramatically scale up work with large data sets.

Other examples, including, as the chairwoman mentioned during her opening remarks, our data science labs' development of a model that can discover incorrect women's contributions mistakenly attributed to men in our collections.

And the Smithsonian Astrophysical Observatories Astro AI Center has more than 50 AI projects planned or underway, creating excitement about our solar system and the universe itself.

Inside the Smithsonian, we recently announced the creation of an AI Community of Practice, which will help us test new capabilities across the institution, as well as build a framework of good governance to determine the best way to use AI effectively and responsibly.

People rely on our reputation as a trustworthy

reservoir of knowledge, so we understand concerns about bias, ethics, safety, and accuracy in available AI outcomes, but there are also fantastic opportunities as well. We and other cultural institutions can collaborate with technology leaders to help improve AI tools for everyone.

We can also examine the many dimensions of AI from a scholarly perspective, convening experts to examine its ethical, social, and economic implications. The nation's 250th birthday in 2026 will be a pivotal moment for us to experiment with innovative AI tools, to enable collaboration with tech leaders in other Government and cultural organizations to make AI more reliable and trustworthy, to educate people about the benefits and the risks of AI.

Ultimately, AI is a tool. Used properly, it will allow the Smithsonian to expand our reach and our impact. Like any new technology, it comes with risks and the potential for unintended consequences.

But by developing internal expertise and partnering with external experts, we can thoughtfully leverage AI, both for the Smithsonian and for the American public. Thank you.

[The prepared statement of Ms. Park follows.]

The Chairman. Thank you very much. I am going to turn it over to Senator Fischer and -- for her questions, and then Senator Padilla.

Senator Fischer. Thank you, Madam Chairman. And thank you all once again for being here today. Can each of you share whether your respective agencies have an official policy that governs its use of AI? Dr. Hayden, do you have an official policy on that?

Dr. Hayden. Yes. The Library has a number of policies for security, IT security in general, and we have an AI working group that has developed guidelines.

And we also follow NIST framework. We also are very cognizant and use the Executive Orders that have also laid that. And so, we are very, I would say zealous on making sure that we have those frameworks and guidelines.

Senator Fischer. You mentioned in your oral testimony, it is in your written as well, about NIST and it aims to improve the trustworthiness of AI applications. Can you give me an example of how that would be used? I think a lot of us, when we think of AI, we don't get down to the nitty gritty.

We don't get down to the basics of what it really can do. We just think, oh, this is going to be so wonderful. It will be sorting data. And so, I think this -- I think some of your answers from all of you would be very helpful

on that.

Dr. Hayden. One of the through lines for each of us is trustworthiness and being -- and with the Library of Congress, we take very seriously being stewards of historic information, current information, that we are a trusted source.

And so, working with the framework that includes looking at a new technology, you map it, you explore, and then you measure its usefulness and effectiveness. And then if it passes quite a few of these things, you would look at how you could implement it.

So, being able to make sure and having these templates. We have questions that we pose when anything is being proposed about AI in particular, what does it do? How can we make sure that humans are in the loop for their quality control?

So, it is a very important part of looking at AI in particular and any newer technology.

Senator Fischer. Thank you. And, Mr. Halpern, can you let me know if you -- do you have a policy?

Mr. Halpern. Absolutely. So, we have a directive in place since last month, since the end of 2023. And one of the things that that created was our AI guidance committee, and I believe that committee is going to have their first kickoff meeting on Monday, where they are going to start

evaluating different technologies for some pilot deployments in the near term so we can start testing some of these in our environment.

Senator Fischer. You know, you mentioned in your testimony, a lot of -- the three interesting examples you gave of how you are using AI right now.

So, how do you decide on a case by case basis when you are ready to test that on that project, whether you think it is safe at that point in time, or if you are just trying to gather information.

Mr. Halpern. So, there are -- we already have a very robust system for evaluating software and hardware for deployment at GPO.

Much of what we do is mission critical to the Congress, so we need to make sure that anything we deploy doesn't interfere with our ability to deliver for our customers, whether that is the House or the Senate, or our State Department customer, or any other customer.

