

**KEEPING UP WITH THE CODES:
USING AI FOR EFFECTIVE REGTECH**

HYBRID HEARING
BEFORE THE
TASK FORCE ON ARTIFICIAL INTELLIGENCE
OF THE
COMMITTEE ON FINANCIAL SERVICES
U.S. HOUSE OF REPRESENTATIVES
ONE HUNDRED SEVENTEENTH CONGRESS
SECOND SESSION

MAY 13, 2022

Printed for the use of the Committee on Financial Services

Serial No. 117-85



U.S. GOVERNMENT PUBLISHING OFFICE

47-880 PDF

WASHINGTON : 2022

HOUSE COMMITTEE ON FINANCIAL SERVICES

MAXINE WATERS, California, *Chairwoman*

CAROLYN B. MALONEY, New York	PATRICK McHENRY, North Carolina,
NYDIA M. VELAZQUEZ, New York	<i>Ranking Member</i>
BRAD SHERMAN, California	FRANK D. LUCAS, Oklahoma
GREGORY W. MEEKS, New York	BILL POSEY, Florida
DAVID SCOTT, Georgia	BLAINE LUETKEMEYER, Missouri
AL GREEN, Texas	BILL HUIZENGA, Michigan
EMANUEL CLEAVER, Missouri	ANN WAGNER, Missouri
ED PERLMUTTER, Colorado	ANDY BARR, Kentucky
JIM A. HIMES, Connecticut	ROGER WILLIAMS, Texas
BILL FOSTER, Illinois	FRENCH HILL, Arkansas
JOYCE BEATTY, Ohio	TOM EMMER, Minnesota
JUAN VARGAS, California	LEE M. ZELDIN, New York
JOSH GOTTHEIMER, New Jersey	BARRY LOUDERMILK, Georgia
VICENTE GONZALEZ, Texas	ALEXANDER X. MOONEY, West Virginia
AL LAWSON, Florida	WARREN DAVIDSON, Ohio
MICHAEL SAN NICOLAS, Guam	TED BUDD, North Carolina
CINDY AXNE, Iowa	DAVID KUSTOFF, Tennessee
SEAN CASTEN, Illinois	TREY HOLLINGSWORTH, Indiana
AYANNA PRESSLEY, Massachusetts	ANTHONY GONZALEZ, Ohio
RITCHIE TORRES, New York	JOHN ROSE, Tennessee
STEPHEN F. LYNCH, Massachusetts	BRYAN STEIL, Wisconsin
ALMA ADAMS, North Carolina	LANCE GOODEN, Texas
RASHIDA TLAIB, Michigan	WILLIAM TIMMONS, South Carolina
MADELEINE DEAN, Pennsylvania	VAN TAYLOR, Texas
ALEXANDRIA OCASIO-CORTEZ, New York	PETE SESSIONS, Texas
JESÚS “CHUY” GARCIA, Illinois	
SYLVIA GARCIA, Texas	
NIKEMA WILLIAMS, Georgia	
JAKE AUCHINCLOSS, Massachusetts	

CHARLA OUERTATANI, *Staff Director*

TASK FORCE ON ARTIFICIAL INTELLIGENCE

BILL FOSTER, Illinois, *Chairman*

BRAD SHERMAN, California
SEAN CASTEN, Illinois
AYANNA PRESSLEY, Massachusetts
ALMA ADAMS, North Carolina
SYLVIA GARCIA, Texas
JAKE AUCHINCLOSS, Massachusetts

ANTHONY GONZALEZ, Ohio, *Ranking
Member*
BARRY LOUDERMILK, Georgia
TED BUDD, North Carolina
TREY HOLLINGSWORTH, Indiana
VAN TAYLOR, Texas

CONTENTS

Hearing held on:	Page
May 13, 2022	1
Appendix:	
May 13, 2022	21

WITNESSES

FRIDAY, MAY 13, 2022

Greenfield, Kevin, Deputy Comptroller for Operational Risk Policy, Office of the Comptroller of the Currency (OCC)	4
Hall, Melanie, Commissioner, Division of Banking and Financial Institutions, State of Montana; and Chair, Board of Directors, Conference of State Bank Supervisors (CSBS)	5
Lay, Kelly J., Director, Office of Examination and Insurance, National Credit Union Administration (NCUA)	7
Rusu, Jessica, Chief Data, Information and Intelligence Officer, Financial Conduct Authority (FCA), United Kingdom	9

APPENDIX

Prepared statements:	
Greenfield, Kevin	22
Hall, Melanie	38
Lay, Kelly J.	46
Rusu, Jessica	52

ADDITIONAL MATERIAL SUBMITTED FOR THE RECORD

Foster, Hon. Bill:	
Written statement of the Alliance for Innovative Regulation	60
Written statement of the Data Foundation	66
Written statement of the National Association of Federally-Insured Credit Unions	184
Garcia, Hon. Sylvia:	
Written responses to questions for the record from Kevin Greenfield	187
Written responses to questions for the record from Kelly Lay	189
Gonzalez, Hon. Anthony:	
Written statement of Security Scorecard	190
Maloney, Hon. Carolyn:	
Written statement of Security Scorecard	190

KEEPING UP WITH THE CODES: USING AI FOR EFFECTIVE REGTECH

Friday, May 13, 2022

U.S. HOUSE OF REPRESENTATIVES,
TASK FORCE ON ARTIFICIAL INTELLIGENCE,
COMMITTEE ON FINANCIAL SERVICES,
Washington, D.C.

The task force met, pursuant to notice, at 9 a.m., in room 2128, Rayburn House Office Building, Hon. Bill Foster [chairman of the task force] presiding.

Members present: Representatives Foster, Sherman, Adams, Auchincloss; Gonzalez of Ohio, and Loudermilk.

Chairman FOSTER. The Task Force on Artificial Intelligence will now come to order.

Without objection, the Chair is authorized to declare a recess of the task force at any time. Also, without objection, members of the full Financial Services Committee who not members of this task force are authorized to participate in today's hearing.

Today's hearing is entitled, "Keeping Up with the Codes: Using AI for Effective RegTech."

I now recognize myself for 5 minutes to give an opening statement.

Thank you, everyone, for joining us today for what should be a very interesting hearing of the task force. Today, we are looking to explore how different financial regulators have been leveraging AI technology to improve their regulatory and supervisory efforts. The use of artificial intelligence (AI) and machine learning (ML) by financial institutions, known as RegTech, as well as the use of AI and ML by financial regulators, known as SupTech, are two areas that are rapidly gaining regulatory focus.

To give our audience some idea of what we are talking about, imagine that a regulator comes up with a simple, rules-based method for automatically generating, say, suspicious activity reports (SARs). Such a system will soon be flooded with false positives, that is, activities that appear suspicious but are in fact not. So, an experienced human regulator will then be assigned to sort through these to determine which are spurious and which are truly suspicious activities.

And this will rapidly generate what is called a tag, a dataset that can be used by machine-learning algorithms to approximately replicate the human's complex judgment at scale and at low cost, which all sounds great until you encounter all of the classic problems with AI and machine learning. That is, if the human inspector was racist, the machine-learning algorithm will be racist. And so-

phisticated bad actors will have the ability to pollute the dataset with thousands of non-suspicious activities so as to trick the ML algorithm to give a pass to a small number of serious violations.

This is a complex space, and taken together in the financial services space, RegTech and SupTech may help financial institutions and government regulators to monitor transactions, evaluate risk, detect noncompliance, identify illicit finance, and implement regulatory changes at a rate never seen before. The AI community anticipates that the use of RegTech and SupTech will grow significantly in the coming years, with a global RegTech market projected to reach over \$55 billion by 2025. This rapid growth exemplifies why it is so important that we get these technologies right, that we need to make sure that we leverage these powerful tools in the right way.

Just as we work to ensure that private industries utilize AI and ML in a responsible manner, we need to be sure that our government regulators and entities embrace such technology to improve oversight and strengthen supervision of the financial services industry. Financial institutions' risk management and compliance operations rely heavily on synthesizing large datasets and detecting both trends and uncharacteristic activity. A large portion of this calculus depends on the ability of these firms and regulators to successfully predict whether wrongdoing or other risks may occur in the future.

With this great predictive ability, AI must remain accountable to broad regulatory schemes. Those employing AI in their business processes or regulatory efforts cannot avoid responsibility when unintended or unwanted results occur. As with any AI, the use and implementation must be carefully monitored and evaluated in the end by a natural person in some capacity. Similarly, regulators and inspectors must be able to look under the hood of business operations utilizing AI, especially in sensitive and impactful industries such as the financial services space. These algorithms and programs must be accountable, explainable, and adjustable.

So, I am glad we are gathered here to discuss such an important topic and to better ascertain how we can promote responsible innovation and adoption amongst both financial services participants and regulators.

And with that, I would like to recognize the ranking member of the task force, my friend from Ohio, Mr. Gonzalez, for 5 minutes.

Mr. GONZALEZ OF OHIO. Thank you, Chairman Foster, and thank you, as always, for your leadership on this task force. It's always a pleasure to work together, and I share your interest and excitement for the possibilities of increased use of emerging technologies in our regulatory system.

Technology has transformed the way that Americans interact with financial institutions, for the better. Today, Americans can access banking services, buy and sell stock, and get approved for a mortgage faster and cheaper than ever before without ever having to step foot into a local bank or credit union. In our work on this committee, we should be doing more to promote the adoption of new technologies to further optimize operations and open finance to more Americans. The expanded use of AI and machine learning

will help financial institutions use data more efficiently and reach more Americans.

That brings us to today's hearing. In one of my first hearings as a member of this committee, I advocated for our financial regulators to explore the applications of emerging technologies like AI and machine learning. As financial institutions have made great strides in innovation, so, too, are financial regulators.

To me, the benefits are twofold. First, it will help decrease regulatory costs on financial institutions, particularly our smallest financial institutions which are often the bedrock of local communities. A 2015 Federal Reserve Bank of St. Louis study found that compliance costs are disproportionately high for small banks, at an estimated cost of 22 percent of net income. Globally, the cost of compliance for banks is estimated to be at least \$100 billion. This prevents institutions from being able to invest those dollars into further innovations, but, most importantly, makes financial products more expensive for everyday customers. Our goal should be to reduce these regulatory costs and give institutions, especially smaller financial institutions, the flexibility to deploy their capital more efficiently.

It is important to note that decreasing regulatory cost does not necessarily mean decreased regulation. This leads me to my second point. Increased use and deployment of RegTech and SupTech will lead to more effective and efficient regulations. As the financial institutions rapidly increase their use of AI and machine learning, we run the risk of traditional financial regulatory systems being vulnerable to failure. In my view, the use of RegTech can play a critical role in enhancing consumer protection, containing systemic risk, reducing financial crimes, and creating greater regulatory feedback between financial institutions and Federal agencies.

In today's hearing, I am excited to hear from our witnesses on the current utilization of AI and machine learning by private industry and Federal agencies, and the barriers facing regulators, including whether agencies have sufficient technical staff resources and sufficient congressional authority. We must also confront the potential difficulties of the greater use of RegTech, including bias, algorithm challenges and implementation, and the protection of Americans' personally identifiable information.

Again, I want to thank Chairman Foster for convening today's hearing. I look forward to hearing from our witnesses, and I yield back.

Chairman FOSTER. Thank you. And today, we welcome the testimony of our distinguished witnesses: Kevin Greenfield, the Deputy Comptroller for Operational Risk Policy at the Office of the Comptroller of the Currency; Melanie Hall, the Commissioner of the Montana Division of Banking and Financial Institutions, as well as the Chair of the Board of Directors of the Conference of State Bank Supervisors; Kelly Lay, the Director of the Office of Examination and Insurance at the National Credit Union Administration; and Jessica Rusu, the Chief Data, Information, and Intelligence Officer of the UK Financial Conduct Authority.

Witnesses are reminded that their oral testimony will be limited to 5 minutes. You should be able to see a timer on your screen that will indicate how much time you have left, and a chime will go off

at the end of your time. I would ask that you be mindful of the timer, and quickly wrap up your testimony if you hear the chime, so that we can be respectful of both the witnesses' and the task force members' time.

And without objection, your entire written statements will be made a part of the record.

Mr. Greenfield, you are now recognized for 5 minutes to give an oral presentation of your testimony.

STATEMENT OF KEVIN GREENFIELD, DEPUTY COMPTROLLER FOR OPERATIONAL RISK POLICY, OFFICE OF THE COMPTROLLER OF THE CURRENCY (OCC)

Mr. GREENFIELD. Chairman Foster, Ranking Member Gonzalez, and members of the task force, thank you for the opportunity to appear today and discuss artificial intelligence tools used by national banks, Federal savings associations, and Federal branches and agencies of foreign banks supervised by the Office of the Comptroller of the Currency (OCC). I appreciate this invitation to discuss the opportunities, benefits, and challenges that artificial intelligence or, more commonly, AI, presents for banks, and our approach to supervising those activities.

I serve as the OCC's Deputy Comptroller for Operational Risk Policy, and I am responsible for overseeing the development of policy and examination procedures addressing bank operational risk, which includes understanding and monitoring the risks of AI. I also represent the OCC in international forums, such as the Basel Committee on Banking Supervision's Financial Technology Group, which coordinates the sharing of regulatory practices on tech issues, including those related to the use of AI.

Technological changes and rapidly-evolving consumer preferences are reshaping the financial services industry, creating new opportunities to provide consumers, businesses, and communities with more access to financial products and services. The OCC promotes responsible innovation in the banking industry to expand access to credit and capital, improve operations, and support full and fair participation in the American banking system. Over the years, we have adapted our supervisory approach to address the increase in banks' use of technological innovations, such as AI.

Today, banks can use AI tools to strengthen their safety and soundness, enhance consumer protections, and increase fair access to products and services. AI can be used to enhance the customer experience, such as assisting with online account openings and product selection. AI can also be used to support more efficient credit underwriting and other banking operations. Used in appropriate ways, these approaches have the potential to promote greater access to banking services by underserved communities. Use of advanced analytical tools is not new, and banks have been employing mathematical models to support operations for some time. However, banks are now introducing AI tools to support even more complex operations and increased automation.

While we have seen many large banks develop these tools internally, AI tools and services are becoming more widely available as third-party firms are increasingly offering AI products and services to banks of all sizes. While AI tools can present benefits, we must

also be mindful of the risks that banks' use of AI is not properly managed and controlled. Potential adverse outcomes are caused by poorly-designed models, faulty data, inadequate testing, or limited human oversight. Banks need effective risk management and controls for model validation and explainability data management, privacy, and security, regardless of whether a bank develops AI tools internally or purchases through a third party.

The OCC follows a risk-based supervision model focused on safe, sound, and fair banking practices, as well as focused on assessing compliance with laws and regulations. OCC examiners have significant experience in supervising banks' use of sophisticated mathematical models and tools. This includes evaluating fair lending concerns and other consumer protection issues, such as unfair or deceptive acts or practices. The OCC expects banks to monitor for and identify outcomes that could create unwarranted risks or adversely impact the fair treatment of customers.

If we identify any concerns, risks, or deficiencies during our examinations, the OCC has a range of tools available and will take supervisory or enforcement action as appropriate. But just as banks are increasingly using sophisticated technologies and tools to enhance bank capabilities, the OCC is similarly engaged in assessing how innovative technologies can strengthen our supervisory processes. The OCC employs a number of analytical and technology tools to support banks' supervision, and work is currently underway to materially upgrade our core supervision systems to further enhance this ability to monitor risks in the banking system. Moreover, the OCC is considering the use of AI tools as part of this effort.

I want to thank the task force for its leadership on this important issue and for inviting the OCC to testify today. I look forward to answering your questions.

[The prepared statement of Deputy Comptroller Greenfield can be found on page 22 of the appendix.]

Chairman FOSTER. Thank you, Mr. Greenfield.

Ms. Hall, you are now recognized for 5 minutes to give an oral presentation of your testimony.

STATEMENT OF MELANIE HALL, COMMISSIONER, DIVISION OF BANKING AND FINANCIAL INSTITUTIONS, STATE OF MONTANA; AND CHAIR, BOARD OF DIRECTORS, CONFERENCE OF STATE BANK SUPERVISORS (CSBS)

Ms. HALL. Good morning, Chairman Foster, Ranking Member Gonzalez, and members of the task force. Thank you for holding this hearing. My name is Melanie Hall, and I am the Commissioner of Montana's Division of Banking and Financial Institutions, and the Chair of the Conference of State Bank Supervisors' Board of Directors. It is my pleasure to testify today on behalf of CSBS on how the State system uses technology today and how we see it shaping the future through networked supervision, a federated regulatory approach to further evolve and streamline the State system.

First, I would like to address how States are using technology platforms as well as data analytics to enhance States' supervision today. CSBS operates a regulatory licensing platform called the

Nationwide Multistate Licensing System, or NMLS, on behalf of State regulators. NMLS became part of the SAFE Act in 2008, but got its start in the early days of the mortgage crisis as State regulators recognized the need to stop bad actors from leaving one State just to move to another. Today, nearly 640,000 companies and individuals use NMLS annually to manage their business licensing and registration. These companies and individuals span the mortgage, consumer finance, debt, and money services business (MSB) industries, with the mortgage industry accounting for more than 80 percent of NMLS use.

The NMLS is also used to register mortgage loan originators who work at banks and credit unions. This RegTech has helped State regulators become more efficient and risk-focused. NMLS data and information is used to identify trends in licensing and supervisory activities. In particular, the data has given State regulators a deeper perspective into the mortgage industry and helped to identify applications and companies that may require more scrutiny. NMLS data further helps State regulators analyze nationwide trends and identify risks through quarterly reports on both the mortgage and MSB industries. CSBS is also in the early stages of developing a call report for consumer finance.

Building on the success of NMLS, in 2020, the States launched a new technology platform called the State Examination System (SES). SES is the first nationwide system to bring regulators and companies into the same technology space for examinations. State agencies can conduct exams and investigations, process consumer complaints, and do other supervision work through this secured platform, and share that information with industry and other States. SES provides uniformity and efficiency, reducing the regulatory burden for multistate companies.

In addition to these two important technology platforms, CSBS has a dedicated data analytics team that works with State regulators to find new ways to anticipate and mitigate risk. Some of the areas that State regulators are exploring include using a technology tool to understand emerging trends from examiners in the field, using predictive modeling as an early detection tool for bank risk, and piloting a predictive analytics program, and understanding how AI could be used to review loan files for consumer compliance and more.

Networked supervision will evolve the State system to one where communication occurs in real time. Knowledge and expertise flows across the States, and regulation becomes streamlined. Last year, the CSBS board identified networked supervision priorities to advance this wide-ranging initiative. The priorities laid the foundation for future collaboration, and further work started under Vision 2020 to streamline the licensing and supervision of money services businesses. Networked supervision requires timely and robust data and information sharing between State and Federal agencies. In addition to developing platforms that support this objective, CSBS is identifying and modernizing the necessary legal underpinnings to enable greater data sharing. The State system has information-sharing agreements with numerous Federal agencies, and we are pursuing and would appreciate your support of more of these arrangements.

States are working toward a future where technology platforms and data analytics allow 54 State financial agencies to operate as one seamless supervisory network. This networked approach will transform financial regulation, giving State regulators an even greater ability to identify and understand local risks before they threaten consumers in the financial system at a national level. As noted in greater detail in my written testimony, the States are committed to implementing technology solutions and collaborating in new ways to improve oversight and enhance consumer protections while reducing regulatory burden.

Thank you for the opportunity to testify today. I look forward to answering your questions.

[The prepared statement of Commissioner Hall can be found on page 38 of the appendix.]

Chairman FOSTER. Thank you, Ms. Hall.

Ms. Lay, you are now recognized for 5 minutes to give an oral presentation of your testimony.

STATEMENT OF KELLY J. LAY, DIRECTOR, OFFICE OF EXAMINATION AND INSURANCE, NATIONAL CREDIT UNION ADMINISTRATION (NCUA)

Ms. LAY. Chairman Foster, Ranking Member Gonzalez, and members of the Task Force on Artificial Intelligence, thank you for conducting this hearing on the effective use of AI and RegTech, and for the opportunity to testify before you today. My name is Kelly Lay, and I am the Director of the Office of Examination and Insurance at the National Credit Union Administration (NCUA). I started my career with NCUA as an examiner in the field and have held various positions throughout the agency. Most recently, I was the NCUA's Director of the Office of Business Innovation, where I led the development and implementation of the agency's new examination platform, the Modern Examination and Risk Identification Tool, also known as MERIT.

In my testimony today, I will first focus on the agency's examination modernization efforts. Second, I will highlight research NCUA has conducted in the realm of AI and RegTech and the NCUA's challenges to incorporate AI and RegTech technologies in the credit union industry. Third, I will discuss last year's request for information on the institutional use of AI. And I will conclude with a legislative request for the NCUA to receive third-party vendor authority.

In 2015, the NCUA formed the Enterprise Solution Modernization Program to help NCUA staff regulate and supervise credit unions more efficiently. The program aims to modernize the NCUA's technology solutions to create an integrated examination and data environment and facilitate a safe and sound credit union system. As an initial step, the NCUA prioritized replacement of the legacy examination application, also known as AIRES, which was over 25-years-old.

After several pilot phases, NCUA rolled out MERIT, including enhanced integrated analytics utilizing a business intelligence tool, and our new secure central user interface, called NCUA Connect, to the NCUA, State supervisory authorities, and credit unions in the second half of 2021. Currently, the NCUA has focused on help-

ing users through this significant transition while deploying system enhancements. In 2017, the NCUA board also approved virtual examination exploration and research funding. Currently, the virtual examination project is in the research and discovery phase.

The agency's goal is to transition, within the next 5 to 10 years, the examination and supervision program into a predominantly virtual one for credit unions that are compatible with this approach. The NCUA is in the testing phase of deploying a machine-learning model to perform data validation more efficiently, with quarterly call reports and profile submissions. Deployment of this new technique is expected to occur in the next 4 quarters and should result in more reliable and consistent call report filing across the industry.

The NCUA recognizes the importance and benefits of technological changes and has incorporated organizational change management strategies into our initiatives. However, there are challenges. In addition to dedicated resources for development and testing, expanding the NCUA's use of RegTech and AI would require the agency to train examiners and credit unions, as applicable, and revise our examination policies and procedures. In addition, while the NCUA supports and encourages innovation and the growth of the industry, we also must protect the interests of credit union members in terms of privacy and security and not compromise the industry's safety and soundness.

Furthermore, most federally-insured credit unions have less than \$100 million in assets. These small credit unions fulfill a vital role in their communities but are usually short-staffed and lack the expertise and resources necessary to keep abreast of changing technology. Generally, the smaller institutions have neither the economies of scale nor the expertise necessary for sophisticated analytics.

Last year, the NCUA joined the OCC, the FDIC, the Federal Reserve, and the Consumer Financial Protection Bureau (CFPB) in a request for information on the institutional use of AI and related challenges. We collectively received responses from financial institutions, vendors, industry trade groups, academic communities, and consumer advocacy organizations. In total, we only received 32 comments, and of those, only 4 were from the credit union industry.

Finally, any examination of technology in NCUA is incomplete without discussing the significant challenges the agency has confronted since the 2002 expiration of its third-party vendor authority. While there are many advantages for credit unions to use these service providers, the continued transfer of operations to credit union service organizations and other third-party vendors diminishes the ability of the NCUA to accurately assess all of the risks present in the credit union system and to determine if current credit union service organization or third-party vendor risk mitigation strategies are adequate.

I would like to thank Chairman Foster for introducing the Strengthening Cybersecurity for the Financial Sector Act to give the NCUA third-party vendor examination authority. I urge the members of this task force to review this legislation and consider

adding their support to close this growing regulatory blind spot that the NCUA continues to confront.

This concludes my statement. I look forward to your questions. [The prepared statement of Director Lay can be found on page 46 of the appendix.]

Chairman FOSTER. Thank you. And Ms. Rusu, you are now recognized for 5 minutes.

STATEMENT OF JESSICA RUSU, CHIEF DATA, INFORMATION AND INTELLIGENCE OFFICER, FINANCIAL CONDUCT AUTHORITY (FCA), UNITED KINGDOM

Ms. RUSU. Good morning, Chairman Foster, Ranking Member Gonzalez, and members of the task force. Thank you for the invitation to appear virtually today. I am currently serving as the Chief Data, Information and Intelligence Officer at the Financial Conduct Authority (FCA). For the committee's background, the FCA is the conduct regulator for approximately 51,000 financial services firms in the U.K. The FCA is responsible for ensuring that relevant markets function well, as well as operational objectives to protect consumers and promote effective competition.

In my role at the FCA, I am focused on building digital supervision technologies and leveraging data science and intelligence capabilities. As stated in our 2022 business plan, we believe that an increasingly data-driven industry should have a regulator. Therefore, the use of AI both by industry as RegTech, and for the purposes of regulatory supervision, or SupTech, is an important area of focus for us. The data technology and innovation division that I lead engages with firms' subject matter experts and fellow regulators to drive positive transformation in how we regulate.

The FCA's innovation services include TechSprints, digital and regulatory sandbox activities, innovation pathways, as well as our scalebox and early oversight for new firms. Our TechSprints are events where we convene industry experts to develop proofs-of-concept to address specific challenges, such as AML and financial crime. The regulatory sandbox allows businesses to test new propositions in the live market with real customers and regulatory oversight, whereas the digital sandbox enables proofs-of-concept to be developed using complex synthetic datasets. Recent digital sandbox participants have focused on ESG data and fraud prevention.

Turning to the focus of today's hearing, we believe that new technologies can bring positive benefits to consumers and markets. As part of our work on AI, we want to facilitate debate on the risks and ethical questions associated with its use. The FCA is actively exploring how we can use AI techniques as well for supervisory and enforcement purposes, including leveraging advanced analytics techniques in our intelligence work, which seeks to extract insights from FCA data to increase the speed and accuracy of decision making, which we will further embed with triaging and intervention models.

Externally, we have partnered with the Bank of England on the development of the AI Public-Private Forum, established in October of 2020, to share information and understand the practical challenges of using AI in financial services, as well as the barriers to deployment and potential risks. The FCA also collaborated with

the Alan Turing Institute on a year-long project which explored the practical application of a high-level framework for responsible AI.

Currently, we are working with the Bank of England to issue a joint public discussion paper on AI, supported by new research that will help us to deepen our understanding of how AI is changing U.K. financial markets. In terms of the high-level outcomes from the work thus far, we see that existing model risk management frameworks reinforced that organizations must take responsibility for algorithmic decision making, regardless of the technology used. And in terms of risk management, we see that AI forums are advocating that human-in-the-loop processes exist. Data is a key building block of responsible AI. We require firms to ensure they demonstrate robust controls, consider data quality, including provenance and recency of data utilized, as well as cyber and data security when implementing new technologies. Governance and accountability are, therefore, core to the way the FCA thinks about AI.

The wider FCA and I would be happy to remain engaged with the committee and with U.S. regulators to continue this discussion. Thank you very much.

[The prepared statement of Ms. Rusu can be found on page 52 of the appendix.]

Chairman FOSTER. Thank you, Ms. Rusu, and I want to thank you also for your excellent written testimony, including many interesting links to all of the great work, things like TechSprints that you are doing, that you are involved in. Unfortunately, it kept me awake way too late last night. Now, to our Members, I would like to say that we anticipate a second round of questions should be possible, so you can keep that in mind.

I will now recognize myself for 5 minutes for questions.

I would like to start by quickly responding to Ms. Lay's important points regarding third-party vendors. In FSOC's annual reports for 2015, 2016, 2017, 2018, 2019, 2020, and again in 2021, which happens to span two Democratic and one Republican Administrations, the Council has highlighted the fact that we have a regulatory blind spot with respect to the oversight of third-party vendors of NCUA and FHFA's regulated entities. Federal banking regulators are able to examine and oversee banks' third-party vendors, which can help ensure those third parties, especially technology firms that banks may utilize, so that they do not oppose cybersecurity vulnerability or other risks to the safety and soundness of the banking system.

The Examination Parity Act gave NCUA and FHFA this very authority from 1991 until it was sunsetted in December 2001. And since then, both agencies, the GAO, and FSOC themselves, have repeatedly and explicitly requested that this authority be reinstated. And I have introduced the Strengthening Cybersecurity for the Financial Sector Act of 2022 to address these regulatory gaps at the NCUA.

Ms. Lay, the NCUA report concludes that the NCUA's lack of authority over third-party vendors is a growing regulatory blind spot and has the potential to trigger cascading consequences throughout the credit union industry and the financial services sector that may

result in significant losses to the NCUA. Can you elaborate a little bit on this issue?

Ms. LAY. Yes. Thank you for the question, Chairman Foster. Currently, the NCUA does not have examination authority over third-party vendors, and so we are unable to implement corrective action on any third-party vendor if we find issues. We do go into third-party vendors of credit union service organizations voluntarily and can provide corrective or recommendations for corrective action. However, we have had instances where those third-party vendors of credit union service organizations do not respond to that corrective action that we put in place. We have a number of small credit unions. Over two-thirds of our credit unions are less than \$100 million in assets, and they really would rely on our ability to provide help to them with our due diligence of third-party vendors if we could have this third-party vendor authority.

Chairman FOSTER. Thank you, and I will probably get back to this issue in the follow-up questions. Ms. Hall and Ms. Lay, how do you handle the whole issue of explainability and—well, actually all of our witnesses? During the financial crisis, I talked to some gentlemen who had been running some of the banks that tragically failed and asked them what it was like as the regulators came in and closed their bank, and they at least knew the formulas and the tasks that they were failing. But if someone comes in to your bank or credit union and says, I'm sorry, our neural network predicts that you are going to fail, how do you explain this, and how do you handle that whole problem with explainability at all levels?

Ms. Hall, do you want to take a swing at that?

Ms. HALL. Chairman Foster, thank you for the question. I think that explainability in AI is something that can be challenging, particularly for community financial institutions. And ultimately, over time, we just have to continue to prove out what has been established by the theories. For instance, in Montana, we were one of a few States that did not have any bank failures during the financial crisis, despite some of those AI predictive technologies showing that we would have bank failures.

And so, I think that we have to focus on the data inputs in order to ensure that those inputs are actually reflective of expected outcomes, and I think that requires constant change. And I think that is where AI could really help financial regulators if there is an effective feedback loop. I think that, in a lot of ways, the regulatory agencies have struggled with that feedback loop and plugging back in what the ultimate outcome was in order to determine whether the models themselves worked.

Chairman FOSTER. Thank you, and my time has essentially expired. So, I will recognize the ranking member for 5 minutes.

Mr. GONZALEZ OF OHIO. Thank you, Chairman Foster, and thanks again to our panel for being here and for your testimonies.

As I mentioned in my opening statement, I believe we need to be doing more to encourage the use of AI for regulatory purposes, both by the financial institutions themselves, but also within our regulatory agencies.

Ms. Lay, I am going to start my questioning with you. In your testimony, you discussed that the NCUA is investigating the use of natural language processing, which transforms unstructured

data into structured data, increasing the uses and applicability of data. What are the barriers facing the NCUA in implementing this technology at present?

Ms. LAY. Thank you for the question. I believe that one of the barriers we face is that AI is just very expensive and those costs would fall to our credit unions as they pay for our budget. Technology, AI, is very expensive. I think another barrier is just the fact that many of our credit unions are less than \$100 million in assets, and so they also don't have the sophistication sometimes and the level of staff to be able to adopt these technologies. That would also be a barrier.

Mr. GONZALEZ OF OHIO. What, if anything, are you all able to do to help mitigate those barriers? Like, the cost would make sense to me, for sure, but at the same time, if all of the bigger players are adopting the use of some of these technologies, we obviously don't want our credit unions to fall behind on that front. What, if anything, are you all able to do to help mitigate that cost issue?

Ms. LAY. One of the things I think that we could do is, if the agency were granted a third-party vendor authority, that would allow us to conduct examinations of any third-party vendors that credit unions would be using to implement artificial intelligence technologies. And I think that would assist our credit unions in being able to have the ability to see our reports of examination of those third-party vendors and assist them in their due diligence process.

Mr. GONZALEZ OF OHIO. Got it. So, help them on the front end in terms of the diligence side. That makes a lot of sense.

Ms. RUSU. I am going to switch to you. On this committee, we often talk about the concerns of algorithmic bias and the potential impact it could have on decision-making processes. How do you all handle that in the U.K.? I am just curious, because we talk about it a lot, and I would just be curious for more of an international perspective on that issue specifically.

Ms. RUSU. Sure. Thank you for the question, Ranking Member Gonzalez. In the U.K., I would clarify that there is a distinction between discrimination and bias. In the concept of algorithmic bias, we think about whether or not groups could be disproportionately impacted, primarily through bias that would exist in the underlying data. And I think it is important, and as you referenced in your earlier opening remarks, in terms of general model risk management, you have to control both the inputs that go into the model as well as the outputs. And that is how we are thinking about bias and algorithms. We understand the complexity and the challenges in understanding how bias and algorithms can lead to unfair outcomes that might privilege one group of users over another.

Mr. GONZALEZ OF OHIO. Thank you. And, Mr. Greenfield, picking up on that line of questioning, how are regulators working with private industry to prevent the use of biases in their models?

Mr. GREENFIELD. Through the ongoing supervision process with both financial institutions as well as our work with many of the banks' service providers, we are very focused on banks having effective risk management and governance in place for the use of these models, which will include controls for the model development, validation of the model, and testing of the model, both ini-

tially and when in production. But what is also very important is continued oversight of the model over time as assumptions change and data quality can change over time. We very much look at how that is being monitored, and those outcomes are very, very closely monitored. We also engage in those discussions with financial institutions as well as the model developers and we put out guidance such as the model risk management guidance that the banking agencies have been using for some time now.

Mr. GONZALEZ OF OHIO. Thank you, and I yield back.

Chairman FOSTER. Thank you. The gentlewoman from North Carolina, Ms. Adams, is now recognized for 5 minutes.

Ms. ADAMS. Thank you very much, Mr. Chairman, and thank you to our witnesses, and to our ranking member as well.

First of all, the Bank Secrecy Act (BSA) regulatory failures and penalties over the last 10 years have been due to a failure to detect and report suspicious activity, among other violations. I hear regularly that financial institutions, especially smaller entities, are both accountable for and at the mercy of the RegTech service providers.

So, Mr. Greenfield, Ms. Lay, and Ms. Hall, in that order, for Bank Secrecy Act/Anti-Money Laundering (BSA/AML) compliance, if not in other areas, are your agencies' oversight activities appropriately balanced? Mr. Greenfield, first.

Mr. GREENFIELD. Yes, I believe so. We are very focused on how banks are setting up their risk management compliance frameworks to manage the risk, and, as mentioned earlier, we take a very risk-based approach. So, depending on the size and complexity of the institution and the services it offers, that level of oversight and that level of risk management supervision would be commensurate with the activities of the bank.

Ms. ADAMS. Go ahead. Finish.

Mr. GREENFIELD. Okay. We do engage in ongoing communications. And we do encourage, especially smaller community banks, to work together to be able to leverage services more effectively, more efficiently, and more economically, and also do focus on the service providers to make sure that they are providing that level of service to those banks.

Ms. ADAMS. Thank you. Ms. Lay, what would you say about that?

Ms. LAY. Yes, thank you for that question. BSA and AML, for our smaller credit unions, is definitely burdensome and something that we know that they absolutely need to follow. For our smaller credit unions that only have one or maybe two staff persons, they will need to bring in artificial intelligence to help them with that compliance could certainly be a benefit. I think the agency—

Ms. ADAMS. Okay. Thanks. How would you respond, Ms. Hall?

Ms. HALL. Thank you so much for the question. I will just say, first of all, State regulators really appreciated Congress enacting BSA reform that supports greater use of technologies. State regulators supervise a large percentage of smaller banks and smaller credit unions, and BSA/AML compliance certainly is a tremendous cost to them, often without a solid feedback loop to let them know how that information is being used. BSA/AML is a perfect example of where AI could be really helpful, because AI is really good at anomaly detection. However, what we really need is a strong feedback loop with law enforcement in the Federal agencies in order to

improve that AI and make it more accessible to smaller institutions in order to help with the costs of BSA compliance.

Ms. ADAMS. Okay. Let me briefly ask each of you, should RegTech firms themselves regulate or engage in financial institution oversight in a different manner? Mr. Greenfield?

Mr. GREENFIELD. If I understand the question, should RegTech firms be engaged with banks in a different manner or oversight of—

Ms. ADAMS. Correct.

Mr. GREENFIELD. We believe RegTech firms should be in communication with their client base, which would be the financial institutions, in meeting their needs to ensure compliance in an economical and efficient manner. We have conversations with—

Ms. ADAMS. Okay. I just have a few more seconds left. So, Ms. Lay and Ms. Hall, I want to at least get a response from each of you as well.

Ms. LAY. We are in the early stages of looking at AI for RegTech. We do believe that we would need to consult and speak with many of our credit unions in the industry before—

Ms. ADAMS. Okay. Ms. Hall?

Ms. HALL. Representative Adams, State regulators would urge passage of H.R. 2270, which would allow greater coordination and information sharing between the Federal and State regulators on third-party service provider exams. That bill is working its way through Congress. And a lot of States have their own laws that say they can regulate these third-party service providers, but the Federal agencies are unsure as to how much they can coordinate and share information with us. And so, that bill would really go a long way to helping to ensure that there is good coordination and information.

Ms. ADAMS. Thank you, ma'am. I am out of time, and, Mr. Chairman, I yield back.

Chairman FOSTER. Thank you. The gentleman from Georgia, Mr. Loudermilk, is now recognized for 5 minutes.

Mr. LOUDERMILK. Thank you, Mr. Chairman, and I thank everyone on the panel for being here. Some of my colleagues and even other observers have raised concerns recently that artificial intelligence and machine learning can exacerbate bias. However, I believe that if used properly, artificial intelligence and machine learning can actually be used to reduce unfair bias. Some of the essential components for obtaining unbiased results are through record-keeping of what goes into algorithms: robust testing and strong risk management.

Ms. Hall, are there potential risks with using artificial intelligence exclusive to AI? Are they inherent to any model risk management framework?

Ms. HALL. Congressman, I believe that bias is always a part of any kind of model and predictive modeling. I do think that there is the capacity for machine learning hopefully to eliminate that bias faster than we have been able to eliminate it in humans themselves. If there are appropriate feedback loops, if there is appropriate information and data gathering there, I do believe that machine learning could help to eliminate that bias readily as long as there are appropriate feedback loops.

Mr. LOUDERMILK. So in reality, bias exists everywhere. It is not just in the artificial intelligence, but with proper testing, checking, and data analysis, you believe we can eliminate, for the most part, unfair bias?

Ms. HALL. Congressman, I am not an AI or machine-learning expert myself, but I certainly would believe that it is faster than humans, as we have proven as humans to not be all that fast in our bias elimination. I would think that machine learning, with evidence showing the actual outcomes, could potentially be much quicker in eliminating that bias. And I don't think that there is a way to necessarily eliminate that bias on the very front end, but hopefully, the learning process of machines is faster than our own.

Mr. LOUDERMILK. Okay. And I would submit that there is inherent bias in human opinion and decisions, and you can eliminate that through the machine if you have the proper data.

Mr. Greenfield, can you describe how the existing bank regulatory structure already accounts for model risks associated with AI?

Mr. GREENFIELD. Yes. I was going to say, so we do. Yes, we have extensive experience in history with model risk management. We have supervisory guidance that was jointly put out by the banking agencies that provide some expectations for banks as it relates to risk management, governance, testing, and validation, control, and oversight of these models. We have examination programs that focus on this as well as we take an integrated supervision approach that when assessing AI or model risk management within financial institutions, we will bring in Ph.D. economists, subject matter or topic experts, whether it be on fair lending, credit underwriting, or whatever the activity being conducted, as well as technology experts that work together in order to identify potential risks or concerns with model risk management and communicate that to the financial institution with expectations for corrective action.

Mr. LOUDERMILK. When you do bring that to the attention of the financial services, business, or organizations, are they examined and supervised in a way that would require them to address these risks before they go forward? In other words, does the government oversee how they address those?

Mr. GREENFIELD. Yes, we have a number of supervisory tools available to us, ranging from matters requiring attention and reports of examination to enforcement actions. But when we identify deficiencies, we will require corrective action and follow-up, and follow-through to ensure it has been done effectively.

Mr. LOUDERMILK. Okay. Thank you. Ms. Lay, I have been concerned about the government's resistance to adopt certain technology. In fact, the FDIC CIO resignation a few months ago was alarming because he addressed the resistance to change. If financial regulatory agencies are technologically stagnant, doesn't that make it difficult to keep up with the changing nature of the companies they regulate?

Ms. LAY. We agree the ability for NCUA and our credit unions to adapt to new financial technologies is very important. The NCUA does not want to hamper innovation in our agency or in our credit union industry. One of the things that I have testified here today is that many of our credit unions are small, less than \$100

million in assets. And we will need to rely on artificial intelligence or rely on third-party vendors to get into the artificial intelligence space. So, for the agency to have third-party vendor authority to help those credit unions with our due diligence for those companies would be very helpful for the agency. I will just add that we have been going through a technology modernization at the NCUA for the past 5 years, and our NCUA Board and executive leadership have been very supportive of that modernization.

Mr. LOUDERMILK. Thank you. I yield back.

Chairman FOSTER. Thank you. The gentleman from Massachusetts, Mr. Auchincloss, who is also the Vice Chair of the Full Committee, is now recognized for 5 minutes.

Mr. AUCHINCLOSS. Thank you, Mr. Chairman, for organizing this important hearing.

My question is for Mr. Greenfield and for Ms. Rusu, but other witnesses are welcome to jump in, too. In these last 10 days, we have seen that algorithmic stablecoins are not so stable. And it is clear that we are going to need both updated auditing and disclosure regulation from Congress for the stablecoin industry, but also for regulators to be able to track the redeemability of stablecoins, if these continue to be an important part of the modern economy.

Mr. Greenfield, while it is not quite AI—obviously it is deep tech—what tools does the OCC have at its disposal to be monitoring the redeemability and liquidity of algorithmic stablecoins?

And, Ms. Rusu, knowing that the United Kingdom has been really at the forefront of much of this legislation in the crypto space, what advice might you offer us here in the United States on this front?

Mr. GREENFIELD. I will start off by just stating that I am not aware of any banks directly dealing with algorithmic stablecoins. However, as you note, it is very much a key topic, and the OCC is very focused on the development and use of stablecoins throughout the financial sector. We do have a number of policy initiatives and research underway looking at the use of crypto assets throughout the financial sector and within the national banking system. Our Office of Innovation is very focused on this, and we are currently engaged in what we have referred to as crypto policy sprints with FDIC and Federal Reserve colleagues.

Mr. AUCHINCLOSS. Mr. Greenfield, if the OCC were vested by Congress with the authority and the mandate to supervise stablecoins, both algorithmic and non-fiat-backed, is it within the capabilities of the OCC to do that?

Mr. GREENFIELD. We have put out recommendations on a framework for stablecoins as part of the Presidential Working Group report that was published last year. It is something that we are very focused on developing, and Acting Comptroller Hsu has spoken extensively on stablecoins. So, it is something that we are very focused on and looking at what a potential regulatory framework would look like.

Mr. AUCHINCLOSS. And you think the OCC has an important part to play in that?

Mr. GREENFIELD. Yes.

Mr. AUCHINCLOSS. Ms. Rusu?

Ms. RUSU. Yes. Thank you for the question. As you know, we do not yet regulate assets except through the anti-money laundering regulation, but we are following up on this area. And this week, I held the first crypto policy sprint, and we considered three problem statements around crypto asset disclosures to investors to address the inadequacy of information shared. We looked at centralized versus decentralized regulation approaches and gaps in the existing regulatory framework for custodians and the complexities of ownership around crypto assets, and we expect to share the findings from the policy sprint later this summer.

We have also started a project using a web scraper to identify websites that are promoting crypto assets and using text analysis to identify risk indicators on the sites. So, just recognizing that it is a complex area for regulation and the algorithms involved share all of the same complexities that AI algorithms share as well and recognize—

Mr. AUCHINCLOSS. Ms. Rusu, is it the opinion of the majority of U.K. financial regulators that algorithmic stablecoins have a place at all in a stablecoin ecosystem, or are you coalescing behind only fiat-backed stablecoins?