So, this is another layer on top of that. And looking at some of these new technologies, both to see if there are improvements to the way we do things that can leverage benefit -- or deliver benefits to our customers and minimize the risk from those same technologies of introducing variables into our product. So, for instance --

Senator Fischer. What does that mean, the risks of introducing variables --?

Mr. Halpern. So, we have all read about the issues with ChatGPT hallucinating facts that aren't there. The good news is GPO doesn't generate a lot of public information. We make other people's public information available.

So, much of where we have got some -- we can leverage these technologies is really in the quality assurance world. But, by the same -- for the same reason you can't -- you need to double check a ChatGPT's output, we need to make sure that the optical system that is checking the quality of our passport identity pages isn't introducing or missing quality problems.

So, from that perspective, we really need to put these systems -- deploy them in a limited way so that we can make sure they are doing what they are advertised to do so that we aren't delivering a product to our customer that is of lesser quality than it should be.

Senator Fischer. Can I have the Smithsonian response as well.

Ms. Park. Of course.

Senator Fischer. Thank you. Ms. Park.

Ms. Park. Great. Thank you. So, I will address maybe two aspects of your question. Internally at the

Smithsonian, we absolutely are focused on governance and also the sharing of best practices.

So, this community of practice that we have formed is really designed to do a number of things. One is to identify policies that we already have that need to be updated with what is a very quickly evolving technology. And also, things like protecting privacy.

You asked about specific examples of how does this actually play out in an organization. You have to make sure that if employees are using ChatGPT or some other generative AI tool, that we are thinking about how that data is being prepared and made available to AI to analyze.

So, there is a lot of thought going into how we go about doing that internally. And then externally, we also have a public mandate, as several of my colleagues have talked about here, as trusted sources for the American public.

We want to make sure that we are available to them, that we have our 177 years of experience with history, culture, science, education as a way to inform these tools, and also as a way to help evaluate the accuracy of information.

So that is still aspirational for us at this point, but we are thinking about both aspects of our internal governance, our internal opportunities for efficiency, as

well as our ability to be a convener and to help in the more public realm.

Senator Fischer. Thank you.

The Chairman. Okay, Senator Padilla. Thank you.

Senator Padilla. Thank you, Madam Chair. As I have had a chance to share in this committee before -- before entering the Senate, I served as California's Secretary of State, where a part of the portfolio was serving as a custodian of the State's archives.

At the time, I made it a priority to dramatically increase public access to our State's archives and this involved a significant undertaking to digitize our archive materials to make them easily viewable, not just to the people of California, but to the world.

Question, first question is for Ms. Park. You noted that prioritizing digitization is one of the Smithsonian's key considerations with respect to AI. How do you envision leveraging AI to increase public access to the Smithsonian?

Ms. Park. Indeed. And I will maybe make a couple of comments in response to your question. The first is that digitization is absolutely one of our priorities as an institution. We have been at it actually for a while. It is a behemoth task with over 150 million objects, and we are working hard to identify ways to prioritize that. We have a long ways to go in terms of having those materials

digitized.

So, when you refer to your own experience in California, the first step, obviously, with AI is even getting the records available and readable and analyzable in a digital format. So that is job number one for us.

And the second part, is we believe that, unlike a search engine that would be able to help the public discover certain items that we might have in our collection and learn about them, AI has the ability for the public to make connections among the documents and objects and items that we have in ways that we can't do as humans.

That we can use generative AI, and we hope to make it available to the public as well, to be able to discover things that we didn't even realize we had in our collections. So, I think there is some really exciting opportunities here, but as you noted, job one is getting more of our items digitized and readable in a way that AI can make use.

Senator Padilla. And, let you in on a secret, maybe a good strategy. Whatever is most popular already, that is where to start.

Ms. Park. The ruby slippers.

Senator Padilla. Because that tends to be the most popular to everybody else. And just imagine the curating opportunities that are possible when you get the metadata

and the keywords right. So, there is some good experts. That will be helpful. Dr. Hayden, same question to you.

Dr. Hayden. The Library has been very, really at the forefront in the library field with looking at how digitized material could be made more discoverable, because if it is not discoverable, it is really not accessible.