Ms. RUSU. I don't think we have reached a decision yet on that point, but we are certainly looking at all of the different categories of crypto assets.

Mr. AUCHINCLOSS. And, Mr. Greenfield, do you have an opinion on that question?

Mr. GREENFIELD. To the point of our focus on stablecoins, we have been very focused on understanding the transparency in reserves, redeemability issues, as you have noted, as well as looking at the importance of having liquid assets in reserve as part of the stablecoin framework.

Mr. AUCHINCLOSS. Do you think that there is a role for algorithmic alongside fiat-backed stablecoins, or is that a to-be-determined question?

Mr. GREENFIELD. I believe that is to be determined. That is something as we look at the development—

Mr. AUCHINCLOSS. I yield back my time, Mr. Chairman.

Chairman FOSTER. Thank you. And now, we will begin our second and final round of questions here.

Ms. Rusu, one of the most interesting links in your written testimony was dealing with efforts towards what is called federated learning. This addresses a problem that occurs really at all levels of financial regulation, where regulators have access to the detailed information on individual entities that they regulate. They would like to share that information with sibling regulators in other States or other countries, but privacy concerns prevents anything more than very general trends. And the federated learning, as I understand it, is an attempt to use access to encrypted datasets and to train the neural networks across regulatory boundaries or even national boundaries, and there is a potential solution to this.

My question is, is this viewed as something that is really ready for prime time? Are there examples of real-world implementation of federated learning between different regulators, or does this feel like something that is at the talking stage?

Ms. RUSU. Thank you for the question, Chairman Foster. We are participating with other regulatory bodies and looking at, for example, the Digital Regulation Cooperation Forum (DRCF), to share learning and approaches on this. We are also looking at AML through the course of TechSprints, and we are focused on building solutions and sharing common approaches.

Chairman FOSTER. Okay. Are there any examples that any of our witnesses are aware of, where that is being looked at in detail in the U.S.? Going once, going twice?

[No response.]

Chairman FOSTER. Okay. I think that is a very promising area, which, if it works technically, is going to really solve a lot of the political problems with data sharing across national boundaries. There was also discussion, Ms. Rusu, in some of the struggles with determining ultimate beneficial ownership and how that works. The heart of that is the issue of having a secured digital ID for market participants that works across national boundaries. During all the discussions in your TechSprints and so on, is there any discussion of what amounts to a crypto driver's license or something that would allow participants to anonymously identify themselves in a way that the regulators could see, but market participants could not?

Ms. RUSU. I know that some of the TechSprints have looked at privacy-enhancing techniques (PETs), as well as different types of encryption, like homomorphic encryption, and I think certainly there is a lot of work to be done in that area. I think we have some more focused areas coming up in TechSprints later this year to delve into some of that. And I would say that I also saw some of those solutions, or issues that you referenced in terms of ownership, were addressed this week in the 2-day crypto sprint that we held. So, it is something that we don't have an answer to, but we are certainly investigating.

Chairman FOSTER. Okay. Were there any other high-level conclusions from that work?

Ms. RUSU. In terms of the crypto, I think we will be in a position to share some of the findings later this summer.

Chairman FOSTER. Thank you. Because the whole issue with crypto, and secure digital identity, and synonymous, but legally traced, but yet legally traceable access to crypto transactions is really, at least in the U.S., I think it is at the heart of the discussion going on right now.

I will now recognize the ranking member of the task force, Mr. Gonzalez, for 5 minutes.

Mr. GONZALEZ OF OHIO. Thank you. Mr. Greenfield, let's start with you. I have heard from advocates for greater use of emerging technology in the regulatory system and also from the private industry that it is sometimes difficult for regulators to work with private industry on testing and acquiring new technology for pilot programs. Is this something that you have experienced in the OCC, and what ideas might you have to help solve that?

Mr. GREENFIELD. Sure. This is something that we have looked at extensively. Again, we support responsible innovation in the banking industry, and part of having that innovation is institutions' and industries' ability to develop and test new products and services.

We do have supervisory guidance that helps set expectations for banks' engagement in new or modified products and services. It talks about the importance of risk management, governance, stakeholder review, as part of these processes.

Our Office of Innovation is very engaged, not just with financial institutions, but many of the FinTech and emerging tech companies that are helping to develop these products and services, and bringing them in for one-on-one discussions or as part of office hours to discuss what it is like to operate within a banking environment, expectations around management control, and to really respond to their questions and allow them to better develop the products and services that they are going to be offering to the financial institutions that we supervise.

Mr. GONZALEZ OF OHIO. Thank you. I think that makes a lot of sense, and, hopefully, those interactions are done in a productive way. I know for a lot of emerging tech companies, there is a fear of coming to Washington and working with regulators because what you will hear oftentimes is that some of these conversations turned into a predicate for an investigation, when they were really just looking to get some simple answers. That wasn't an accusation, by the way. I was just sharing observations about conversations I have had.

Mr. GREENFIELD. Yes. It is one of the reasons why we have our Office of Innovation that is separate from our supervision group. It is an open invitation not only to come in to D.C. and speak with us, but there also will be office hours in many of the tech cities around the country.

Mr. GONZALEZ OF OHIO. That is great. I am staying with you, Mr. Greenfield. One of the more interesting applications of AI and machine learning, in my view, is the ability to crack down on illicit finance. Can you discuss how banks are currently using these technologies to better track financial crimes, and what more our regulatory agencies can be doing to promote the use of this technology?

Mr. GREENFIELD. Sure. I think one of the biggest areas of RegTech development that we have seen are advances in the products and services developed both by banks as well as being offered by third-party service providers to allow for better and more efficient identification and determination on suspicious activity, and for ensuring adherence with bank secrecy and anti-money laundering laws. Banks have often been challenged with, again, going back to the fundamentals of validation and testing, when it comes to ensuring not only that the model is picking up, but that it is also not overreaching and having a lot of false positives, and really being able to adjust those models over time.

And it is something that banks continue to have challenges with, but we are seeing a lot of advancement in this area. And there is a lot of opportunity, because as many of the other panelists here today have commented, anti-money laundering laws are there for a reason, and they are very important. And it is important for adherence, but they can be challenging and burdensome, especially for smaller community banks. And use of these technologies can help provide the opportunity that we enforce these laws, and enforce them as they are intended, but also reduce the burden.

Mr. GONZALEZ OF OHIO. Thank you. That is all I have. I yield back.

Chairman FOSTER. Thank you, and that will conclude our second round of questions. I would like to thank the witnesses for their testimony today.

The Chair notes that some Members may have additional questions for these witnesses, which they may wish to submit in writing. Without objection, the hearing record will remain open for 5 legislative days for Members to submit written questions to these witnesses and to place their responses in the record. Also, without objection, Members will have 5 legislative days to submit extraneous materials to the Chair for inclusion in the record.

And this hearing is adjourned.

[Whereupon, at 10:05 a.m., the hearing was adjourned.]

A P P E N D I X

May 13, 2022

For Release Upon Delivery
9:00 am

STATEMENT OF
KEVIN GREENFIELD
DEPUTY COMPTROLLER FOR OPERATIONAL RISK POLICY
OFFICE OF THE COMPTROLLER OF THE CURRENCY
before the
TASK FORCE ON ARTIFICIAL INTELLIGENCE
COMMITTEE ON FINANCIAL SERVICES
UNITED STATES HOUSE OF REPRESENTATIVES
May 13, 2022

Statement Required by 12 U.S.C. § 250:
The views expressed herein are those of the Office of the Comptroller of the Currency and
do not necessarily represent the views of the President.

Chairman Foster, Ranking Member Gonzalez, and members of the Task Force, thank you for the opportunity to appear today to discuss Artificial Intelligence (AI) approaches used by national banks, federal savings associations, and federal branches and agencies of foreign banks (hereafter “banks”) supervised by the Office of the Comptroller of the Currency (OCC). I appreciate this invitation to discuss the opportunities, benefits and challenges AI presents for banks and the OCC’s approach to supervising those activities. For purposes of this testimony, AI is broadly defined as the application of computational tools to address tasks traditionally requiring human analysis and decision making.

I serve as the OCC’s Deputy Comptroller for Operational Risk Policy, and I am responsible for overseeing the development of policy and examination procedures addressing bank operational risk. This includes bank information technology and cybersecurity, critical infrastructure, payments systems, third party risk, and operational risk governance. In this capacity, I oversee staff whose duties include understanding and monitoring AI risks and benefits and I participate in the development of examiner and industry policy and guidance. I also participate in the OCC’s international work in this area by serving as the Co-Chair of the Basel Committee on Bank Supervision’s Financial Technology Group which coordinates the sharing of regulatory approaches and practices across foreign jurisdictions, including those related to the use of AI.

In my testimony today, I will describe the agency’s approach to responsible innovation and our supervisory expectations for banks’ use of AI, including regulatory compliance. I will also discuss our coordinated interagency approach related to banks’ use of AI, as evidenced by the 2021 *Request for Information and Comment on Financial Institutions’ Use of AI, including*

*Machine Learning*¹ as a means to gather stakeholders' views on how banks can use AI and the key benefits and risks associated with its use. Finally, I will briefly discuss the OCC's ongoing efforts to update our own technological framework to support our bank supervision mandate.

OCC Supports Responsible Innovation

The OCC charters, supervises, and regulates 1,109 banks which cover a broad spectrum of asset sizes and business models. These banks range in size from very small community banks to the largest, most globally active banks operating in the United States. The vast majority have less than \$1 billion in assets, while more than 60 have greater than \$10 billion in assets. Together, they hold \$14.5 trillion in assets—almost 70 percent of all the assets of commercial U.S. banks.

Technological changes and rapidly evolving consumer preferences are reshaping the financial services industry and creating new opportunities to provide consumers, businesses, and communities with more access to and options for products and services. Over the years, we have adapted our supervisory approach to address these technological innovations by banks, including an increase in banks' use of AI. AI can provide efficiencies in operations and back office functions, support lending decisions, and improve a bank's ability to monitor compliance with regulatory requirements including under the Bank Secrecy Act (BSA) and the Home Mortgage Disclosure Act. However, the OCC also recognizes that there are risks associated with the use of AI and vigilantly supervises banks to ensure their use of AI does not adversely affect safety and soundness, impede fair access to financial services, or result in unfair treatment of consumers, or violate applicable laws and regulations.

¹ See OCC Bulletin 2021-17, issued March 31, 2021, [Artificial Intelligence: Request for Information on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning | OCC](#)

The OCC promotes responsible innovation in the federal banking industry to expand access to credit and capital, improve operations, and support full and fair participation in America's banking system. In March 2016, the OCC published a paper entitled *Supporting Responsible Innovation in the Federal Banking System: An OCC Perspective*² which outlines our approach to responsible innovation. Our guiding principles for responsible innovation emphasize safety and soundness, compliance with applicable laws and regulations, and fair access and fair treatment. The OCC expects bank management and each bank's board of directors to understand the impact and associated risks of enabling technologies, including on their bank's financial performance, strategic planning process, risk profiles, and traditional banking models.

The OCC established a dedicated Office of Innovation in 2016. The Office of Innovation conducts external meetings and outreach activities, including formal office hours, with industry stakeholders, including banks and nonbanks, to promote responsible innovation in the federal banking system. This also allows the OCC to stay abreast of current trends and developments and improve the OCC's ability to respond more quickly, efficiently, and effectively to inquiries regarding new products and services. In addition, the Office of Innovation fosters internal OCC staff awareness of responsible innovation and emerging trends, and enhances the skills of examiners.

AI Presents Opportunities, Benefits and Risks

As the premise of this hearing suggests, the use of technology to aid banks in meeting regulatory requirements is not a new concept and it is not surprising to see new and innovative

²See OCC issuance March 2016, [Supporting Responsible Innovation in the Federal Banking System: An OCC Perspective | OCC](#)

approaches, such as AI, be applied for this purpose. Such use of innovative technologies to facilitate regulatory compliance, commonly referred to as RegTech, can increase bank productivity and improve accuracy in analysis, decision making to support risk management and regulatory compliance monitoring and internal controls. AI has the potential to strengthen safety and soundness, enhance consumer protections, improve the effectiveness of compliance functions, and increase fairness in access to the financial services when implemented in an effective manner. Examples of where AI can be used to enhance bank customer services range from the use of chat-bots that aid in directing customer inquiries and assisting with online account openings, assist in selection of beneficial products and investments for consumers, to supporting more efficient credit underwriting, all of which have the potential to promote greater access to banking services by underserved communities.

Technological innovations have also increased access to greater computing power, including through cloud computing resources, coupled with the availability of increasing volumes of structured and unstructured data. This has created an environment where the resources needed to develop AI tools and services are more widely available to banks of all sizes. This can be seen in the development of advanced tools to improve the quality of fraud prevention controls, increase the effectiveness of anti-money laundering and the countering of terrorist financing (AML/CFT) monitoring activities, and help to identify and mitigate the risk of fair lending violations. AI can augment overall risk management, compliance monitoring and management, and internal controls. However, the key to ensuring these benefits are achieved and unintended risks are not introduced is by validating that banks have effective governance processes and controls in place for the planning, implementation and operation of these innovative solutions.

One promising use of AI relates to using alternative data, such as utility or rent payments as opposed to loan payments, in traditional credit models or AI applications. The OCC has championed the use of innovative technologies and approaches to expand access to credit as observed in our Project REACH (Roundtable for Economic Access and Change) initiative. Project REACH brings together leaders from banking, business, technology, and national civil rights organizations to reduce barriers that prevent full, equal, and fair participation in the nation's economy. One of the project's workstreams is focused on addressing the challenge of credit invisibles and promoting partnerships that open the door to mainstream financial services for economically-disadvantaged communities. Participants in this workstream are working to build a utility that will enable financial institutions to share customer permissioned data – including alternative data that includes information not typically found in the consumer's credit files – to be the basis of credit decisions for people who previously lacked opportunities. Using alternative data in AI applications may improve the speed and accuracy of credit decisions and may help firms evaluate the creditworthiness of consumers who may not otherwise obtain credit in the mainstream credit system.

More broadly, banks have been employing mathematical models to support operations for some time and are now introducing more sophisticated AI tools to support activities, including those mentioned above, in addition to trading algorithms and automation, financial marketing analysis, strengthening cybersecurity, and suspicious activity monitoring and customer due diligence. While we have seen many large banks develop these tools internally, third party firms are also increasingly offering sophisticated products and services to banks of all sizes.

The increasing availability of third-party firms and fintechs allows banks to access specialized skills and benefits from economies of scale in the delivery of products and services.

The OCC encourages the responsible use of third-party products and services, with appropriate risk management and controls, by banks of all sizes and business models. However, third-party relationships can increase operational complexity or present unique challenges, and banks' use of third parties does not diminish the responsibility of the board and management to implement and operate new products and services in a safe and sound manner and in compliance with applicable laws and regulations.³ Where appropriate, the OCC has the authority to conduct examinations of services provided by third parties under the Bank Service Company Act, which could, depending on the facts and circumstances, include the banking services supported by AI.

The OCC also remains focused on the potential risks of adverse outcomes if banks' use of AI is not properly managed and controlled. Potential adverse outcomes can be caused by poorly designed underlying mathematical models, faulty data, changes in model assumptions over time, inadequate model validation or testing, and limited human oversight, as well as the absence of adequate planning and due diligence in utilizing AI from third parties. Key risks in the use of AI include:

- ***Explainability.*** The extent to which AI decisioning processes are reasonably understood and bank personnel can explain outcomes is critical. Lack of explainability can hinder bank management's understanding of the conceptual soundness of the technology, which may inhibit management's ability to express credible challenge to models used or understand the quality of the theory, design, methodology, data, or testing. This may also inhibit the bank's ability to confirm that an AI approach is appropriate for the intended use. Lack of

³ The OCC, Board of Governors of the Federal Reserve System, and FDIC requested comment on proposed interagency guidance to help banks manage risks associated with third-party relationships on July 13, 2021. The comment period closed on October 18, 2021 and the agencies are still reviewing the comments received. This would update and replace the current OCC Bulletin 2013-29 Third Party Relationships: Risk Management Guidance.

explainability can also make compliance with laws and regulations, including consumer protection requirements, more challenging. For example, should a bank management team not adequately validate or attain reasonable assurance that a given AI model or tool being used to support credit underwriting is sound, a bias inherent in a given model's algorithmic design could exist. Such bias could ultimately perpetuate discriminatory credit decisioning practices and go unnoticed until a well-established pattern and practice has become evident and caused harm to consumers or caused the bank to be in noncompliance with applicable consumer protection requirements. Banks will find it important to reassess and, where necessary, enhance their risk management frameworks, including validation and internal audit, to achieve appropriate levels of explainability and manage inherent and residual risk appropriately for use of AI and other sophisticated analytical tools.

- **Data Management.** Understanding data origins, use, and governance when adopting traditional models, advanced analytics, and AI is also critical. Data analytics and governance are particularly important when AI involves dynamic updating or algorithms that identify patterns and correlations in training the data⁴ without human context or intervention, and then uses that information to generate predictions or categorizations. Because the AI algorithm is dependent upon the quality of the data used and effectiveness of training, an AI system generally reflects the limitations of that dataset. To an even greater extent than other systems, AI may perpetuate or even amplify bias or inaccuracies inherent in the data or make incorrect predictions if a data set is incomplete, non-representative or otherwise flawed. This potential for unintended or illegal outcomes increases the importance of enhanced

⁴ Training data typically refers to data sets used to teach and confirm AI and machine learning models on expected outcomes.

understanding, monitoring and review of AI systems that are used for customer-focused activities such as credit underwriting.

- ***Privacy and Security.*** Consumer privacy and data security are also important. The use of AI is predicated on access to large volumes of data, much of which may be sensitive. Banks must comply with applicable privacy and information security requirements when using AI. We also expect banks to practice sound cyber hygiene and maintain effective cybersecurity practices to prevent or limit the impact of corrupted and contaminated data that may compromise the AI application and to safeguard sensitive data against breaches. The OCC recognizes the paramount importance of protecting sensitive data and consumer privacy, particularly given the use of consumer data and expanded data sets in some AI applications.
- ***Third-Party Risk.*** As previously noted, many banks rely on third party vendors, service providers and expertise for the development and implementation of AI tools and services. As part of an effective third-party risk management program, banks are expected to have robust due diligence, effective contract management and ongoing oversight of third parties based on the criticality of the services being provided.⁵ This includes ensuring effective controls over aspects relevant to many AI services, including use of cloud-based entities, availability of documentation on models used, establishing roles and responsibilities and defining data ownership and permitted uses, security, privacy and limitations of any data that is shared with or exchanged among parties and other key governance expectations for the delivery of AI services. It is important for banks to monitor a third party's performance over time, and have controls to ensure data is used consistent with what the consumer originally

⁵ See OCC Bulletin 2013-29, October 30, 2013, [“Third-Party Relationships: Risk Management Guidance”](#)

permissioned, and that the results of independent assessments are available to assess if the AI service performs as intended.

While this is not an exhaustive list of risks, it does reflect key issues banks should manage appropriately in their use of AI. Failure to adequately manage risks can result in adverse outcomes for the bank or its customers that may include, but are not limited to, erroneous results, misinformed decisions, unauthorized access to customer information, failed transactions, or violations of consumer protection requirements, including fair lending laws.

OCC Approach to Supervision of AI

Many banks have plans to increase use of AI technologies and are investing or considering investing in AI research and applications to automate, augment, or replicate human analysis and decision-making tasks. As such, the OCC is continuing to update supervisory guidance, examination programs and examiner skills to respond to AI's growing use.

The OCC follows a risk-based supervision model focused on safe, sound, and fair banking practices, as well as compliance with laws and regulations, including fair lending and other consumer protection requirements. This risk-based approach includes developing supervisory strategies based upon an individual bank's risk profile and examiners' review of new, modified, or expanded products and services (collectively, new activities). Examiners generally consider new activities' effect on banks' risk profiles and the effectiveness of banks' governance and risk management systems as a key step in scoping each bank examination.⁶

⁶ In 2014, the OCC issued guidelines to establish minimum standards for the design and implementation of a risk governance framework that applies to any bank with total consolidated assets equal to or greater than \$50 billion. See OCC Guidelines Establishing Heightened Standards for Certain Large Insured national Banks, Insured Federal Savings Associations and Insured Federal Branches 12 C.F.R. Part 30, App. D.

Failure to adequately manage risks can result in adverse outcomes for the bank or to its customers. In order to aid banks with developing robust risk management and governance processes for the implementation of AI solutions, the OCC has published a number of informational resources individually and as part of interagency communications to the industry and examiners.

With respect to the complex mathematical models AI tools and techniques encompass, the OCC, in collaboration with the other federal banking agencies, issued OCC Bulletin 2011-12, “Supervisory Guidance on Model Risk Management Guidance” (MRMG) in 2011.⁷ The bulletin articulated the elements of a sound program for effective management of risks that arise when using quantitative models in bank decision making. It also provides guidance to OCC examiners and regulated institutions on prudent model risk management policies, procedures, practices, and standards.

Many of the AI tools being employed would be considered models under the MRMG. The key principles in the 2011 guidance apply to AI uses today, and the interagency statement is a key resource leveraged by the industry. OCC examiners have significant experience supervising banks’ use of sophisticated mathematical models and we are incorporating AI concepts into several of our recently published guidance and statements.

In October 2017, the OCC issued guidance⁸ to the banking industry addressing the implementation of new activities, such as AI. This guidance stresses the importance of establishing appropriate risk management processes for new activities, such as the introduction

⁷ See [OCC Bulletin 2011-12](#) April 4, 2011, [Sound Practices for Model Risk Management: Supervisory Guidance on Model Risk Management](#)

⁸ See OCC Bulletin 2017-43, [New, Modified, or Expanded Bank Products and Services: Risk Management Principles](#) | OCC

of AI, including the ability to effectively identify, measure, monitor, report and control the risks associated with those activities. Effective risk management includes appropriate due diligence and risk assessment, sufficient and qualified staffing, governance (e.g., approval authority, parameters for use, independent validation processes), and controls. These fundamental practices apply to the risks posed by both AI applications, and the functions that AI supports regardless of whether the AI tool was developed by the bank or obtained through a third party.

More recently, the OCC has coordinated with other agencies and published statements addressing key governance and risk management practices for the use of innovative solutions in specific areas of banking. Examples include:

- *Interagency Statement on the Use of Alternative Data in Credit Underwriting*⁹, issued on December 3, 2019, by the OCC, Board of Governors of the Federal Reserve System, the Consumer Financial Protection Bureau, Federal Deposit Insurance Corporation, and National Credit Union Administration to outline consumer protection implications of the use of alternative data in underwriting, thus highlighting potential benefits and risks. The statement seeks to encourage the responsible use of alternative data; explain that many factors associated with using alternative data, including cash flow data, may increase or decrease consumer protection risks; and explain that a well-designed compliance management program provides for a thorough analysis of relevant consumer protection laws and regulations to ensure firms understand the opportunities, risks, and compliance requirements before using alternative data.

⁹ See [Interagency Statement on the Use of Alternative Data in Credit Underwriting](#)

- *Interagency Statement on Model Risk Management for Bank Systems Supporting BSA/AML Compliance and Request for Information*, issued in April 2021, by the federal banking agencies, in consultation with FinCEN and the NCUA. This joint statement addressed questions from the industry regarding how banks can apply the MRMG when using technology systems or models to comply with BSA/AML laws and regulations. The joint statement clarified that the MRMG provided flexibility for banks when developing, validating, implementing, and updating systems or models used to comply with BSA/AML laws and regulations. While the joint statement responded to industry interest in using advanced models for BSA/AML compliance, it also emphasized the broader point that the practical application of any principle discussed in the MRMG depends, in part, on a bank's reliance on sound risk management.
- *OCC Comptroller's Handbook, Model Risk Management*¹⁰ booklet, published in August 2021. This new examination booklet does not create new guidance but informs and educates examiners on the sound model risk management practices, addressed in the interagency MRMG, that should be assessed during an examination. The booklet also provides information important to planning and coordinating examinations on model risk management, identifying deficient practices, and conducting appropriate follow-up. AI concepts and the relationship to the MRMG are discussed in the booklet.

In addition to providing guidance, the OCC is focused on educating examiners on a wide range of AI uses and risks including risks associated with third parties, information security and resilience, compliance, BSA, credit underwriting, and fair lending and data governance, as part of training courses and other educational resources. The OCC also maintains an Economics and

¹⁰ See [Comptroller's Handbook: Model Risk Management](#).

Risk Analysis Division to ensure appropriate technical expertise and knowledge is available to support ongoing supervision and targeted examinations of banks' use of complex models and AI. This team is staffed with PhDs in economics, finance, and other sciences whose primary objective is to assess the technical strengths and limitations of a broad range of models and innovative tools developed, managed, and used by banks. Key supervisory areas this team focuses on include fair lending, BSA/AML, retail and commercial credit, capital markets and other banking operations. This staff is focused on the ongoing understanding and identification of emerging risks, financial markets, and industry behavior in the use of these sophisticated tools, such as AI, to inform OCC policy and supervisory strategies.

With these resources, the OCC will continue to perform robust supervision of banks' use of AI, whether directly or through a third-party relationship, to ensure banks operate in a safe and sound manner, comply with applicable laws, provide fair access to financial services, and treat their customers fairly. This includes evaluating fair lending concerns and other consumer protection issues such as unfair or deceptive acts or practices. Banks should maintain a well-designed risk management and compliance management program as well as monitor for and identify outcomes that create unwarranted risks or violate consumer protection laws. If these outcomes occur, the OCC has a range of tools available and will take supervisory or enforcement actions as appropriate, to achieve corrective actions and address potential adverse consumer impacts.

Interagency Request for Information

Recognizing the growing interest and use of AI in the banking sector and the potential impact on banks' operations, in 2021, the OCC, in collaboration with the Board of Governors of the Federal Reserve System, Consumer Financial Protection Bureau, Federal Deposit Insurance

Corporation and National Credit Union Administration, issued a *Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, including Machine Learning* (RFI).¹¹ This RFI included specific questions related to these advanced technologies and requested respondents' views on the use of AI by financial institutions in their provision of services to customers and for other business or operational purposes; appropriate governance, risk management, and controls over AI; and any challenges in developing, adopting, and managing AI. The RFI solicited views on the use of AI in financial services to assist in determining whether any clarifications would be helpful for the use of AI in a safe and sound manner and in compliance with applicable laws and regulations, including those related to consumer protection.

More than 70 comments were submitted in response to the RFI. The OCC, working with the other agencies, identified several key themes from the comments, consistent with the risk categories described in this testimony. The OCC has organized workstreams around each of the key themes to complete a gap analysis which will include identifying existing OCC pronouncements applicable to AI and machine learning; outlining potential gaps in guidance, supervision, or regulation; and developing recommendations to address any identified gaps as appropriate. We will continue to engage in interagency discussions to determine next steps.

OCC's Developments in SupTech

While the banks we supervise are using technology to enhance their capabilities, the OCC is also engaged in innovative approaches to enhance the planning and execution of our supervisory responsibilities, risk identification, and policy development. Currently, the OCC is

¹¹See OCC bulletin 2021-17, issued March 31, 2021, [Artificial Intelligence: Request for Information on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning | OCC](#)

engaged in a significant initiative to materially upgrade core supervision systems. The OCC is committed to evaluating and exploring use of advanced technologies, including AI capabilities, that can deliver improved insights to our supervisory, policy staff, and risk analysis teams as part of this supervisory system upgrade. We recognize that these technologies rely on quality data to generate unique insights. Accordingly, the OCC has developed and is in the process of deploying a strong data governance program. Additionally, we are concurrently enhancing our technology architecture which will include upgraded network and security capabilities. These enterprise level initiatives are foundational to the implementation of advanced technologies such as AI. These OCC strategic priorities will help gain examination efficiencies and enhance our ability to identify, measure, and monitor risk in the federal banking system. Finally, the OCC recognizes the importance of to developing and maintaining a robust human capital capability to successfully deploy advanced technical solutions and continues to focus on hiring people with strong analytical capabilities for our supervision analysis staff as well as data scientists.

Conclusion

The OCC supports continuing efforts by national banks and federal savings associations to explore safe and sound uses of AI. While AI opens opportunities and can provide many benefits, we are also mindful of the challenges and risks associated with AI applications and services that we expect our regulated institutions to address. The OCC will continue to conduct robust, risk-based supervision, monitor the industry, and research efforts to keep pace with changes in AI use in the financial sector.

I want to thank the Task Force for its leadership on this important issue and for inviting the OCC to testify today.

38

TESTIMONY OF

MELANIE HALL

COMMISSIONER

MONTANA DIVISION OF BANKING AND FINANCIAL INSTITUTIONS

ON BEHALF OF THE

CONFERENCE OF STATE BANK SUPERVISORS

TO THE

HOUSE FINANCIAL SERVICES COMMITTEE ARTIFICIAL INTELLIGENCE TASK FORCE

OF THE

U.S. HOUSE COMMITTEE ON FINANCIAL SERVICES

HEARING ON

“KEEPING UP WITH THE CODES – USING AI FOR EFFECTIVE REGTECH”

MAY 13, 2022

Introduction

Thank you, Chairman Foster, Ranking Member Gonzalez and distinguished members of the Task Force. My name is Melanie Hall. I am the Commissioner of the Montana Division of Banking and Financial Institutions and serve as the Chair of the Conference of State Bank Supervisors' (CSBS) Board of Directors. CSBS thanks the House Financial Services Committee Artificial Intelligence Task Force for convening this important hearing entitled "Keeping Up with the Codes – Using AI for Effective RegTech."

It is my pleasure to testify on behalf of CSBS on how the state system uses technology for financial supervision today and how we see that shaping financial supervision in the future through Networked Supervision, our design to create a federated state system through greater use of technology and enhanced state-to-state and state-to-federal partnerships.

CSBS is the national organization of state banking and financial regulators that educates and informs policy makers, consumers and other stakeholders about issues affecting states' banking and financial services markets. Representing a network of financial regulators in all 50 U.S. states and territories, CSBS also provides a range of professional development and training programs and serves as a forum for collaborating and developing policies that strengthen financial regulation for the benefit of consumers and communities. CSBS supports state regulators in their mission to ensure safety and soundness; protect consumers; promote economic growth; and foster innovative, responsive supervision.

State regulators charter and supervise approximately 3,981 banks with more than \$7.15 trillion in combined assets, representing 79% of all U.S. banks, half of all small business lending and two-thirds of all agriculture lending in America. State regulators also are the primary licensing authority for nonbank financial services providers, including mortgage lenders, money transmitters and consumer lenders.

CSBS, on behalf of state regulators, operates the Nationwide Multistate Licensing System (NMLS), a regulatory licensing platform that launched in 2008 and is now used for state-licensed nonbank financial services providers in the money services, mortgage, consumer finance and debt industries. In the mortgage space, for example, all 579,000 mortgage loan originators operating in the United States are licensed or registered in NMLS. CSBS continues to enhance its technology platforms as states identify and develop new ways to work together and improve their supervisory capabilities. Two years ago, we launched the State Examination System (SES), an online platform that connects state agencies and the companies they supervise in the examination process.

Technology has been an important part of the state system's evolution and will continue to play a critical role as financial transactions become faster and more complex. To that end, CSBS is focused on finding new and improved ways for our data, technology and analytical products to help state regulators mitigate risk and protect consumers from bad actors. In addition, states

are focused on building the regulatory system of the future through Networked Supervision where technology platforms and data analytics allow 50+ state financial agencies to operate as one seamless supervisory network. We appreciate the opportunity to talk about how we are using “regtech” and “suptech” to strengthen regulation today and to envision the regulatory system of the future.

Specifically, this testimony will address:

- How States are Building a Network of Supervision
- How States use NMLS as a Regtech Resource
- The Evolution of Suptech Supervision through SES
- How States are Improving Supervision Using Data Analytics
- Results of Networked Supervision are Already Being Realized

BUILDING A NETWORK OF SUPERVISION

State regulators have been working together for decades to identify and develop tools and resources to streamline the regulation of multistate companies. That collaboration has steadily increased over the years with changes in the financial services space and advancements in technology. Some of the milestones toward greater collaboration include the Nationwide Cooperative Agreement signed in 1997, NMLS created in 2008, and the Vision 2020 initiative adopted by the states in 2017 that led to Networked Supervision.

CSBS is driving greater technology use through Networked Supervision, a regulatory approach that encourages diversity in size and scope, streamlines licensing, reduces regulatory burden, and enables responsible innovation that benefits consumers and local economies alike. This single strategic approach will evolve the state system to one where communication occurs in real time, knowledge and expertise flows across the states and regulation becomes streamlined throughout the industry.

NMLS, SES and data analytics are critical technologies to enable the state systems’ move toward Network Supervision – both by supporting today’s initiatives and establishing the technological foundation for future collaboration among the states.

How States use NMLS as a Regtech Resource

NMLS emerged at a critical juncture for state supervision and regtech. More than a decade ago, state regulators recognized growing problems in the mortgage industry as bad actors were taking advantage of a lack of regulatory coordination. Working together, states created uniform mortgage loan originator license application forms in 2006. A year later, CSBS began building a common licensing platform to better manage and monitor licensed mortgage lenders, mortgage brokers and individual mortgage loan originators (MLOs) doing business in one or

multiple states. That became the Nationwide Mortgage Licensing System, launched in January 2008 – one of the first common technology platforms used by multiple state regulators in the financial regulatory space.

Congress recognized its value and incorporated what is now called the Nationwide Multistate Licensing System in passage of the Secure and Fair Enforcement for Mortgage Licensing Act of 2008. State regulators began using NMLS in 2012 to license a broader range of nonbank financial services providers.

Today, NMLS is a comprehensive system of licensing for all state-licensed mortgage companies and MLOs, and registration for MLOs working in all depository institutions, including banks and credit unions. Nearly 640,000 companies and individuals use NMLS annually to manage their business licensing or registration. These companies and individuals span the mortgage, consumer finance, debt and money services businesses (MSB) industries – with the mortgage industry accounting for more than 80% of NMLS use.

In addition to serving as a regulatory platform, NMLS has a public facing portal (nmlsconsumeraccess.org) where consumers review individual and company licensing status and publicly available regulatory actions. NMLS also supports state efforts to improve regulatory data and information sharing about nonbank financial services providers. States are using NMLS data to understand and evaluate trends and risks in their regulated industries and to better risk-scope licensing and supervisory priorities and activities.

In 2011, NMLS launched the Mortgage Call Report, which is a quarterly report of originations covering more companies than is covered by the Home Mortgage Disclosure Act. In 2017, we launched the MSB Call Report, which is the first and only nationwide report of MSB information, especially important in understanding the money transmission industry. These reports create a standardized reporting requirement across all participating states that allow for nationwide trend analysis and risk identification. We are also in the early stages of developing a Consumer Finance Call Report.

The data has given state regulators a deeper perspective into the mortgage industry landscape and has helped us identify applications that might require more scrutiny. As a result, state regulators have become more efficient and risk focused. NMLS is useful for federal regulators as well. The Consumer Financial Protection Bureau relies on NMLS to register mortgage loan originators who work in banks and credit unions.

NMLS is an important system, which is why we have begun to rebuild NMLS to advance Networked Supervision. Our goal is to transition all state agencies to the new modernized system. Currently, CSBS is creating common information standards to be agreed upon by all states to support the software development for the modernized system and in the beginning stages of the functional development. Some of the standards that have already been finalized

include: the core data and documents that all MSBs must provide, creating a national standard of information collection; and information standards for the MSBs.

The State Examination System

Building on the success of NMLS as a licensing system, the states developed a new technology platform called the State Examination System. Through SES, launched in February 2020, state regulators have a powerful tool to advance uniformity, efficiency and a less burdensome supervisory process for companies. This secure, end-to-end technology platform is the first nationwide system to bring regulators and companies into the same technology space for examinations, fostering greater transparency throughout supervisory processes.

SES provides a platform for state agencies to conduct exams, investigations, consumer complaints and other supervision work and to share that information within their agency and within the broader state regulatory network. Examiners can use SES for every aspect of the exam, from scheduling to reviewing reports and everything in between. SES allows for greater partnering with other state agencies, reducing the regulatory burden for companies that are supervised by more than one agency. Having a view into the supervisory findings from one agency allows another agency to better allocate resources and schedule their own supervisory activities. State regulators can also use shared exam manuals and procedures, which will help drive uniformity and consistency across the states.

SES improves how state agencies and companies address consumer complaints by providing a single platform for all involved. State agencies can enter and process complaints about financial entities under their supervision. Summaries of all complaints are available to any state regulator using the complaints system. State regulators investigate the complaints and work with the company on behalf of the consumer to resolve the issue. This not only accelerates the process, but it also creates a data set that gives a fuller picture of complaints, reveals areas of potential consumer harm, and enables a risk focused examination approach.

Currently, 49 state agencies have used the system for individual state exams. We continue to improve and expand upon SES, which in 2021 underwent five large-scale enhancements that provide agencies and companies with new or improved functions, leading to a better overall user experience.

IMPROVING SUPERVISION THROUGH DATA ANALYTICS

Data analytics is an important area for CSBS. To support our data and analytics efforts, the CSBS Board chartered a group of regulators known as the Data Analytics Taskforce, which has been exploring options for how states can advance supervision by harnessing real-time, or more-frequent, data.

CSBS has worked with a group of data-savvy state regulators since 2019 to explore how predictive modeling could be used as an early detection tool for bank risk. This year, staff

completed work on the first version of its predictive analytics solution, which includes two models, two predictive dashboards and extensive model documentation, and is slated for a pilot program in 2022.

The CSBS Risk Identification team is comprised of a group of regulators who are leveraging technology to understand emerging trends from examiners in the field. Using the CSBS Risk ID Data Collection Tool, examiners are asked to submit data each quarter on examination trends like frequently upgraded/downgraded categories, examination time trends, topical elements like CECL and LIBOR transition and other emerging risk areas.

RESULTS OF NETWORKED SUPERVISION ARE ALREADY BEING REALIZED

Networked Supervision is a wide-ranging transformation that will require the adoption of common policies and practices to fully realize the benefits of technology. Recognizing Networked Supervision's large scope, the CSBS Board in early 2021 approved eight priorities that both lay the groundwork for future initiatives and further work started under Vision 2020, which includes: the [Model Money Transmission Modernization Act](#); the MSB One Company, One Exam supervisory program; and the Multistate MSB Licensing Agreement.

Last year, the CSBS Board approved the Money Transmitter Model Law for state adoption. This model law establishes one set of regulatory requirements, creating greater clarity for businesses and streamlining the regulation of nationwide payments companies. The law will also inform the MSB experience in the modernized NMLS initiative described above. Several states have begun the process of state-level adoption.

States advanced the MSB One Company, One Exam pilot to a supervisory program by performing single exams for 74 nationwide payments companies, reducing duplicative exams by 78%. States have also reduced the time it takes to license a new MSB by 25% through the Multistate MSB Licensing Agreement (MMLA). Currently, 29 states have adopted the MMLA, and 10 more have committed to join the agreement in 2022 and 2023. By reducing duplicative exams and licensing time, these programs are addressing significant pain points identified by the Vision 2020 Fintech Industry Advisory Panel.

In the mortgage area, and in support of Networked Supervision, SES facilitated a multi-state exam for the One Company, One Exam mortgage pilot. Fifty-one state agencies were part of the pilot, with 27 states participating, nine states leveraging portions of the exam, 10 states accepting and five agreeing to a moratorium. An after-action report on the One Company, One Exam mortgage pilot will yield valuable information to guide more effective and efficient multi-state exams in the coming years.

In addition, states approved prudential standards for nonbank mortgage servicers to coordinate individual state authority into nationwide requirements for financial condition and corporate governance over the fast-growing area of nonbank mortgage servicing. To encourage uniform

coverage, CSBS staff are working with state agencies to understand the path to adoption of common prudential standards in their state.

Networked Supervision requires timely and robust information sharing between state and federal agencies. In addition to developing platforms that support this objective, CSBS and the states are identifying and seeking to modernize the necessary legal underpinnings in state and federal law or processes to promote greater legal certainty to enable a networked approach to regulatory and supervisory information.

The elimination of data silos is critical to achieving more efficient supervisory processes. For example, the primary prudential supervisors of nonbank mortgage servicers - state regulators, the Federal Housing Finance Agency (FHFA) and Ginnie Mae - each have proposed prudential standards. However, our lack of information sharing protocols means that these standards were developed and will be administered in vacuums. The state system has had substantive information sharing platforms in place with the Consumer Financial Protection Bureau (CFPB) since 2011 and with the U.S. Department of Housing and Urban Development (HUD) since 2014. State and federal information sharing is accomplished through memorandums of understanding facilitated by CSBS, and the process works very well. All these agencies collect and use data to monitor nonbanks. The Financial Stability Oversight Council itself has been recommending collaboration between the state system and FHFA since 2014.

In addition to working with our federal counterparts to develop information sharing agreements, state regulators are also asking Congress to support the Bank Service Company Examination Coordination Act (HR 2270/S 1230) which will enhance state and federal regulators' ability to coordinate their oversight of bank technology partners, such as fintech lending platforms or cloud service providers.

The Networked Supervision priorities are important for building the system of the future, but the state system is already operating as a network in other areas. For example, states discovered a new way to use the CSBS data analytics platform. As part of the CSBS Data Working Group, the Iowa Division of Banking shared that it was developing a comprehensive risk analysis dashboard for banks using the self-service section of the CSBS Analytics platform. Recognizing this would be useful for other states, CSBS developed a launchpad capability to quickly deploy select state-created reports to all end-users in a special "Member Created Reports" section of our analytics platform. As a result, 15 states have already used Iowa's risk dashboard. We plan to expand this crowd-sourced capability so states can more seamlessly network with one another on analytics development activities and reduce duplicative efforts.

In another example, authentication technology employed by CSBS recently uncovered massive fraud that violated the Secure and Fair Enforcement for Mortgage Licensing Act (SAFE Act) education requirements for mortgage loan originators. CSBS notified the state agencies, who had the authority to enforce their SAFE Act requirements. As a result of this discovery, 45 state financial agencies reached settlements with more than 400 mortgage loan originators nationwide, resulting in \$1.2 million in penalties, and several state financial agencies reached a

settlement with the education administrator for both providing false certificates and taking courses on behalf of the mortgage loan originators through other education providers. Working with each other, state agencies ensured not just that the SAFE Act is followed but that consumers are protected nationwide.

CONCLUSION

As we have shown, technology plays an important role in today's state regulation of financial institutions. Whether using regtech or suptech, state regulators are finding ways to innovate and enhance information sharing, creating a network of supervision that benefits all stakeholders. We look forward to working with Congress and our federal partners on technology solutions that enhance regulation and supervision.

###

**Statement by Kelly J. Lay
Director of the Office of Examination and Insurance
National Credit Union Administration**

**before the Task Force on Artificial Intelligence
United States House of Representatives Committee on Financial Services**

“Keeping Up with the Codes – Using AI for Effective RegTech”

May 13, 2022

Chairman Foster, Ranking Member Gonzalez, and members of the Task Force on Artificial Intelligence, thank you for conducting this hearing on the effective use of Artificial Intelligence (AI) in RegTech and the opportunity to testify before you today.

My name is Kelly Lay, and I am the Director of the Office of Examination and Insurance at the National Credit Union Administration (NCUA). I started my career with the NCUA as an examiner in the field and have held various positions throughout the agency. Most recently, I was the NCUA’s Director of the Office of Business Innovation and led the development and implementation of the agency’s new examination platform, the Modern Examination and Risk Identification Tool, also known as MERIT.

The NCUA’s mission is to protect our nation’s system of cooperative credit and its member-owners through effective chartering, supervision, regulation, and insurance. To achieve this, the agency’s examination program focuses on risks to the broader system and the National Credit Union Share Insurance Fund. The credit union industry is comprised of 4,942 federally insured credit unions, with 129.6 million credit union member owners, with total assets of \$2.06 trillion dollars.¹

In my testimony today, I will first focus on the agency’s examination modernization efforts and highlight the research the NCUA has conducted in the realm of AI and RegTech. Second, I will discuss the NCUA’s challenges to incorporate AI and RegTech in the credit union industry. Third, I will relay the findings of last year’s request for information on the institutional use of AI. Finally, I will conclude with a legislative request for third-party vendor authority for the NCUA.