And so, one of the projects that I mentioned in my testimony was taking a 20 year project of digitizing and making available newspapers from all 50 States, but adding an AI instrument, a navigator. You can now do even more research into those newspapers. The photos that are in there, names, all types of things.

And so, we have seen that that has made what was just basically making things, putting them into digital form from the vast amount of analog material we have in print material going back hundreds of years, just basically digitizing them.

Yes, that is an accessible aspect. But having tools that have people be able to get into them has proven very popular, especially when we allow the public to use material in a creative way, like our Citizen DJ program, where they can -- we put the material up there, music, all types of things, and people can use them.

So, these tools can be very helpful for having people make those connections between our various collections

-- we have collections in 470 languages. And so, being able to use AI tools to help --

Senator Padilla. A follow up question to both of you. Do you envision institution wide or even specific initiatives to always be contained within your respective institutions, or are there opportunities for public, private partnerships that you are envisioning at this point? Dr. Hayden and then Ms. Park.

Dr. Hayden. Well, the Library of Congress and the Smithsonian are co-chairing a Secretariat in this coming year that will look at more ways that we can all collaborate and, of course, opportunities with the private sector when we are making sure there are guardrails to them.

Senator Padilla. Correct.

Ms. Park. I would just add that absolutely, there are opportunities to work with all sorts of external groups. Higher education, companies, all are interested in this space.

And we have formed several consortiums with different organizations to try to figure out if we can convene together to talk about these kinds of issues, to provide space for people to discuss the complexities of AI. So, absolutely, there are ways for us to partner with others.

Senator Padilla. Okay. Thank you very much.

The Chairman. Very good. Thank you. Senator Hagerty.

Senator Hagerty. Thank you, Chair Klobuchar and Ranking Member Fischer. Appreciate you holding this hearing. Dr. Hayden, it is good to see you back here again -- to all three of our witnesses. Dr. Hayden, I would like to start with you, though, if I might. We all know the Library of Congress houses the U.S. Copyright Office.

Dr. Hayden. Yes.

Senator Hagerty. And I think, as a result, you are in a unique position to ensure the proper implementation of AI. And I think, we all know that AI offers many promising benefits, but with its increasing prevalence, it is not without drawbacks.

So, I would like to take this home, and you and I have talked about my affinity for the songwriting and music business in my home State of Tennessee. We are actually home to thousands of performing artists there in Tennessee, songwriters, artists, all of whom have the potential to be greatly impacted by artificial intelligence.

I am concerned that AI generated deepfake recordings, voice clones, or other infringements on the works of copyright holders could cause serious damage to an artist's reputation and to their livelihood.

Earlier this month, Governor Bill Lee announced plans

for legislation to protect the voices of performing artists by adding their voices and sound to Tennessee's protection of personal rights law.

I am pleased that my home State of Tennessee is leading the nation in providing legal protection for our artists and for our songwriters. And my first question, Dr. Hayden, is given the growing pervasiveness of AI, how does the Copyright Office plan to protect the copyrighted works and uses of voices of performing artist at the Federal level?

Dr. Hayden. The Copyright Office has been in the forefront of looking at the issues involved with machine generated works. And in fact, they have developed a separate webpage for people who are creators, but just giving so much guidance on that.

They have hosted a number of presentations and webinars, and our Register of Copyright Shira Perlmutter has been in the last year in many, many international copyright sessions and things. And so, they are still -- they just issued a rule to get information, and they had over 10,000 responses.

And they are going to be issuing more information about what they are hearing from stakeholders about how that can really be used, but also be protecting.

Senator Hagerty. Can I get into -- let's drill into

this a little bit more in terms of information you are getting back from stakeholders. In August, the Copyright Office issued a notice of inquiry on copyright and artificial intelligence. The notice sought information and comments on a number of issues, including the use of copyrighted works to train AI models.

The appropriate levels of transparency and disclosure that are required with respect to the use of copyrighted works, the legal status of AI generated outputs, and the appropriate treatment of AI generated outputs that mimic the personal attributes of human artist.

I am sure that many Tennessee songwriters and artists submitted information to the Copyright Office urging the implementation of several things. One, the requirements for obtaining appropriate licenses or authorization to ingest or otherwise use materials that are copyrighted, or that implicate a person's rights of publicity or privacy.

Second, requiring adequate recordkeeping and auditing. And third, ensuring that appropriate transparency is maintained by AI companies. So, Dr. Hayden, are these remedies that I just listed, are those the types that the Copyright Office is considering?

Dr. Hayden. Well, the office is currently reviewing those comments, over 10,000, and they cover the full range. And so, what they are saying, and I have the information

here, that they are really looking at what could be infringement, the treatment of these types of AI generated outputs, and currently they are preparing that report. So, I can't -- further than that but they will --

Senator Hagerty. I hope they will take into account the three points that I have raised that are certainly of interest to our industry in Tennessee. Rights to the voices and likenesses of performers are generally protected by State law rather than copyright law.

That includes laws providing a right of publicity, privacy protections, and remedies for misappropriation. Given this, in your opinion, Dr. Hayden, are statutory changes needed to adequately protect the name, the image, the likeness of performing artists?

Dr. Hayden. Our Register of Copyrights, Shira Perlmutter, is very involved with that. And so, she will be working with Congress on developing any potential legislation or any other things that would impact the copyright law as it stands. Right now, it is producing works by a human being. That is the standard.

Senator Hagerty. I certainly hope you will view my office as a resource and the people of Tennessee as a resource here. I think we can add significantly to the conversation. Thank you, Dr. Hayden, and I yield back my time. Thank you.

The Chairman. Very good. Thank you. Some of the risks of AI are already clear. I talked about what just happened with those robocalls and what we are seeing on -- really in fake videos on both sides of the aisle. It isn't just one side. I know the Republican Attorney General is investigating what happened in New Hampshire.

But we know that there is going to be other security issues, from our infrastructure, cyber-attacks. Could you talk about what your agencies have done to guard against cyber security threats to ensure that your systems remain strong in the face of potential AI threats? Each one of you could answer. Dr. Hayden.

Dr. Hayden. The Library of Congress, with the help of Congress, has been able to build a very, very strong cybersecurity foundation over a number of years.

And in fact, our CIO, Judith Conklin, is a cybersecurity expert. And so that has been a major part of what we have been able to bring to looking at the risks with this particular technology, AI. And so, it fits into our governance -- IT governance framework already, and we are making sure that we are not on the cutting edge with AI.

We are looking at our -- and using the experiences that we have had. But there is a technology governance board within the Library, and AI now has a working group

within that board to make sure that everything is aligned with that. We are very concerned about the trustworthiness of products that have are.

The Chairman. Okay. Very good. Mr. Halpern.

Mr. Halpern. Absolutely. So we, our team works very hard to make sure that our systems are as impregnable as possible. But we are also looking at some AI tools as helpful in those efforts.

And one of the things that AI excels at is pattern matching. So, they may be able to see patterns in our logs that a human being might not readily identify. So that is actually a benefit. One of the other areas that is sort of tangentially related is in the privacy sphere. Searching materials produced by others for personally identifiable information.

Just like many, many other organizations, we have had issues in the past where it has been hard for us to catch every incidence of PII in our materials that we make public. But again, these AI based tools that can recognize whether it is a Social Security number or address or another piece of PII in the Federal Register, the Congressional Record, or anything else we publish, that is an additional layer of security that these tools can help us with, so that we don't inadvertently put something out in the public domain that isn't supposed to be there.

The Chairman. Very good. Ms. Park.

Ms. Park. So, we take our cyber security responsibilities very seriously, both in terms of our own internal practices, but also in terms of our public domain. We do comply with NIST standards when it comes to our cyber profile.

So, our CIO is very much checking constantly to make sure that we are following those standards. Our Inspector General is also there to regularly check on our progress. And the other thing we do is we regularly report and, participate in Federal forums where you share information about the latest threats that might be coming through various channels, including, the emergence of AI and how it is being used.

For one, for example, phishing has become much more sophisticated now with the emergence of AI and the ability to create much more easily emails and other things that appear to be real. So, there are -- so we are stepping up our testing of our own workforce in their terms of their ability to recognize these things.