NCUA Examination Modernization and Research

In 2015, the NCUA formed the Enterprise Solution Modernization Program to help its staff regulate and supervise credit unions more efficiently. The program aims to modernize the NCUA’s technology solutions to create an integrated examination and data environment and facilitate a safe and sound credit union system. This modernization initiative aligns with the NCUA’s efforts to:

¹ See <https://www.ncua.gov/files/publications/analysis/industry-at-a-glance-december-2021.pdf>.

- Reduce the burden on credit unions and increase agency efficiency by reducing onsite examination time;
- Improve offsite supervision capabilities with better data and analytics;
- Provide consistency and standardization for the examination and supervision process with modernized, flexible applications; and
- Improve communications between the NCUA, credit unions, and State Supervisory Authorities.

As an initial step, the NCUA prioritized replacement of the legacy examination application, the Automated Integrated Regulatory Examination System (AIRES), which was over 25 years old. The agency also prioritized the need for a single point of entry for interacting with and accessing NCUA applications. The goal was to create a secure, modernized infrastructure to serve as a technical foundation for all future application modernizations.

After several pilot phases, the NCUA rolled out MERIT, including enhanced, integrated analytics utilizing a business intelligence tool, and our new, secure central user interface, called NCUA Connect, to the NCUA, State Supervisory Authorities, and credit unions in the second half of 2021. Currently, the NCUA has focused on helping users through this significant transition while deploying system enhancements.

As the NCUA's modernization efforts continue, the agency has an opportunity to integrate new technological solutions, applications, and data from other sources for centralized access and view into a credit union. There is a potential opportunity for offsite work and earlier risk identification with enhanced analytics capabilities. The NCUA has another ongoing modernization project that focuses on a virtual examination model to prepare for this potential development.

In 2017, the NCUA Board also approved virtual examination exploration and research funding. The agency's goal is to transition, within the next five to ten years, the examination and supervision function into a predominantly virtual one for credit unions that are compatible with this approach.

The virtual examination model should lead to greater use of standardized interaction protocols and enhance advanced analytical capabilities. For subject matter experts, the benefits are more consistent and accurate supervisory determinations and greater clarity about how the NCUA conducts supervisory oversight between the agency and credit union staff.

Currently, the virtual examination project is in the research and discovery phase. During this phase, staff identifies new and emerging data sources and methods to access the data, assesses advancements in analytical techniques, and considers harnessing other technologies to automate or streamline various aspects of the examination process. So far, research efforts have been focused on deploying AI solutions.

Specifically, the NCUA is in the testing phase of deploying a machine learning (ML) model to perform data validation more efficiently with quarterly Call Reports and profile submissions. This technique automatically clusters credit unions into various buckets and is more appropriate

for time-series data. Further, this technique employs forecasting models comparing actual and predicted values to identify outliers. Deployment of this new technique is expected to occur in the next four quarters and should result in more reliable and consistent Call Report filing across the credit union industry.

Another AI solution under investigation is Natural Language Processing (NLP). NLP transforms unstructured data into structured data; therefore, allowing end-users to leverage the data through analysis. Applied to the NCUA's examination process, NLP could take unstructured data, such as information found in board minutes, internal and external audit reports, and file maintenance reports, and turn it into structured data. Structured data is easier to consume, evaluate, and analyze.

The NCUA is also researching whether it could utilize Process Robotics Automation to perform an array of repetitive or routine tasks during examinations. Examiners could use this tool during examinations to perform various scope steps that currently require extensive manual review or input. The NCUA could save examination time, increase productivity, and reduce human errors by deploying this technology.

Finally, the NCUA has embraced a data-driven supervisory initiative with its largest credit unions. Through this approach, the agency improves its ability to assess risk to the Share Insurance Fund. For example, the NCUA conducts stress testing of its largest credit unions annually to evaluate their capital resiliency. Throughout the pandemic and the many months of social distancing, the agency estimated the potential impact on credit unions from adverse economic conditions and changes in borrower behaviors. Data-driven supervision offers opportunities to be more efficient by conducting supervisory activities offsite. The agency can integrate financial analysis and risk modeling results into our supervisory scope. These activities are conducted offsite, reducing time and costs traveling to credit unions.

Challenges for the NCUA and Credit Unions

The NCUA recognizes the importance and benefits of technological changes and has incorporated organizational change-management strategies into our initiatives. In addition to dedicated resources for development and testing, expanding the NCUA's use of RegTech and AI would require the agency to train examiners and credit unions, as applicable, and revise our examination policies and procedures. The NCUA supports and encourages innovation and the growth of the credit union industry, while at the same time ensuring it:

- Serves the needs of credit union members;
- Protects the interests of those members in terms of privacy and security; and
- Does not compromise the commitment to the industry's safety and soundness.

More importantly, AI and ML algorithms must be tested to prevent the intrusion of underlying historical bias that may result in discriminatory practices. We must be cautious when deploying AI tools, to avoid exacerbating systemic inequities.

As a mid-sized agency, the NCUA utilizes an Information Technology Oversight Council to review proposed IT investments to align with the NCUA's strategic plan. The NCUA considers our modernization roadmap, personnel resources, other information technology requests, implementation timelines, dependencies, and costs. Advanced technologies, such as those discussed today, may have data dependencies and integration with other initiatives that must be prioritized, organized, and systematically implemented to maximize the benefits to the NCUA and our stakeholders.

Such advanced technologies can be expensive to implement and maintain for both federal agencies and regulated entities. Currently, the NCUA does not have a budget dedicated strictly for AI, and the acquisition of such technology requires immense resources.

Likewise, small credit unions face similar, if not more significant, challenges. Of note, nearly two-thirds of credit unions (3,222) are smaller than \$100 million in assets, meaning they have very limited resources, and average only seven employees per institution. Of all credit unions, nearly 500 operate without a single full-time employee, and more than 100 of those credit unions operate exclusively with volunteers.

In general, credit unions are small, not-for-profit institutions and may not possess sufficient expertise to properly conduct due diligence on what is rapidly becoming a very complex ecosystem of third-party vendors. These smaller institutions may neither have the economies of scale nor the expertise necessary for sophisticated analytics. While small credit unions play a vital role in their communities, they are commonly short-staffed and may lack the resources required to keep abreast of evolving AI technologies.

Request for Information Results and Responses

Last year, the NCUA joined the Office of the Comptroller of the Currency, Federal Deposit Insurance Corporation, Federal Reserve Board, and Consumer Financial Protection Bureau in a request for information on the institutional use of AI and related challenges. We collectively received responses from financial institutions, vendors, industry trade groups, academic communities, and consumer advocacy organizations.

While the NCUA received 32 comments, only four were distinct from those received by the other regulators. Three of these letters were from credit union trade associations and one was from a natural person credit union. The broad range of respondents provided information on their uses of, and challenges with, AI and ML models in operations, including consumer-facing services and compliance management.

Respondents noted that AI/ML-based tools present many benefits, including the potential to expand access opportunities for consumers and create greater efficiencies for institutions. Likewise, many commenters noted that AI/ML-based tools and models contain new and expanded risks. For example, many community financial institutions, including credit unions, may not possess sufficient resources to internally develop and manage such tools, resulting in increased reliance on third-party vendors and outsourced solutions. In addition, there is concern

around the current regulatory framework, particularly regarding model bias, data quality, and fairness, as it relates to consumer facing products and services.

More work is needed to understand the impact of such innovation. Additionally, the development and deployment of innovative tools that affect consumers must be fair, protect consumers, and ensure compliance with fair lending laws.

In summary, the initial request for information provided preliminary insight into the industry and stakeholder views of how this technology should interact within the financial services sector; however, as we previously noted, the volume of responses was low. The Agencies are currently reviewing the responses they received in order to determine whether any clarification would be helpful for financial institutions' use of AI in a safe and sound manner and in compliance with applicable laws and regulations, including those related to consumer financial protection.

Third-Party Vendor Authority

Any examination of technology and the NCUA is incomplete without discussing the challenges the agency has confronted since the 2002 expiration of its third-party vendor authority. The NCUA asks Congress to enact legislation restoring the agency's examination and enforcement authority over third-party vendors, including credit union service organizations (CUSOs). The NCUA requires third-party vendor authority to safeguard not just the Share Insurance Fund, but the credit union system overall, which is a major pillar of our national economic system. The inability of the NCUA to supervise or examine third parties poses numerous systemic risks.

The NCUA, at present, has no authority to enforce third-party vendors' compliance with federal consumer financial protection regulations or prudential standards, like anti-discrimination laws, concentration limits, maximum loan-to-value ratios, minimum capital levels, and some cybersecurity and anti-money laundering efforts. While there are advantages for credit unions to use these service providers, the high concentration of credit union services within third-party vendors presents safety and soundness risks for the industry. The continued transfer of operations to CUSOs and other third parties hampers the ability of the NCUA to accurately assess the risks present in the credit union system and determine if current CUSO or third-party vendor risk-mitigation strategies are adequate.

With third-party vendor authority, the NCUA could better collaborate with vendors to streamline data collection for examination purposes and provide better analytics for identifying potential risks to credit unions and the Share Insurance Fund. We appreciate the continued support of this committee and thank Chairman Foster for introducing the [Strengthening Cybersecurity for the Financial Sector Act](#), which restores the NCUA's third-party vendor examination authority. The enactment of this legislation would close this growing regulatory blind spot the NCUA continues to confront.

In 1998, the NCUA was temporarily granted third-party vendor authority to address the Y2K changeover, but that authority expired in 2002. Since then, the NCUA's Inspector General, the Financial Stability Oversight Council, and the Government Accountability Office have all called

for the restoration of this authority.² Moreover, on March 4, 2022, in the NCUA Inspector General's [Top Management and Performance Challenges Facing the National Credit Union Administration for 2022](#) report, the NCUA Inspector General reiterated the need for third-party vendor authority as one of the agency's top seven challenges. Risks posed by third-party service providers ranked third of seven.

Because the NCUA lacks general statutory supervisory and enforcement authority, vendors and CUSOs can reject the agency's recommendations to implement appropriate corrective actions to mitigate identified risks. For example, several vendors have refused to implement the NCUA's recommendations to improve network security and safeguard sensitive member information due to cost concerns. By regulation, the NCUA requires federally insured credit unions to obtain a written agreement with a CUSO before investing or lending. These contractual provisions require the CUSO to provide the NCUA complete access to the CUSO's books and records, and the ability to review its internal controls, among other requirements. However, contractual enforcement alone is not optimal to perform the NCUA's oversight functions.

As such, the NCUA requests comparable authority as our counterparts on the Federal Financial Institutions Examination Council to examine third-party vendors. With such authority, the NCUA can better address the risks that arise from vendor relationships and safeguard the Share Insurance Fund.

Conclusion

In conclusion, AI and ML have the potential to improve the supervision of credit unions, but there are also challenges in implementing such technology. Thank you for the opportunity to provide input on this topic.

The NCUA also appreciates the support of the Task Force on Artificial Intelligence and the House Committee on Financial Services in promoting a robust credit union system and protecting its member owners. Again, I would like to thank the Chairman for introducing the bill to reestablish NCUA third-party vendor authority and request the members of this Task Force support the legislation when the committee considers it.

I look forward to your questions.

² See U.S. Government Accountability Office, GGD-99-91 "Enhancing Oversight of Internet Banking" (July 1999) <https://www.gao.gov/assets/ggd-99-91.pdf>. Office of Inspector General, OIG-20-07, "Audit of the NCUA's Examination and Oversight Authority over Credit Union Service Organizations and Vendors" www.ncua.gov/files/audit-reports/oig-audit-cusos-vendors-2020.pdf. Annual Reports of the Financial Stability Oversight Council 2015, 2016, 2017, 2018, available at <https://home.treasury.gov/policy-issues/financial-markets-financial-institutions-and-fiscal-service/financial-stability-oversight-council/studies-and-reports/annual-reports/fsoc-annual-reports-archive>. U.S. Government Accountability Office, GAO-04-91, "Financial Condition Has Improved, but Opportunities Exist to Enhance Oversight and Share Insurance Management" (October 2003) <https://www.gao.gov/products/gao-04-91>.

FCA Public



**Testimony of Jessica Rusu,
Chief Data, Information and Intelligence Officer,
UK Financial Conduct Authority**

To:

**The US House of Representatives; Committee on Financial Services; Task
Force on Artificial Intelligence**

"Keeping Up with the Codes – Using AI for Effective RegTech"

May 13, 2022

Washington, D.C. (Virtual Attendee)

FCA Public



Summary

This testimony will set out some of the key work and supporting documents that the Financial Conduct Authority (FCA) and other UK agencies have published in relation to our innovation work, particularly related to Artificial Intelligence (AI), Regulatory Technology (RegTech) and Supervisory Technology (SupTech).

For the purposes of this document, we define RegTech as industry solutions designed to make specific regulatory processes more effective and efficient, or larger, multi-faceted technology projects designed to modernise regulatory compliance functions across organisations. We define SupTech as the application of innovative technologies by regulators to enhance their supervisory and market oversight functions. This includes technologies that support both data analytics and data collection processes, across regulatory reporting, document analysis, prudential analytics, misconduct analytics and others.

Approach to Innovation

1. Building a digital regulator: how the FCA is riding the innovation wave (April 2022)

- This is a speech by Jessica Rusu at the Innovate Finance Global Summit 2022.
- The role of data and technology is increasingly pertinent as we are faced with volatile geopolitical conditions.
- There are three themes at the centre of our work to ensure we can effectively carry out our role as a regulator:
 - i. Fostering innovation – to create an enabling environment for innovation in financial services, while proactively scanning the horizon to get on the front foot of emerging issues.
 - ii. Being intelligence-led – to leverage data science and advanced analytics to inform our decision-making and help us regulate at scale.
 - iii. Preparing for the future – providing insights to ensure that our regulations are fit for purpose in a rapidly changing market.
- Synthetic data has been a key feature of our digital sandbox, and we have recently published [a Call for Input](#) to gather views and assess the potential of synthetic data to further spur innovation in the market.
- In May, we hosted our first ever CryptoSprint, engaging with the industry to seek their ideas to inform our regulatory policy thinking. This will be the start of a programme of industry engagement as we work together to develop a dynamic framework that supports innovation while protecting consumers.
- Later this year we will host a joint TechSprint with the Payment Systems Regulator (PSR) on Authorised Push Payment Fraud. It will explore solutions to identify and prevent APP Fraud, for example through identification of suspicious social media advertising and scam promotions.

2. Drivers of change in the financial services industry and how we are responding (November 2021)

- This is a speech by Jessica Rusu at the CDO Exchange for Financial Services in 2021.
- The threat landscape has shifted for consumers, with fraudsters and scammers benefitting from new technologies and new consumers being

drawn to high-risk markets and products, motivated by competition with friends, family, acquaintances and the influence of social media.

- To tackle the challenges faced by consumers and industry, we must make the best use of our own resources – connecting the dots in terms of intelligence across the organisation, drawing on strategies and approaches from data science, and leveraging data to create new tools and techniques which allow us to detect harm and intervene more quickly.
- We have a vital role in enabling the FinTech sector to achieve its potential and position the UK at the forefront of technological growth.

3. Levelling the playing field – innovation in the service of consumers and the market (April 2021)

- This is speech by FCA CEO Nikhil Rathi at UK FinTech Week in 2021.
- We believe that success in financial innovation has been enabled by regulatory open-mindedness.
- Support for innovation has been matched by action to protect consumers and markets.
- Tackling many of the issues we face – whether it is finding market abuse in transaction data, mapping access to cash against consumer vulnerability, scraping the web for poor advertising or the quicker raising of red flags about those we regulate – relies on the innovative use of technology and data, in many cases drawing on lessons from others in industry.
- In order to meet our objectives over the long term we need not only to improve our structure, technology and data analytics capabilities, but also to adapt our culture and our risk appetite, and for us to take innovative advantage of our powers. This is the motivation for bringing this together under a new DTI division led by Jessica Rusu.

4. FCA Data Strategy (January 2020, updated August 2020)

- The FCA's refreshed data strategy builds on learnings from our [first data strategy](#), published in 2013. It sets out how the organisation will focus on new technology and advance analytics techniques that are available and look at how we can apply these to improve our own efficiency and effectiveness.
- The data strategy has three strategic objectives:
 - i. Gain deeper understanding of how markets work and consumer behave to shape our interventions.
 - ii. Identify and respond to problems in firms and markets to intervene and minimise harm earlier.
 - iii. Build a flexible, future-ready organisation to react faster to market changes and shifting priorities.
- The portfolio of work to deliver on these priorities is structured around the following areas:
 - i. Projects and Key Business Initiatives in every single division that exploit our capabilities and ways of working to bring benefit to us all.
 - ii. Changes to our culture and skills to provide the environment, training and recognition for all employees to identify and champion opportunities to fully exploit data and deliver improvements in their areas.
 - iii. A new FCA data operating model that provides central capabilities and services to push forward our transformation, while growing data and analytical capability throughout the organisation.

FCA Public



- iv. Improved Central Data Services that support our governance, control and supply of data and enable us to make better use of it by providing support across a range of disciplines.
- v. New Data Management and Analytical tools to give us control, flexibility and power in the way that we use our data.
- vi. Core enabling technology platforms giving us the foundational technologies and infrastructure to fully harness the power of data.
- We have committed to publishing updates on our progress and how we deliver continue to support our mission.

Artificial Intelligence / Machine Learning

1. FCA and Bank of England (BoE) - AI Public-Private Forum: [Final report](#) (February 2022)

- The FCA and BoE launched the Artificial Intelligence Public-Private Forum (AIPPF) on 12 October 2020 ([Terms of Reference](#)). The purpose was to further dialogue on AI innovation between the public and private sectors. More specifically, the AIPPF sought to:
 - i. Share information and understand the practical challenges of using AI within financial services, as well as the barriers to deployment and potential risks.
 - ii. Gather views on potential areas where principles, guidance or good practice examples could be useful in supporting safe adoption of these technologies.
 - iii. Consider whether ongoing industry input could be useful and what form this could take.
- The AIPPF ran for one year, with four quarterly meetings and a number of workshops. It brought together a diverse group of experts from across financial services, the tech sector and academia, along with public sector observers from other UK regulators and government.
- The Final Report explores the various barriers to adoption, challenges and risks at three levels within AI systems: Data, Model Risk and Governance. It also explores potential ways to address such barriers and challenges, as well as mitigate potential risks. These are highlighted in the key findings and examples of best practice.
- Ultimately, this report aims to advance the collective understanding and promote further discussions amongst academics, practitioners and regulators to support the safe adoption of AI in financial services.

2. 'AI Transparency in financial services – why, what, who and when?' – [FCA Insight Article](#) (February 2020)

- The FCA and The Alan Turing Institute worked together on a year-long collaboration on AI transparency.
- This is a blog post where we explain the motivation for pursuing such a project and present an initial framework for thinking about transparency needs in relation to machine learning in financial markets.

3. [AI in Financial Services](#) – The Alan Turing Institute report commissioned by the FCA (June 2021)

- The FCA and The Alan Turing Institute worked together on a year-long collaboration on AI transparency. As part of this work, the FCA

commissioned a report to inform and advance the debate about responsible AI in the context of financial services.

- The report provides an introduction to relevant technological concepts, discusses general challenges and guiding principles for the adoption of AI, maps out potential benefits and harms associated with the use of AI in financial services, and examines the fundamental role of AI transparency in pursuing responsible innovation.

4. **FCA and BoE Joint - Machine Learning in UK Financial Services [Research Note](#) (October 2019)**

- The FCA and BoE conducted a joint survey in 2019 to better understand the current use of machine learning (ML) in UK financial services, including the current state of deployment, maturity of applications, use cases, benefits and risks.
- This is a research note presenting our analysis of the responses to the survey. It includes a qualitative overview of the use of ML across the respondent firms
 - i. A quantitative overview of the use of ML across the respondent firms
 - ii. The ML implementation strategies of firms that responded to the survey
 - iii. Approaches to the governance of ML
 - iv. The share of applications developed by third-party providers
 - v. Respondents' views on the benefits of ML
 - vi. Perceptions of risks and ethical considerations
 - vii. Perspectives on constraints to development and deployment of ML
 - viii. A snapshot of the use of different methods, data, safeguards performance metrics, validation techniques and perceived levels of complexity
- We will be re-running the joint survey with the BoE later this year.

5. **[The future of regulation: AI for consumer good](#) – Speech (July 2019)**

- This is speech delivered by former FCA Executive Director of Strategy and Competition Christopher Woolard at the Alan Turing Institute's AI ethics in the financial sector conference.
- As the regulator, we consider the use of AI in financial services from three main perspectives:
 - i. Firstly, which parts of the debate are novel and where is there continuity.
 - ii. Secondly, how can we ensure AI is creating value for citizens.
 - iii. And lastly, how can we work with others to develop a shared understanding that will determine our approach over the years ahead.

Supervisory Technology (SupTech), including Advanced Analytics and Data Science

1. **FCA RegTech Forum – FCA Approach to Innovation & SupTech [Event Recording](#) (February 2022)**

- This is an on-demand recording of the FCA's RegTech Forum on our approach to regulatory innovation and SupTech.
- The event showcased how our end-to-end approach to innovation has evolved over the years and lessons learned, followed by a more detailed overview of our SupTech projects and how we are utilising new technologies to meet our objectives.

- The event featured a panel with domestic and international regulators and academia to discuss the importance of innovation for regulators, how the field of SupTech is evolving, and what the future might hold for it.

2. **FCA Approach to Innovation & SupTech** - [Transcript](#) and [Video Recording](#) (February 2022)

- This is a transcript and video recording of the speech given by Jessica Rusu at the aforementioned RegTech Forum 'Approach to Innovation & SupTech' event in February 2022.
- Across the Data, Technology, and Innovation division, we are investing heavily in the skills and technology foundations that will make the FCA a regulator fit for a digital future.
- We are developing a new Digital Unified Intelligence Environment (DUIE) that fundamentally re-imagines how we collect intelligence, analyse data, and will help the FCA take action in a more joined up way.
- Underpinning the Data Strategy is the cultivation of a spirit of innovation, of which we believe there are three key ingredients: culture, experimentation, and collaboration.
- Building a diverse and inclusive culture is an imperative to create the environment that is receptive to new ideas – an environment that has not only the necessary skills, but also the agile mindset to adapt to change.
- Supporting early innovation is about supporting experimentation, and two ways the FCA does this through our digital sandbox initiative and our TechSprint model.
- We believe it is important to work collaboratively with industry on shared challenges. Our innovation culture will be influenced by our external network, and collaboration with innovative firms, academics, and industry forums in building the future of regulation. This is what we refer to as the 'RegTech-SupTech bridge'.