So, there is a lot of work going on there. And finally, I would mention there is an internal group led by our CIO right now that is really trying to develop or evolve our policy structure to make sure that we are prepared for whatever threats that might be coming from

emerging technologies, including AI.

The Chairman. And so, are you all taking part in cybersecurity training? Do you think that will be important as we see this changing world, both opportunities and risks, with your staff?

Dr. Hayden. Definitely --

The Chairman. Is your staff doing -- okay. Yes.

Mr. Halpern. Absolutely. And we are evaluating vendors now for some specific AI related training so that folks understand these tools, and both the good and the bad that they can bring.

Ms. Park. We have an annual training requirement for all staff who have access to our network.

The Chairman. Okay. Dr. Hayden, Senator Hagerty talked to you a bit about copyrights voices. When you appeared before the committee last year, we discussed the Copyright Office's new initiative to examine copyright law and policy issues raised by AI. What are the lessons learned by the Copyright Office over the past year? And have they continued to see increases in the amount of copyright applications for AI generated content?

Dr. Hayden. Yes, they have. And they have actually had a number of cases. There have been, as I said, about 100 works that have had AI generated and human authorship, and they have been registered by the office.

And they are also taking a leadership role internationally when looking at what is the proportion of the AI generated content and human, as well. So, they have definitely been on the forefront of looking at that. But right now, they have rejected several applications because of the amount of AI authorship.

The Chairman. One last question of you, and then I will have just one more. Your written testimony highlights the work being done to test potential AI tools for the National Library Service for the Blind and Print Disabled, including improving the accessibility of book descriptions.

And we have talked about this way back, some of the work we have done with the blind. What guardrails has the Library put in place to ensure the content being produced by these tools is accurate for its users?

Dr. Hayden. And that is part of the LC Labs, and working directly with a service unit, NLS, Library for the Blind and Print Disabled. And so, they have had quite a bit of success with the e-reader rollout.

And so, they are looking at how can AI be used to help this population. And there are some exciting opportunities, but they are working hand in hand to make sure that what is being presented to that community is trustworthy.

And then you see that that is the main theme with, I

think all of us, is that we want to use this tool for a number of reasons, for efficiency, effectiveness, but also we have to make sure that there is that oversight --

The Chairman. Okay. Last, Ms. Park, you talked about the assembling the group of experts from across the Smithsonian to identify best practices, opportunities. Could you talk about what next steps are after that, what the timetable is?

Ms. Park. Yes. So, this group is charged with doing a number of things. The governance piece is an important one, reviewing policies, and also finding what is working within the units. So, we have multiple museums and research institutes.

And so, figuring out where we can do some things to scale rather than have individual units try to do their own thing where we can, for efficiency sake. I have asked for them to do their initial work and come back with some recommendations for what we might need to do in the short term.

What are some of the immediate stopgaps that need to be put in place, whether it is governance related or a best practice related issue that we want to share. And then I suspect there will be some longer term things that will be -- take a little longer to address, whether it is systems related, business process, or applications. So, I expect

it to be a bit of a phased approach.

The Chairman. My last question is, since Judy Garland's shoes came up, do you know where she was born, Judy Garland?

Ms. Park. Oh, no, this is a quiz.

The Chairman. Yes. Well, Grand Rapids, Minnesota.

Ms. Park. Yes, I was going to say -- I was going to say Minnesota.

The Chairman. And they just sort of solved the crime of the missing shoes that were there, which I would just like to note for the record, finally. That they have been returned. It is a sordid story. Yes, but they are back. They are back.

So, in any case, thank you very much. Do you want to ask any other questions, Senator Fischer? We have a vote that has been called and thank you for the work that you are doing, continuing to do, leading your agencies.

I speak for myself, I know other members of the committee, we feel good about your stewardship of these very important agencies, and we look forward to working with you on the important work you are doing.

As we discuss opportunities, risks, I am sure we will have more discussions about this. The hearing record will remain open for one week, and we are adjourned. Thank you.

[Whereupon, at 3:12 p.m., the hearing was adjourned.]

WORD INDEX