3. [Meeting the pace of technological change](#) – Speech (November 2019)

- This is a speech delivered by former FCA Director of Innovation Nick Cook at the Chief Data Officer Exchange Financial Services conference in London.
- The changing technology landscape means that regulators, such as the FCA, need to continually adapt to remain fit for purpose for both the markets and firms we oversee and the consumers we serve to protect.
- Our innovation services started as a 2-person initiative called 'Project Innovate' offering regulatory feedback to innovative business models, and it has grown into an entire division dedicated to bringing disruption to financial markets in the interest of consumers.
- We are continually asking ourselves whether FCA innovation services could be enhanced to broaden relevance and appeal, including by providing greater support to the RegTech market. We are looking at ways to do this through the creation of a 'digital sandbox'.
- Our TechSprints have provided opportunities to collaborate with RegTechs on focussed proofs-of-concept, including in the area of AML and financial crime. Our [2019 Global AML and Financial Crime TechSprint](#) brought tech firms and the industry together to explore the potential of Privacy Enhancing Technologies, or PETS (e.g. federated learning; complex scenario modelling and scenario simulation), to make improvements in the area of financial crime detection.

4. [From Innovation Hub to Innovation Culture](#) – Speech (June 2019)

- This is a speech delivered by former FCA Director of Innovation Nick Cook at the 6th Central Bank Executive Summit.
- We believe that regulators should actively stimulate certain innovation within the market that we believe will deliver public value. We have been vocal in our desire to see further innovation and progress in anti-money laundering and financial crime, and we have previously seen the positive impact that publicly identifying a specific area of interest can have on technological development.
- The FCA has an increasing focus on international engagement in innovation in financial services and being a global advocate for innovation. This is why we proposed and launched the Global Financial Innovation Network (GFIN).

Regulatory Technology (RegTech)

1. 'RegTech – a Watershed Moment?' – [FCA Insight Article \(June 2020\)](#)

- Each year the FCA carries out a significant amount of research and analysis to help financial markets work effectively. The FCA Insight section is designed to promote access to some of the most interest and thought-provoking highlights of our work.
- The Covid-19 crisis has brought to the fore the potential for technology to streamline regulatory activity and bring benefits to organisations. The Covid-19 crisis may prove a pivotal moment for the RegTech industry's growth and transformation.
- Regulators have a role in supporting the adoption of technology to solve regulatory challenges while maintain careful oversight of the sector.

2. 'The Future of RegTech – what do firms really want?' – [FCA Insight Article \(June 2021\)](#)

- The use of technology to streamline the regulatory process in financial services continues to be a fast moving and fast-growing sector.
- Data since the start of the pandemic indicates that RegTechs are realising potential found during the Covid-19 crisis to help industry adopt technology solutions for tackling new types of financial crime and respond to changing regulatory rules.
- The FCA commissioned an external agency to conduct quantitative research among large UK financial services firms about the current state of play in the RegTech industry, their perceptions of RegTech and the FCA's role, as well as the impact of Covid-19.
- Through these interviews, we received a good understanding of how firms engage with RegTechs and what the onboarding journey looks like - a crucial part of the adoption process.
- It was observed within the RegTech ecosystem that RegTech firms have the challenge of making the leap from 'proof of concept' to demonstrating clear 'proof of value'. The FCA's digital sandbox has been designed to assist innovators to develop new solutions and products through access to synthetic data assets, and demonstrate the potential value of these products to the market by way of an observation platform open to market participants.

3. [Fostering innovation through collaboration: The evolution of the FCA TechSprint Approach Report \(March 2020\)](#)

- This is a report we published in 2020 exploring the lessons we have learned from hosting seven TechSprints across a range of subjects. As part of our commitment to foster innovative and collaborative global financial markets,

FCA Public



we want to share our experience and support other organisations in their efforts to address key industry problems.

- The report includes information on how the FCA goes about planning TechSprints that may be of interest to regulators considering planning their own events.

4. **FCA Insight Podcast - Using technology to fight financial crime (August 2019)**

- This is an audio recording and [transcript](#) of an FCA Insight Podcast hosted by former FCA Director of Innovation Nick Cook during the FCA's 2019 Global AML and Financial Crime TechSprint.
- In the podcast, Nick Cook is joined by Global Head of Financial Crime Threat Mitigation at HSBC, Jennifer Calvery, and Co-Founder and CEO at 11:FS, David Brear, to discuss how new technologies can be used to prevent and detect financial crime.

5. **Technology and global ties: turning the tide on financial crime – Speech (May 2018)**

- This is a speech delivered by former FCA Executive Director of Strategy and Competition Christopher Woolard at the Anti-Money Laundering TechSprint in 2018.
- Key points from this speech include:
 - i. The financial system is the gateway through which criminals must pass to make crime pay.
 - ii. New technologies have huge potential to monitor, analyse and prevent financial crime. Machine learning to improve the detection of suspicious activity, artificial intelligence driven anti-impersonation checks and distributed ledger technology to improve the traceability of transactions are just some of the real, practicable solutions technology has gifted us.
 - iii. A clean, competitive financial system cannot be achieved in isolation – only by working together can we make a real impact. It is only by working together, pooling our resources and sharing our expertise, that we will achieve the real and tangible change we are all seeking.



May 17, 2022

The Honorable Maxine Waters
Chair
U.S. House Committee on Financial Services
2129 Rayburn House Office Building
Washington, DC 20515

The Honorable Patrick McHenry
Ranking Member
U.S. House Committee on Financial Services
2129 Rayburn House Office Building
Washington, DC 20515

The Honorable Bill Foster
Chair
Task Force on Artificial Intelligence
U.S. House Committee on Financial Services
2129 Rayburn House Office Building
Washington, DC 20515

The Honorable Anthony Gonzalez
Ranking Member
Task Force on Artificial Intelligence
U.S. Committee on Financial Services
2129 Rayburn House Office Building
Washington, DC 20515

Dear Chairs Waters and Foster, Ranking Members McHenry and Gonzalez, and members of the Task Force on Artificial Intelligence:

We appreciate the opportunity to submit this letter in relation to the hearing of the House Financial Services Committee Task Force on Artificial Intelligence, held on Friday, May 15, 2022, to explore the use of AI in regtech. We commend the Task Force for holding this important discussion and respectfully request that our letter be included in the hearing record.

The Alliance for Innovative Regulation (AIR) is a nonpartisan, non-membership 501c3 organization founded in 2019 to spread awareness about the need for regulators and compliance professionals to adopt cutting-edge digital technology to keep pace with changes taking shape in the financial industry.¹ Through AIR's thought leadership, sponsorship of problem-solving competitions known as "TechSprints," and other activities, we encourage the regulatory community to make changes that we believe will result in a fairer and more inclusive financial system.

Nearly 90 years after the British mathematician Alan Turing introduced the concept of stored-memory machines, numerous industry and government sectors continue to find new practical uses for artificial intelligence.² The banking sector's compliance with laws and regulations, and regulators' oversight of the financial system, together are Exhibit A in unlocking the real-world applicability of AI. Policymakers should consistently encourage the federal financial regulatory agencies to embrace the power of machine learning, natural language processing and other AI tools to monitor a financial sector that is on the path to being fully digitized.

¹ <https://regulationinnovation.org/>

² How Alan Turing Invented the Computer Age, by Ian Watson, Scientific American, <https://blogs.scientificamerican.com/guest-blog/how-alan-turing-invented-the-computer-age/>

Until recently, regulatory oversight of the banking industry was fundamentally paper-based. This has led to gaps in the system that still pose obstacles to achieving efficiency in financial supervision and compliance. The industry is light years ahead of the government in upgrading technology infrastructure to offer customers a wider array of digital products. Being late to the party has cost regulators in multiple ways. Agencies find themselves needing to close a knowledge deficit compared to the more advanced industry they oversee, and they are trying to supervise banks' digital makeover with a government infrastructure that is still analog.

If we widen the lens from oversight of traditional banks to emerging financial products like digital assets, the challenge is even more acute. The expectations will likely rise over the next few years for regulators to address the risks and opportunities presented by cryptocurrencies and other new activities. They will likely need to determine how to oversee crypto functions tied to payments and currency, crypto investments, non-fungible tokens (NFTs), decentralized autonomous organizations (DAOs), financial activities in the "metaverse," and even newer iterations of financial instruments that leverage AI and other technologies. These innovations already are breaking the mold of the traditional financial system, while transcending legal frameworks, regulatory agency structures and national boundaries. Regulators will need AI-based tools to monitor these complex, fast-changing markets.

As noted in the Administration's recent Executive Order on digital assets, these innovations have significant upside potential to improve financial markets, but they also carry risks for consumer and investor protection, financial inclusion efforts, mitigating the impact of climate change, fighting financial crime, and maintaining systemic stability.³ Regulators will need regtech tools to monitor and curb these risks.

There is reason for hope that just as financial firms increasingly utilize "Regulatory Technology" (regtech) powered by AI to improve their compliance systems, government agencies are beginning to address their own digital knowledge deficit and access "Supervisory Technology" (suptech) tools that will put regulators on more equal footing with the financial sector.

Nomenclature in the regtech space is still young and evolving. The term usually refers to the industry's efforts to use high-tech compliance methods. Suptech is the term commonly used to describe regulators' adoption of the same kinds of tools. A primary methodology in both categories is AI — and its sub-branches of Machine Learning (ML), Natural Language Processing (NLP), and neural networks — that are deployed to improve regulatory processes. Financial regulatory agencies around the world have made significant strides embedding an AI focus within their traditional supervisory framework. In the United Kingdom, the Financial Conduct Authority (FCA) has honed the use of machine learning and natural language processing to monitor industry activities.⁴

Here in the U.S., the Securities and Exchange Commission (SEC) uses machine-learning algorithms and other AI methods to detect risks in certain investment instruments and to spot anomalies in regulatory

³ Executive Order on Ensuring Responsible Development of Digital Assets, <https://www.whitehouse.gov/briefing-room/presidential-actions/2022/03/09/executive-order-on-ensuring-responsible-development-of-digital-assets/>

⁴ The Financial Conduct Authority's Innovation Journey, by Amy Friend, senior advisor to Alliance for Innovative Regulation, <https://regulationinnovation.org/wp-content/uploads/2021/03/The-FCA's-Innovation-Journey-Moving-Forward-in-the-Face-of-Uncertainty-AIR-Final-2.pdf>

filings.⁵ Other entities such as the Commodity Futures Trading Commission (CFTC), Financial Industry Regulatory Authority (FINRA) and Financial Crimes Enforcement Network (FinCEN) have similarly made progress integrating AI into oversight and compliance processes.

While the U.S. banking agencies have been slower in adopting AI tools, they proved adept at developing digital reporting tools and offsite examination procedures in the face of the COVID-19 pandemic, which signals their ability to make use of digital data in the years to come.

However, the regulatory community has only begun to scratch the surface. Financial innovation is moving forward at such accelerating speeds that regulators and law enforcement authorities will be disadvantaged in detecting risks and catching bad actors if they try to layer digital improvements on top of a legacy analog design. Agency leaders need to think more holistically about establishing a digitally native design.

A prime example of this gap is the monitoring of global transactions to identify financial crimes, where more robust use of regtech and supotech powered by AI would have enormous impact. Despite the billions of dollars that banks spend on suspicious activity reports and other anti-money laundering (AML) requirements, the criminals are still going uncaught. The United Nations has estimated that less than 1% of financial crimes are ever detected and stopped.⁶ Traditional transaction monitoring favors casting as wide a net as possible, but suspicious transaction alerts produce an alarmingly high 90% false-positive rate.⁷

It's not that the regulators lack data about suspicious transactions and other financial risks. On the contrary, digital technology has given them a limitless supply of data. Rather, they do not yet deploy AI-powered tools to analyze and synthesize the data, which could help them focus their efforts on the highest value activity.

Mark Carney, a former governor at the Bank of England, famously said in 2019 that the bank received 65 billion pieces of industry data each year and that getting a handle on what the data means would require “each supervisor reading the complete works of Shakespeare twice a week, every week of the year.”⁸

AI can enable financial institutions and government authorities to synthesize transaction data with more precision to find the needles in the haystack.⁹ This would lead to beneficial outcomes on a global scale. Better efforts aimed at stopping money laundering could shed light on Russian oligarchs — subject to sanctions triggered by the war in Ukraine — hiding their assets in shadowy shell companies. It could flag

⁵ FCA Using NLP, Machine Learning to Regulate Businesses, by Joanna Wright, Waters Technology, <https://www.watertechnology.com/regulation/7551751/fca-using-nlp-machine-learning-to-regulate-businesses>

⁶ Fact Sheet, Money Laundering, United Nations' Office on Drugs and Crime, https://www.unodc.org/documents/hlr/FactSheets/Money_Laundering.pdf

⁷ Banks feel more pressure to upgrade AML tech after 'Fincen Files', by Penny Crosman, American Banker, <https://www.americanbanker.com/news/banks-feel-more-pressure-to-upgrade-aml-tech-after-fincen-files>

⁸ Enable, Empower, Ensure: A New Finance for the New Economy, Speech by Mark Carney, Governor of the Bank of England, <https://www.bankofengland.co.uk/-/media/boe/files/speech/2019/enable-empower-ensure-a-new-finance-for-the-new-economy-speech-by-mark-carney>

⁹ How AI improves AML efforts across the financial services industry, by Clark Frogley, Banking Exchange, <https://m.bankingexchange.com/news-feed/item/9024-how-ai-improves-aml-efforts-across-the-financial-services-industry?Itemid=577>

global transactions associated with human trafficking. AI-powered methods can also help improve the highly regulated process of banks performing due diligence on new customers and identifying their beneficial owners.

AML is not the only use case for bringing a suptech focus to financial regulatory authorities. If regulators incorporated AI into their analysis of lending patterns, it would better reveal incidents of fraudulent loan applications, discriminatory lending patterns or violations of the federal prohibition on unfair, deceptive, or abusive acts or practices (UDAAP). UDAAP enforcement often involves subjective determination about what is a true violation versus something milder. AI-based tools can make detection of violations clearer to all parties.

The mainstream growth of AI in the financial sector, while promising efficiency for both consumers and financial institutions, undoubtedly brings new risks. There are looming unanswered questions about the security of consumers' financial data, who owns that data, whether AI algorithms are prone to bias and discrimination, and whether companies' reliance on machine learning and other tools can pose a threat to safety and soundness. But as the industry and public sector navigate this new frontier, AI will also bring a whole host of benefits to regulators' efforts to oversee the financial sector with more efficiency, precision and fairness.

We recommend that financial regulators accelerate and expand their regtech activities by taking the following steps:

- Adopt suptech strategies at each agency. Some agencies already have this work underway. Each agency should explore how technology, including AI, could close information gaps and equip supervisors and field examiners with better and more timely information in more usable form. We also recommend expansion of the suptech work underway at the Federal Financial Institutions Examinations Council (FFIEC) to coordinate activities among agencies.
- Shift financial regulatory reporting to a digital format. This would involve converting reported information so that it could be submitted and analyzed electronically using AI-based tools. Using a digital reporting format could be phased in as a voluntary option, starting with regulated entities that are equipped to comply easily. This process would enable supervisors to learn to use more AI tools and equip agencies to develop strong reporting systems to be used more widely over time. One model project is underway at FINRA.
- Move toward digital regulatory reporting (DRR) that accesses full data sets and real-time information. Regulators over time should shift away from reliance on sampling of data and toward gaining access to full data in real time. This should be done with appropriate controls to ensure that access is limited to information that is legally available to agencies and to designated, authorized individuals. Some reporting activities could be shifted from a "push" basis — reported by the firm — to a "pull" basis in which agencies can access data as needed with appropriate limits and safeguards. This information would be obtained and analyzed using AI. Such a step could greatly enhance regulatory effectiveness while also reducing industry reporting costs, which are very substantial. The FCA has estimated¹⁰ that shifting to digital

¹⁰ After the Pandemic: The Future of Financial Regulation with UK Financial Conduct Authority Interim CEO Christopher Woolard, Barefoot Innovation Podcast,

regulatory reporting could save U.K. financial companies over one billion pounds annually. For models on digital regulatory reporting, see activities of the Bank for International Settlements Innovation Hub¹¹ and the G-20 2020 Global Regulatory Techsprint.¹²

- Focus attention on the benefits of AI in AML and fair lending. Efforts to fight money laundering and comply with fair-lending rules are perhaps the most advanced use cases for adopting regtech, with the market offering an array of AI-powered tools. These two areas could help regulators learn how to use these new technologies to advance their supervision objectives. New methodologies can be incorporated in both AML and fair-lending reporting requirements and examination procedures.
- Adopt use of NLP. Regulators have found that use of Natural Language Processing offers numerous use cases. In those cases, NLP is relatively easy to adopt and can generate substantial benefits in regulatory outcomes, including better efficiency and awareness of industry risks. For example, regulators are using NLP to review and analyze public comment letters, consumer and investor complaints, and disclosure documents such as SEC filings.
- Encourage agency innovation offices to address regtech and suptech. All of the U.S. federal financial regulatory agencies have undertaken innovation initiatives in some form, as have many states. However, most focus mainly on the industry's innovation activities, conducting functions such as office hours and sometimes experimentation environments. We recommend that these offices, or equivalent ones, also focus on how new technology, including AI, can be used by the agencies themselves for suptech purposes, and to assess and encourage adoption of effective regtech for industry compliance activities. These efforts should be accompanied by updated training and procedures for field examiners.
- Experiment with suptech oversight tools for digital assets, cryptocurrency activities and related industry innovation. Digital assets, crypto, and "decentralized finance" (DeFi) are generating new financial activity at a pace not seen before in finance. As noted in President Biden's Executive Order, these developments offer substantial upside opportunity but are also creating significant risks in areas like consumer and investor protection, financial crime, climate impacts and potential systemic instability. As regulatory policy is clarified, agencies will need to be equipped with suptech tools for overseeing these novel and fast-changing markets. AI will be essential to enable regulators to understand and respond to these developments.
- Develop workforce talent specializing in digital technology, including AI. Federal and state financial regulators will need more technology and AI expertise. We encourage them to assess talent gaps in fields like data science, software engineering and human-centered design, and develop plans to close those gaps. This includes assessing hiring barriers in recruiting procedures and compensation policies. We also suggest that agencies undertake comprehensive training of their personnel in digital technology, including AI.

<https://regulationinnovation.org/podcast/the-future-of-regulation-series-episode-4-with-interim-ceo-of-the-uk-fca-christopher-woolard/>

¹¹ <https://www.bis.org/about/bisih/about.htm>

¹² https://www.bis.org/hub/g20_techsprint.htm

For more details on AIR's work regarding AI, regtech and supotech, we encourage the Task Force to review the following: 1) A series of papers on the financial inclusion benefits of technology-based innovation that I published through a senior fellowship at the Harvard University Kennedy School of Government¹³; and 2) AIR's Regtech Manifesto, a paper calling for the conversion of the financial regulatory system to a "digitally-native" framework¹⁴

If we can be helpful to the Task Force or to the Committee in any way, please let us know.

Sincerely,
Jo Ann Barefoot
Founder and CEO
Alliance for Innovative Regulation

¹³ <https://regulationinnovation.org/harvard-kennedy-school-papers/>

¹⁴ <https://regulationinnovation.org/regtech-manifesto/>



Data Coalition
1100 13th Street NW
Suite 800
Washington, DC 20005

202.964.1120
@DataCoalition
info@datacoalition.org
datacoalition.org

**Task Force on Artificial Intelligence
United States House of Representatives Committee on Financial Services**

May 20, 2022

**Written Testimony of Nick Hart, President, Data Foundation
Subject: Building a Data System for Responsible Implementation of Effective Regulatory
Technology**

Chair Foster, Ranking Member Gonzalez and Members of the Task Force on Artificial Intelligence of the House Committee of Financial Services, thank you for this opportunity to provide outside written testimony on behalf of the Data Foundation.

The Data Foundation champions the use of open data and evidence-informed public policy to make society better for everyone. As a nonpartisan think tank, we conduct research, collaborative thought leadership, and advocacy programs that advance practical policies for the creation and use of accessible, trustworthy data. The Data Coalition Initiative within the Data Foundation includes a broad group of data analysis and technology companies, public sector consulting and accounting firms, and nonprofits. Regulatory Technology (RegTech) and Artificial Intelligence (AI) are two issue areas where the Data Coalition Initiative promotes policies that enable government data to be high quality, accessible, and usable. In particular, our advocacy focuses on the adoption of data standards that improve the quality of information that strengthens the ability to compare, link, share, and analyze data. Earlier in the 117th Congress, the House of Representatives passed the Data Coalition-endorsed Financial Transparency Act (H.R. 2989), which promotes long overdue data standards for financial data. As the Committee considers the use of AI for effective application of RegTech, the Data Foundation and members of the Data Coalition Initiative emphasize the importance of strong data governance frameworks, adequate resourcing and staff capacity, and clear legal authorities that enable the responsible adoption of RegTech across federal agencies. AI technologies that support regulatory functions provide an opportunity for government to ensure it is fulfilling its oversight functions, while reducing the burden on regulated entities.

The Data Coalition Initiative recently hosted a convening where experts across government, nonprofit, academia, and private sector organizations discussed innovative solutions to facilitate an environment to support the use of RegTech in the federal government (*see Appendix A*). The need for both data standards and data governance were recurring themes throughout the day's event. Data standards, as one speaker put it, are "a common understanding of data elements and vocabulary ... [and] can help drive innovations and will support data-informed decision making. Standards facilitate automation and streamlining of processes, which results in lower costs and more effective analysis."¹

¹ Transcript of RegTech 2022 Summit,
<http://www.datacoalition.org/wp-content/uploads/2022/05/RegTech22-Data-Summit-1.docx>



Data Coalition
1100 13th Street NW
Suite 800
Washington, DC 20005

202.964.1120
@DataCoalition
info@datacoalition.org
datacoalition.org

The ability to access and use large amounts of high-quality, standardized data would allow application of innovative technologies to glean new and timely insights from which oversight bodies, agencies, and regulators can use to improve policies. For example, to better analyze important issues facing the country using data such as improving environmental quality, addressing the growing need for more information on climate impacts, and assessing the intersections between economic and environmental impacts, data needs to be available and useful. Technology can help improve analytical capabilities and reduce burdens – and costs – for conducting analysis if the required data is standardized and machine-readable.

Effective data governance practices and clear leadership in agencies can enable federal agencies to implement such standards and develop processes that support the generation and use of high-quality data. Many of these required practices are now required by the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act), which includes the OPEN Government Data Act. The Evidence Act establishes data officials at federal agencies, provides statutory support for improving data management across government, directs agencies to establish comprehensive data inventories and governance processes, requires the publication of open data plans, and establishes practices for standardized, machine-readable data. The Evidence Act's Chief Data Officers, Evaluation Officers, and Statistical Officers are all data leaders who lend expertise to ensure agency data collection, analysis, and use capabilities are applied across program and agency functions. The continuing focus on implementing and enforcing various provisions from the Evidence Act is one way to encourage strong data governance frameworks.

As Congress and Executive Branch agencies move toward a more holistic data management structure, they must also consider potential bias and ethical concerns. During the Data Coalition's "Accelerating AI in the Public Sector" forum (*see Appendix B*), experts shared perspectives and recommendations on both the potential promises and issues surrounding application of AI in government.² Strong data governance and standardized, structured data can mitigate such problems by providing a way for the government to detect the existence of bias in AI algorithms or discriminatory practices within agency programs. For example, if housing data is structured and standardized, data mining would provide a technique for analyzing mortgage, leasing, and eviction data sets to determine if there is a systemic pattern of discrimination within the vast amount of housing data available.

Clearly your committee recognizes the value of AI in RegTech to reduce cost, ease compliance burdens, and improve effectiveness of federal regulations across policy areas. The Financial Transparency Act is a productive first step. To continue that work **your committee should consider how to adequately resource efforts to adopt artificial intelligence for RegTech and provide legal authorities where appropriate.**

² Transcript of "Accelerating AI in Public Sector" Event,
<http://www.datacoalition.org/wp-content/uploads/2022/05/Public-Forum-Accelerating-AI-in-the-Public-Sector-2.doc>



Data Coalition
1100 13th Street NW
Suite 800
Washington, DC 20005

202.964.1120
@DataCoalition
info@datacoalition.org
datacoalition.org

Implementing new technologies, such as RegTech, is a significant undertaking for both regulatory agencies as well as their customers, and the associated costs for using AI to improve regulatory compliance can be resource intensive. Reliable, predictable, and adequate funding is crucial for enhancing agencies and financial institutions' capacity for sophisticated use of technology for regulatory purposes. While some larger organizations are able to adopt and keep up with rapidly evolving technological advances, smaller organizations with less capital may fall behind in serving their customers or complying with regulation.

Staff expertise may also be a key indicator for an agency's or institution's capacity for the level of sophistication in analysis and skills needed to implement and effectively manage RegTech in practice. Smaller institutions with few staff members, who may benefit greatly from incorporating AI into their oversight or compliance processes, may be restricted from adopting new technologies due to their size and resources. Funding to bolster capacity and hire additional staff as well as providing training to ensure the workforce has the necessary skills to adequately use AI for RegTech should be prioritized when considering the effectiveness of introducing these new technologies.

Congress will certainly have a role to play in facilitating a more robust RegTech system by examining new or revised legal authorities. For example, regulators may have the opportunity to ease the reporting burden on their constituents by allowing increased data sharing, particularly between state and federal regulators. Congress can continue to work with state regulators and the Financial Stability Oversight Council to seek strategies in which data silos can be reduced, streamlined, or eliminated.

Further, as off-the-shelf AI solutions become more accessible to smaller entities, they may present new regulatory challenges. These types of third-party solutions offer smaller banks and credit unions the chance to take advantage of the benefits of RegTech solutions, without the sophistication and funding of larger entities. At the same time, while these solutions offer the benefits of RegTech, they also offer new risks that small entities may not be able to evaluate sufficiently on their own. Congress should work with regulatory agencies and small entities that are likely to adopt third-party solutions to see what types of reasonable safeguards are necessary to ensure safe, responsible, and cost-effective RegTech adoption for these entities.

Reliable, quality, and accessible data feed the technologies that can help the government carry out its responsibilities efficiently and effectively. As the Task Force informs the full Committee's consideration of the use of RegTech and AI to improve federal government regulatory processes, hopefully it is clear that data governance, adequate resources, and necessary authority are crucial to the foundation that enables these technologies.

Thank you for the Task Force and Committee's attention to highlight the important issues related to the application of RegTech and AI in our society. Responsible application of these approaches holds substantial promise to benefit the American people, businesses, and our government with improved efficiency, reduced burden, and ultimately better information. We hope Congress will consider and take



Data Coalition
1100 13th Street NW
Suite 800
Washington, DC 20005

202.964.1120
@DataCoalition
info@datacoalition.org
datacoalition.org

actions in the near future to ensure that these technologies are responsibly adopted, regulation is increasingly more efficient and effective, and regulatory burden is decreased.

Thank you for the opportunity to provide written testimony on behalf of the Data Foundation and the members of the Data Coalition Initiative. We look forward to working with you, the Task Force, and the Committee on these important issues in the months and years ahead.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Nick Hart: Nick Hart, president of the Data Foundation. And this event is hosted by the Data Coalition, presented by Donnelley Financial Solutions. The Data Coalition is America's premier voice on data policy, and we're really excited to be able to host today's event. Now, this morning, we had a whole discussion, a public forum on artificial intelligence. And this afternoon, we're going to delve deeply into the topic of regulatory technologies. Before we get started, I want to offer just a few logistics notes. So, this is the Data Coalition's first day of in-person events in over two years, and we're very excited that you all are joining us for those that are in-person. And for those that are joining us virtually, thank you also for joining us in the hybrid event.

For those that are in-person, you showed your COVID vaccination cards on the way in, and thank you for doing that. We'll ask those that would like to, to continue masking out of respect for those who have different levels of COVID sensitivity through the day. There are food and beverages available, and we invite you to share those and partake in those refreshments as you choose out in the atrium. On your name tags, as you checked in, we asked that you indicate your level of COVID sensitivity with a color coding system. So if you have a green indicator, a yellow indicator, or a red indicator, that should be a flag for those around you of whether you want to shake hands or not. So we ask that you just be conscious of what color folks have chosen on their name tags. If you're looking for the wifi password, I can't remember what it is, but it's on the boards on the side. So if you're in the room, take a look at that.

For all of those who are joining us online, if you have any questions at any time, technological challenges, feel free to drop those in the chat box or email our team at info@datafoundation.org. And then finally, for everyone, in-person or online, join us on social media with the hashtag RegTech22. Now, without further ado, I'd like to welcome our partner in today's event, Craig Clay. Craig is a member of the Data Foundation's board, but also a longtime member of the Data Coalition's advisory board. And I'll also say, one of the longest supporters of the Data Coalition. As a partner for the RegTech Data Summit, really thrilled to have Craig's partnership and support for today's event. So Craig, I'd like to welcome you to offer a few words.

Craig Clay: Thank you very much. I'm Craig Clay, president of Donnelley Financial Solutions, DFIN. I'm a green, just FYI. We are the leader in risk and compliance solutions. And what does that mean? It means we help businesses, doing some of their most consequential events, whether that's an IPO, whether it's a SPAC, whether it's ongoing 34 Act reporting or growing through mergers and acquisitions. We help companies comply with this complex growing web of government regulations. And to do this, we really have amazing array of Software as a Service solutions for professionals to use and as well access our professionals. It's a distinct honor for DFIN in partnership with the Data Coalition to sponsor and bring you the second half of today's RegTech22 Data Summit.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Let me thank everyone who attended our sessions this morning, Accelerating AI in the Public Sector. A special thanks to Nick Hart, who's the president of the Data Coalition, and this morning speakers from all four panels, which focused on how best to use AI in the public sector. During this morning sessions, I hope you had a chance to hear DFIN's own [Tori Ken 00:04:06]. She's a corporate attorney and project manager of our eBrevia artificial intelligence powered contract analytics tool, and DFIN's Darren Wray, who's the co-founder of Guardium, which DFIN acquired last year. It is a product that provides data protection, PII, which is an amazing tool that we've added to our portfolio of cyber security solutions.

We are delivering on RegTech solutions in both the public and the private markets in the focus of this year's events, applying data innovation and regulatory technologies to improve the customer experience in government. Couldn't be timelier. This afternoon sessions began with a closer look at how RegTech and data management strategies are being implemented across the financial services sector in government, by improving daily lives and providing transparency. We're going to explore customer-friendly technologies designed to improve the customer experience, assess risk and ensure transparency.

Our second panel dives into the most challenging data management and reporting trends of our time, environmental, social, and governance or ESG. The SCC is taking up this very topic in an open commission meeting to propose their climate disclosure rules, which is set for this coming Monday, March 21st. The SCC announcement actually sets up the next session today, which is entitled, Using RegTech to Improve Climate Data Reporting, Access and Analysis. DFIN's John [Trezalino 00:05:48], attending virtually, was the chair for the Data Coalition's ESG workshop group. In January, the group delivered a recommendations on improving the usefulness of transparency and ESG data by adopting standardized, interoperable and audited ESG data. You can find and read about this report on the Coalition's website. You can also pick it up, so it is available here today. You can read that during the cocktail hour outside.

Panelists today will consider how to open, assess and share data. How it'll benefit regulated entities, particularly in the context of ESG topics, ranging from standardized reporting to develop large scale analytical platforms. Our afternoon keynote speaker, Hugh [Harper 00:06:35], who's director of the US Government Publishing Offices, is going to discuss innovative ways for the US Congress to improve legislative output. We all look forward to that. And our third panel is going to look at citizen engagement and transparency in rule making using RegTech solutions. With existing technologies, including AI and advanced analytic systems and data framework that standardize reporting requirements, RegTech solutions have the potential to transform how government monitors, regulates, analyzes and submits comments and information and interacts with the American public on regulatory actions.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Lastly, we're going to have a fireside chat with Mike Willis, always an engaging conversation. He is chairperson of Regulatory Oversight Committee, he will bring his insights on the use of RegTech and government today and how it's shaping the future for more efficient transparent and useful data for all stakeholders. Please stop by our DFIN table right outside the door and meet our team, talk about virtual data rooms, venue, eBrevia, which is our AI tool, Data Protect, which is our PII tool, Active Disclosure, our SCC reporting platform, and learn more about how DFIN is revolutionizing the RegTech space. I will now turn the session back to Nick Hart, who's the CEO of the Data Coalition. Thank you.

Nick Hart: All right. Thanks, Craig. And thank you to the DFIN team for your partnership on today's event and for setting the stage for the next segment. We're today, looking forward to what I would say is a very important discussion about applying RegTech in a CX context. And I want to come back to this theme. I also want to share some history. This year is actually a very important milestone for our organization.

I joined the Data Coalition as the CEO three years ago, and it's been a pretty incredible three years. That's not actually the milestone by the way, I'll come back to this. So 2019 was my first event as the CEO in New York city. Roughly this time of the year, I met Craig there, it was one of the very first Data Coalition representatives when I had just been hired. I've appreciated our friendship, the partnership, the collaboration over the years, Craig. In chatting with Craig last night, ahead of this event, I learned we actually share a piece of common professional background.

We both spent time working in the White House complex. Craig, however, was smart enough to not spend a decade working there. But I think we both probably learned a common lesson and it's something that's very rare in this town of Washington, D.C. and it's something that I will carry forever from my time at the White House, humility. Fair?

Craig Clay: Yeah.

Nick Hart: So, we learn an incredible amount from those around us, from our partners, our collaborators and our peers. That's also why events like this are so important and the work that we do. That's why the Data Coalition's partnership with organizations like DFIN are so important. They make us stronger, they make us wiser. And honestly, we get the opportunity to learn and grow together. Reflecting back to 2019, that summit up in New York city, our speaker lineup looked a little bit different. The discussion, actually, not that different though.

We framed it in quite a similar way because some of the challenges that we face are actually very common. How is it that we can, from the outside of government, encourage common sense reforms for our government to modernize, facilitate open data, and ultimately better use data to make better decisions? Many of our speakers from back in 2019 are actually working inside

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

government today as political appointees. Matt Reeves over at the FDIC. Ben Harris is an appointee at the Department of Treasury, but a lot has changed since 2019, hopefully you'll agree, much more than where people sit in their jobs. Obviously, 2020, our RegTech summit that we hosted was virtual. 2021, also virtual. Here we are in a new hybrid setting. We have hundreds of people joining us online.

For many in our audience, we also know you've experienced fair number of personal challenges in your own lives. Things that you may never share publicly. We can only express our empathy about what those challenges may have been. Whether it's as a parent, a co-worker, maybe a manager, it's been a difficult couple years. We're going to be honest about that. My own story is no different. It's been rough. I have a four year old son who has basically known nothing other than COVID his entire life. So I've learned a lot from my son in the last couple of years.

It was occurring to me last night. I was thinking about some of the basic things that he's taught me about. He understands data standards because he likes trains. Do you know how many different kinds of standards there are about trains and train tracks? He's four, he understands standards. Why is it so hard for our government to get this right? Fortunately, there's an interesting counter narrative to what we've experienced. And that includes actually about my son, it's that I get to share with him the stories about what we all do together. There's a lot of hope, there's optimism. There's the work that we do through our Data Coalition with all of you and through our bigger data community. So, I want to give you an example.

My son has actually never known a world that didn't exist without something called the Foundations for Evidence-Based Policymaking Act. Some of you have heard me talk about this before. Many of you might actually know what this is. Some of you, this might be a new concept. It's a law that Congress passed just before Christmas of 2018, signed by the last president early 2019. And it was written just before my son was born. It was filed two days after he was born in Congress and quickly made its way through to become law. Why does that matter and why does that matter for the conversation today?

It mandates the publication of machine readable open data. Interesting. In the US, it directs agencies to tell us what data they actually have. It created chief data officers, a whole suite of evaluation activities in government. And it sets up a whole ecosystem for better using data to make evidence, to improve decisions, to drive outcomes, then improve government decision making. So, let me say all of that just a slightly different way. My son has never actually known a world where the US government actually shirk its responsibilities around learning and improvement. Now, you might have political commentary that you could insert here. That's not what I mean.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

He's never known a government where the establishment, the institutions said that's probably good enough. And I hope that he never does. I don't think he ever will, because now we have a legal framework that changes this entire narrative. So for his generation, when it comes to tackling the wicked problems that are in front of us, climate, economic disparities, educational attainment, global security, whatever challenges the world might face in the 20 years from now. One thing to me is very clear. Data will be center stage. So for all of you, I hope this message is not lost. But how can data be center stage if we don't take the actions today to actually position our government to make this real? The Evidence Act was one step. It's a law. Great, we have a law.

The public forum this morning, highlighted advances in artificial intelligence, the momentum around the national workforce, infrastructure needs, and AI offers so many more opportunities. Congress has passed a lot of other data laws over the past decade. In fact, so incredibly helpful in this narrative, yet most of the American public has no idea that that has happened. Things that our organization has worked on, like the Data Act of 2014, improved our ability to report and standardized government spending information. The American public hugely benefits from that. It's not exactly a topic that people are probably sitting around their dining room tables, having conversations about the Data Act, but they probably talk about what it means. Yeah, you're laughing. You know it's true, right?

Speaker 1: I talked about the Data Act [inaudible 00:15:13].

Nick Hart: Okay. Exceptions in this room, right? Well, here's an example of why that matters. In 2019, I was part of a team at the Bipartisan Policy Center where I'm a fellow. We used information from states because of the Data Act and matched it together to analyze opioid spending. This example is probably a non-sequitur for RegTech, but it's an interesting salient one. Our question was, is federal funding for prevention targeted in a way that actually aligns with needs? Interesting question, right? The answer, which we know because we analyze the data made available because of the Data Act, is that there are some incongruences. We also would expect that because federal funding formulas are often imperfect. But we know this because of the Data Act.

We have better information today than we ever have about federal spending. That's replicable across many other domains. So why does that matter today? Transparency. Today, the American people have better access to information about how their government spends funding than at any point in history, accountability. Access to information means they can also ask questions of their government about how we're spending. Innovation, Data Act encourages new technologies and tools for agencies to reduce burdens. Efficiency, this all means cost effectiveness and efficiency for taxpayers across the four trillion dollar budget.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Policy makers are daily making thousands of decisions about allocating federal dollars even after appropriations are made. All right, I said the term four trillion dollars. Does anybody in the room have a way of contextualizing four trillion dollars? Do you really understand what four trillion dollars is? I don't. I used to work in the office that thought about dollars every day. So I'm going to give you an image. Imagine that you are stacking pennies end to end across the United States coast to coast. A penny is three quarters of an inch wide. The budget is four trillion dollars. You wouldn't do it once, you wouldn't do it twice, you wouldn't do it three times. You'd do it 16,000 times. That's four trillion dollars. Have fun.

So, this is important to get writing. The American people deserve to understand how we're spending dollars. We have other laws. The GREAT Act helps us understand the same information for grant recipients. I can go on and on. These are laws that are bipartisan. They're not controversial. And the good news is, they're being implemented today. All right. So let's do a little thought exercise. Close your eyes for a minute, whether you're virtual or in-person, imagine your life pre COVID. We didn't know what COVID was. There's daffodils. We're not debating face masks. We didn't know what PCRs were. Donald Trump's on The Apprentice. Barack Obama's in the Senate. We're way back in time.

If we had our crystal ball, what would we say about the world to come? What would you say? What data issue could you solve for? What would be the big thing that you would solve for? Some of you are inevitably thinking about COVID. You couldn't help it. There are data gaps. There's the reporting chaos that you're predicting. You're thinking about data governance frustrations and data linkage issues. There's ESG, people in the room, so you're thinking about those challenges. Some of you are thinking about the things that happened in the last three weeks around the world, coordinated sanctions, perhaps, around certain countries. Some of you might be thinking about census and some of the enormity of the tasks, whatever these issues around data are, there's tons.

We have tons of challenges around us when it comes to data. Some of you went deeply, and we're probably thinking about the technical, the technocratic stuff. Some of you were thinking about concepts and themes. And right now in this moment, the summary for me is that we're all innovating. We're all developing statistical advances, we're influencing important topics and issues. Whatever you're working on, whether it's anti-poverty, banking, environment, climate, housing, equity, data, are there. So let's acknowledge it, let's embrace it, and let's plan for it. Let's ensure that when we look in our crystal ball going forward 20 years, that we're not blindsided or ineffective, incomplete when we think about our data infrastructure. Okay. So let me bring this full circle and tell you why I'm talking about all this stuff.

The laws that I've just described over the last decade are opportunities. They're all about things that have encouraged open data, evidence informed policy making, and modernized disclosure systems. And that's what we do at the Data

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Coalition. This is our emphasis, this is our priority, and it's why we're America's premier voice on data policy. That's what we've done for the last 10 years. And we've been a driving force in all of the things that I've just described.

Our members and our staff have been part of these processes and dialogues over the last decade. And I say this all with an incredible sense of humility. This is a team sport. And our partners and collaborators, whether you're on government, academia, nonprofits, or industry, are all instrumental to the success of this whole community. And at the end of the day, as an initiative of the nonprofit Data Foundation, we're here to help have a common voice and an advocate for high quality accessible and useful government data.

This is a drum we're going to continue to beat. It's not just a message about us touting the coattails of projects that other people are working on. We're here to develop partnerships and strategic advocacy, and sometimes that even means direct lobbying. Yes, we do that sometimes. We value the expertise of our members, whether you are a small nonprofit or a fortune 50 company, because we need the range of expertise that is out there. And we learn a lot from this expertise. So if it's not clear by now, the milestone that I'm coming back to is actually a celebration.

This is our 10 year anniversary. The Data Coalition has now been around for 10 years. And all of the things that I have just described around the legal, the laws, the successes around the implementation are because in part, we are celebrating that 10 years. So we have the two co-founders of the Data Coalition in the room, and I just want to quickly acknowledge them. On my left, John Runyan, on my right back there in the corner, Hudson Hollister.

So Hudson, John, thank you for many years ago, coming up with this idea. This is an organization though, that is about more than a single company, more than a member or a person. It's about a community, then we're stronger together. We're more effective when we accomplish a shared goal. So at this moment in time, we know our government still needs advocates that call on this improved use of data for our government and for our society. And that's a task we're going to continue to take on. So I hope that you'll continue to join us in this endeavor in the years ahead.

Today we're not here to talk about shiny objects. We're here to talk about solutions. We're here to talk about fixes to the problems in front of us and the years ahead. Some of these are going to be big wicked challenges, and that's really the forte that we seek to endeavor and solve for the magnitude of the problems that are ahead of us. So, flying RegTech and government is that the scale we think that has huge benefits and it's an incredible opportunity. We use the president's CX memo as a way of framing this dialogue, because government has a tense relationship with some of its customers. And we can be honest about that in business, that would be catastrophic.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

This is an area where we can see some comments and some improvements. They're long overdue, and I could keep going on and on and on. I'll just maybe stop there and say, we are in the business of improving our government. We know we have the opportunity to help it improve. That's the business of what the Data Coalition is about. Hopefully, my son will see this government and this reality come true in the years ahead. So let's work together and solve some problems and let the dialogue begin. I'd like to begin by inviting co-founder of the Data Coalition, longtime friend, John Runyan, to the stage to moderate our first panel, which is about data accessibility and customer experience with financial regulators. John, I'll let you take it away.

John Runyan: Sure. Well, good afternoon, everyone. Let me just exercise the prerogative of the microphone. My colleague Hudson and I were graciously acknowledged, but can we just say a word of thanks to Nick and the entire Data Coalition and Data Foundation team for the incredible work that they continue to do day after day throughout the pandemic and beyond. So thank, Nick. Thank you all.

So, I was excited when Nick asked me to moderate this panel at this in-person event. But he said, "Well, you'll be in-person, but your panel will be virtual." So we continue to learn and grow throughout the pandemic. The topic today of our group is Data Accessibility and the Customer Experience with Financial Regulators. We're pleased to have a terrific group of experts, Tammy Roust, Dr. Tamer Roust if you're not familiar with her. But Tammy is the director of the Division of Data and chief data officer at the Commodity Futures Trading Commission, where she's responsible for overseeing the agency's enterprise data strategy and data governance approaches.

Adam Scott is director of design and development at the Consumer Financial Protection Bureau, where he describes his role as developing technology tools for consumers with a positive social impact. And finally, Darren Wray is the co-founder of Guardum, a leading data security and privacy software provider that's now a division of DFIN. Darren has a long history in the implementation of business and IT processes that handle personal data. So let me start by turning and asking each of our panelists, and we'll go in the order of Tammy, Adam and Darren, to fill in a bit of detail about your current role, but also ask each of them to address the question of who are your customers and be they internal or external, and how do they influence the way you package data and relay information? Tamer, Tammy.

Tammy Roust: Yep. Everybody calls me Tammy, including my mom. So, I'm Tammy Roust, I'm the director of the Division of Data and the CFTC's first chief data officer. That means, as John stated, I'm responsible for our data governance approach, our enterprise data catalog, as well as our AI and data science efforts and machine learning. I have to start off by, after I tell you who I am, to give you the obligatory disclaimer that my remarks are my own and the opinions, analyses and conclusions expressed by me do not necessarily reflect the views of the

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Division of Data, the Commission, other commission staff, or the US government itself. There, we've got that out of the way.

So, who are our customers? Unlike my [inaudible 00:26:36] colleague at the CFPB, which I always get wrong, and if I get it wrong, I apologize to Adam ahead of time. Our customers are a little bit different. Although we serve the American taxpayers, we serve them by ensuring that the futures and derivatives markets are fair and transparent and have integrity. That benefit isn't obvious to your average American taxpayer. I can say so from experience. I grew up in a place where the hog report was on the new news every day. That's produced by the Department of Agriculture, but the hog futures contract is regulated by the CFTC. I had no idea who the CFTC was until I actually started working in derivatives.

We operate behind the scenes of finance. We're not right out in front the way the consumer facing agencies are. You can't buy a box of Cheerios without touching the markets the CFTC regulates, but everybody who buys a box of Cheerios does not know who the CFTC is and about how we work to make futures markets fair and transparent, including the oats markets that's incorporated in that. So to make what is really going on in those markets more transparent, we package data and relay information. We generate the weekly swaps report and the commitments of traders report and other reports that we put up on our public facing website.

Those are then packaged more for the business press and institutional entities who are aware of our existence, and some of whom we regulate. They then convert that technical information and relay it further downstream to the people who are end users and producers, people like craft and general mills who makes Cheerios, and your average farmers who are looking to risk mitigate by going into the future's markets, as well as institution entities. Keep in mind, there's nothing preventing your average American tax payer, your hog farmer, or URI from going to the public facing website and downloading that data and parsing it out themselves. It's posted on the public facing website. It's just that it's drilled down into further detail by the institution entities. And that's our customers.

John Runyan: Great. Adam.

Adam Scott: Sure. Thank you. So I'm Adam Scott, I'm the director of design and development at the Consumer Financial Protection Bureau, where I oversee all of our custom web and application development, as well as our design team, user researchers, content managers, our creative arm of the organization. And like Tammy, I'd like to apply the standard disclaimer, that anything I express is my own opinion, not necessarily that of the organization or US government. I want to also acknowledge Tammy. She set the stage for this nicely because I certainly think of our customers at the CFPB as being the American people or consumers.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

The Dodd-Frank Act created the CFPB to protect consumers of financial products or services and to encourage the fair and competitive operation of consumer financial markets. So in my mind, everything that we do, whether it be direct to consumer information, regulatory filing guides, RegTech systems that bring in data is all being done in service of real people. So keeping that in mind really helps us to focus on building great experiences for everyone who interacts with us, whether it be direct to consumers or with industry or researchers. Because ultimately, the goal is to make an effective and worthwhile experience for the people we serve.

John Runyan: Great, thank you. Darren.

Darren Wray: Hi, I'm Darren Wray. I'm the co-founder of a company called Guardum. We are now part of DFIN, as Craig mentioned earlier on. It's a fantastic organization and our customers are widespread. Not doing all the things, all the good governmental things so much as Tammy and Adam have described, but we are working with the law firms and industry, helping them ensure that they comply with regulation certainly around data protection and data privacy and protecting those individuals. This is something that's increasingly important in the United States. California's led the way, Virginia is not far behind now, and Colorado following behind them too.

So our customers are those organizations that are protecting the data, finding that personal information. It's usually a critical time in their business. It might be as part of mergers acquisitions. It might be that they've been approached by regulators and they're being investigated, or they're having to provide information that has to be anonymized or made safe to provide to the public. And that could be law firms, utility companies, large manufacturing organizations, and yes, government. In this case, the UK government using our products to help ensure that they remain compliant and can respond to those regulatory requests. Those are our customers.

John Runyan: Great. Thank you very much. Tammy and Adam, let me ask each of you as representatives of government agencies, how the president's recent executive order on service delivery and the customer experience has altered the way your agency approaches public private interactions? Starting with Tammy, can you each address that for us?

Tammy Roust: Sure. So as I mentioned, we make reports available on our website that are then consumed and interpreted by others further down to provide information downstream. That current service provision is static. The report's provided and left to users to glean the information they want out of it. And as part of our response to the EO on customer service, as part of our overall vision, we're moving towards reporting with visualization and open data portals. That will reduce the time tax that's discussed in the EO on customer service.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

But keeping in mind, however, to [inaudible 00:32:58] back on something Darren said, is that as a regulator of systemically important financial institutions, we carry data that has privacy requirements embedded in it because it contains material non-public information. So while visualization will allow the customer to drill down into the markets they are interested in, the products that they're interested in, there will be limits on what we can share publicly in a data exchange. Adam.

Adam Scott: Yeah, sure. Thank you, Tammy. I thought that was a great question. I think it's really important topic and I was excited to see this executive order come out because it places the emphasis on people. I think the subtitle, the executive order was putting the public first. Now, for my team and I, it has had minimal impact on our day to day because I think this is how we've approached our work already being user centered and bringing that design aspect to our work. But it's certainly added some legitimacy to that and a little extra emphasis.

One of the areas of focus that I really saw on the executive order was on what I might think of as moments for the American people. Things like retirement, financing a business, times of crisis, such as poverty. I think in the long run, where we have a lot of opportunity within the government, is to see more cross agency collaboration in some of those focus areas where we've seen that be successful in times of crisis, such as COVID. But now more and more, how can we as government agencies work across ourselves to build in customer experiences or user experiences that help meet the needs of the public for their needs, not with an agency first mindset, but with a people first mindset?

John Runyan: Great. Adam, just picking up on this, talk about... I know you have both internal as well as external customers, the agency. You've spent a good bit of time on your...

PART 1 OF 6 ENDS [00:35:04]

John Runyan: ... Those external customers, the agency you've spent a good bit of time on your internal side. Can you elaborate about some of the efforts you've done to encourage the use of the website for the internal audience that you have at the agency?

Adam Scott: For the internal audience at the agency? Yeah, sure.

John Runyan: Yeah, I think that was-

Adam Scott: So we do build a number of tools and things for our internal audience and seek to meet their needs. I think that's an area that is really valuable as an organization to look at, where is the data that we're receiving and coming in and how can we help build out tools and meet the needs of internal users the same way we meet the needs of external users, so that they can access those and make sense of that information.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

John Runyan: Question for each of you, talk about the role or how you would describe the interplay between data, technology and oversight. How do you think about that issue?

Tammy Roust: Well, I'll start. Technology is a tool-

John Runyan: Starting with Tammy.

Tammy Roust: ... like a flashlight. Data shows us where to point the flashlight, where to focus our oversight efforts. So it's iterative. Sometimes for us it means having a conversation with the markets about what we're seeing. Sometimes it means rules being written to get more or different data from those markets. Sometimes it means bringing in other tools like other technology based on those conversations. So I go and talk to the markets and ask them what they're seeing in terms of their cybersecurity efforts. And they say, we are using this tool for this particular purpose. And then we go back and figure out whether that would be useful in our paradigm as well, for example.

John Runyan: Great. Adam.

Adam Scott: Yeah, sure. I think of that really as an ecosystem. Traditionally this interplay has been looked at as separate units. We might have technology, we have rule writing, we have data collection, market monitoring and research. And I think within government, we could really do a shift in that space. There's a famous saying that's kind of been popularized in business world that every company is a technology company. But the root of that comes from the influential technologist Anil Dash, who said that every industry and every sector of society is powered by technology today and being transformed by the choices made by technologists. And I think that's a really a critical mental shift that regulators can make, where we no longer think of technology as a byproduct of our initiatives, but as something that's core to our work. And the best way I see that gap being bridged is by having technologists in the room with decision makers and with those involved in oversight data collection.

And that can help shape policies in ways that are data enabled, anticipate technical needs and I'd ultimately build far reaching across cutting solutions. So the more we can integrate that. And I think a key aspect is that technologists shouldn't always include programmers and data scientists. I'm biased, I come from being a software developer, but I also lead a team of designers. And I think great designers bring a focus on forming solutions for real people. And having designers perform user research early on often as part of policy making, getting feedback and working with supervisors and technologists to iterate prototype, get feedback, and introduce human centered design principles could really go a long way towards integrating all those ideas.

John Runyan: Thanks. Darren, before I ask you to answer that same question, I just want to put the audience on notice if there are some questions from the room, just be

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

thinking about that. And in a moment we'll have a microphone for you to pose those questions. So Darren, please go ahead.

Darren Wray: Sure. Look, my view is perhaps coming at a slightly different angle. The way that I've seen things develop around the world particularly in aspects of information security and cyber security and data privacy, where I spend a lot of my time, the oversight and the regulatory requirements and legislation that's been passed, it's really helped organizations focus. And that focus has required them to adopt more and more technology to be able to do this job more efficiently and automatically wherever possible, because that brings that consistency in terms of the results you produce and the speed that you can produce it at. Now, the other benefits that come through that though, is the consistency of data and sure the protection piece that I've already spoken about is exceedingly important. But it's also about being able to identify that information and classify that information.

And whether the end purpose be for ESG or data privacy or in response to a data breach, all of those aspects, all of those things require that greater understanding analysis of data. And with the organizations that I speak to regularly, that's very much being driven from an oversight requirement because there's a regulatory requirement for them to be able to report on this. And the technology, I love Tammy's example by the way, or analogy of it being a flashlight. I always refer to it as being a magnifier. It means it's a tool that you can, in the same way, as a wheelbarrow allows you to ship more things and move things around or a truck allows you to do so, a machine and software and technology enable you to do these things. So one person can do all that much more good and be more powerful and more empowered to actually do that work with technology. It's essential in the world that we live in today with the amount of data that we are dealing with, we couldn't do it by hand. Some people still try unfortunately, but we can't do it by hand.

John Runyan: Great. Thank you, Darren. I don't see a microphone in anyone's hand. Do we have some questions from the floor? Yep. This one over here. I guess, bear with us one second. Great. Go ahead and state your name and organization too if you would please. And then tell us who you're directing your question to.

Karthik: Sure. My name's Karthik, I'm from AWS. So I'm kind of intrigued to ask Tammy Kochan around, because you mentioned about this spotlight thing. I'm always curious to find out because the pace of or the acceleration, because previous in the morning we had discussion about acceleration, the pace of outcome that we tend to produce from the public sector sort of things, whether it's accessibility or AI or customer experience, and correct me if I'm wrong, it tend to be a little bit slower than what we see in the rest of the world, in the pace of development. And that is framework and process and policies and procedures in the public sector. And they really should be awake to accelerate that and faster and foster the outcome for the benefits of citizen. But for somehow or for some reason, we tend to shy away from that acceleration and the pace of

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

development. So what do you think is the synergy that we can bring along as a community, whether it's public or non non-public sector to accelerate that growth?

Tammy Roust: Wow, that's a tough question because I haven't figured out in like 30 years of public and private sector service, how to accelerate the public sector to the speed of the private sector. I won't say you're wrong. I think the public sector does move at a speed that lags that of the private sector. And I actually don't think it's a bad thing. It's very expensive to be on the bleeding edge. And to echo something Nick had talked about earlier about the cost effectiveness of spending, when you're on the bleeding edge it's expensive and you tend to move in one direction and then have to completely change direction in response to external forces.

To your point, how to accelerate the public sector faster than we are doing? So currently again, I've been working in this for 30 years and I don't have a great solution. I think if we had more public private partnerships and more academic partnerships, that we would have a better spotlight, better technology in seeing what is actually going on in the private sector. So that would accelerate our efforts. We're still not going to be at the speed of the private sector, we don't have the funding, we don't have the human capital to be in that direction. But as I said, I think it's more cost effective. I've said since I left hedge funds, that the great part about being a regulator is that I can see the bleeding edge. I don't have to be on it. Being on the bleeding edge means you get cut.

John Runyan: Adam or Darren, any comments?

Adam Scott: Yeah. I'd be happy to add to that. I think Tammy is exactly right. I really agree with a lot of things she said. I think of the old Kool-Aid man commercials, I don't know how many of you remember that where the Kool-Aid man would burst through the wall and give everyone a glass of Kool-Aid. And that was great. Everyone had a nice cold drink, but they also had a big hole in the wall. And I think in government, working within government often, we want to open the door and bring everyone a drink. And that might take a little bit longer, but we want to still meet the needs and not break things along the way. So that requires a little more caution. And the thing I've seen really work well for us is to hire great talent and really recruit great talent, to try to put ourselves on par with public sector in terms of technologists and others within the organization. And give them the tools and runway to do great work. And that has been a great way to accelerate innovation and change within our organization without breaking too many walls.

John Runyan: Darren, anything to add?

Darren Wray: I'd completely agree with Adam. I think it's about talent and knowing the bounds of where you can make change, where change can be achieved quickly. But as Tammy says, it's sometimes a good thing that the federal government

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

doesn't move quickly and break things because the consequences can be catastrophic with that. But hiring in good and smart people from private sector with those kinds of experiences can really help accelerate things, giving them the bounds and the scope to actually do good things. Completely agree.

John Runyan: Question back here.

Mike: Hi, Mike Anderson with Informatica. This is for Tammy and Adam. Generally speaking, king and queen for a day, given the data wrangling you do in your day job, what is a technology that you're aware of that's pretty advanced in industry that you would like to get your hands on to be able to do your mission better?. Thank you.

Tammy Roust: Just one? I would really like to get my hands on natural language processing toolkits. Yes, we have open source available in Python and with the digestion of all of the unstructured data that we have. And I would love to set them loose and have them do data discovery throughout our organization so that we have what Peggy was talking about with the data in motion. We have an enterprise data catalog that feeds from the data that is in motion when it comes in, so we know that we have it when we have it. And that we are ingesting all of this unstructured data that we get in and pushing it into a structured data form, because we have a tremendous amount of futures market experience in this organization. And it is not their subject matter expertise to be trolling through PDF files, to find the one nugget of information they're interested in. That is not the best use of human capital at the government or elsewhere.

Adam Scott: And I could say, I would love to have platforms as a service really. How could we not think about infrastructure and have the tools that we need in the cloud environment to quickly add and test new tools and new capabilities. Find out what works and what doesn't for us. And to think about that as a service, not as sort of modern day equivalent of a server. That would be my dream state. I think we're working towards that in a lot of government agencies and I've seen that happen. So I'm hopeful. And also I might be influenced by the previous question of where we can move a little more quickly.

John Runyan: Great. Thanks. Another question in the room right here.

David: Hi, this one could be for anyone who would want to jump on it, because you've all talked about really, really talented people. I should introduce, David Carlos, [Peace Work 00:49:55] we're a data for good NGO. These really, really, really talented people we're all talking about, they all have jobs. And I have had rather limited success at going out and hiring these incredibly talented people who are all there and most of them are all have jobs and most of them are pretty happy with it. So my question is how do we work to grow our own, to develop our own resources, what structure we can put in place to create the data scientists of the future?

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

John Runyan: Tammy, you want to take a crack?

Tammy Roust: Yes. Thank you. I heard David talk this morning and I also heard the last speaker in the morning session talk about the difficulty in hiring data scientists in these organizations and talking about data for good. And how do you build that talent stream? You start at the beginning, you build intern pools and relationships with universities so that you are there on the ground floor, they have a relationship with you and you can build on those interns. This gets back to the AWS speakers remark about how you accelerate change and accelerate technology growth.

I had a point and it was this, David's point was about using data for good. How do I keep those people once they're here? Which is, I make them care about the mission. When I worked at a hedge fund, I used to say, "I don't take money from widows and orphans. I make it for them." I cared about the mission. I was managing pension fund money. I am at the CFTC because of the mission. We make sure that the agricultural and swaps markets are fair for the American taxpayers and when you and I go and buy a box of Cheerios, it's not \$6 unless we're buying it at 7-Eleven or something like that.

Adam Scott: Yeah. I wanted to give Tammy a high five through my computer screen when she said that care about the mission. I think that, that's something I really found is to recruit people that are already mission focused and to talk about the scale of impact that one person with that skillset and our organization can have is tremendous, to touch millions of people and have a tremendous impact there. And then I think the other piece that I've really found is to lean into what makes us unique. Working within government, we're not working to generate profit. And we're often working to share information. So at the CPB, we have a great open source policy where our code is open sourced by default. And that's a great recruitment strategy for technologists because everything they do can be seen by others, they can contribute and encourage to contribute to the open source ecosystem and be a part of that in their day job, because we see value in that both to recruit, but also in what we produce and contribute to.

So I think the more we can lean into those unique factors that you can have in government or in nonprofits, that's another just recruitment strategy to pull folks in.

John Runyan: Great. Thanks. And Darren, you must have some thoughts on that.

Darren Wray: Yeah, absolutely. I'd agree with what's been said, but I'd also say that one of the big differentiators is the scope and scale of the challenge. A lot of these smart people are looking for those big challenges and helping them understand the scope and scale and the difference that they can make, it's another way of talking about that difference. But really talking about it from the challenges that are there and the opportunities to make a difference and actually deal with scale that you can't deal with in industry necessarily. I think those things are

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

attractive to technical people and to these talented people that you're looking for. Whether they be data scientists or your other skillsets too.

John Runyan: Great. Another question from the back.

Alex: Yeah, thank you. Alex Light from [Maida 00:54:16]. Adam, you mentioned moving to, I guess based on the customer experience, EO kind of a, maybe we'll call it like a life events type of model for service delivery. I was wondering if you could talk about what your kind of long term goals are for CFPB, and also what are some of the kind of bureaucratic and technical challenges around that?

Adam Scott: Yeah, I think that's a tremendous question. I think that's an area that there's opportunity and that we're not fully exploring as a government wide. I think ideally it's done in partnership with other organizations that share similar mandates if we think of a student life cycle, CPB might have considerations around student lending, while the department of education has similar also lending but also educational opportunities and grants. And so where are we collaborating across government? So we're almost seen as a single organization or providing a single experience. What we've been doing at CFPB within what we can control is really centering on the users and focusing on design sprints, where we're doing quick iterative work, prototyping and doing constant user research. So we're doing the research, looking at our analytics numbers but also just talking and connecting to real users so we can see the feedback real time and adjust accordingly.

John Runyan: Great. Darren, let me ask you this. You've obviously had big focus in your career on privacy and which cuts through all of this. Can you elaborate a little bit more on your perspective on the role of privacy and data security for compliance reporting?

Darren Wray: Yeah, absolutely. It's a massive aspect globally, not just within the United States. And as more data is required whether it be for financial reporting, tax reporting, ESG, any of those aspects, more and more personal data is there, more and more personal data could be breached, shared inappropriately, misused, all of those things. So it's a really important aspect to have the right type of controls, the right kind of monitoring in place. So this takes place automatically. It's not someone making a decision on a case by case basis at all. The volumes are too great. The risks are too high. The consequences can be devastating. And we've seen this with lots of different organizations and governments around the world where mistakes have been made and data is leaked inappropriately. Whether it be from controversial organizations whose data gets leaked for various different reasons.

But however the data gets leaked, for whatever purpose it's devastating. So you have to have the right level of controls. You have to go into this eyes open understanding the nature of the world that we live in, is so data driven these days. And it comes back to a point I can't remember whether it's Tammy or

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Adam made earlier on about having the right people in place, understanding the decisions that are being made and the controls that are being put in place. Both at a regulation level, but also within your departments or organizations, it's absolutely critical. You can't have the level of regulatory reporting that's going to be required today or it's required today and is going to be required in future, without people trusting that their data is going to be safe. Either on organizational level or indeed a governmental level. So it's absolutely critical.

John Runyan: Tammy or Adam, any comments on that from your perspective as a regulator?

Tammy Roust: Yeah. So as I started off with, the privacy requirements are utmost in mind. As I had discussed earlier, our customer facing component is where we publicly produce data. And that data has been scrubbed to make sure that we're meeting all of those privacy requirements, but there's also surveillance monitoring enforcement efforts. And all of those efforts have to be kind of Chinese walled to make sure that the privacy requirements for those groups are separate from the public ones that we report out. I think probably Adam has the same problem.

Adam Scott: Yeah. Yeah. I would say just, echo what Tammy and Darren were saying, privacy has to be of paramount importance to us. And the more that we are conducting data collection, the more critical that is.

John Runyan: All right. Well, folks, this has been terrific and we really appreciate the panel's time. Any final short thoughts, from any of you before we wrap up?

Tammy Roust: RegTech now, RegTech forever.

John Runyan: Let's give them a... RegTech forever. All right. Well, let's give them a round of applause. Thank you all very much for this [inaudible 00:59:56]. Back to you Nick.

Speaker 2: All right. Well, thank you all so much for sharing your great insights. This is a fantastic discussion to get us going. We're now going to take a short break. And for those who are here in person, we have some beverages and snacks outside for you. We will return here at 2:00 PM Eastern Time and we'll get started at 2:00 PM sharp for an all person live panel. So we'll see you at 2:00.

Using regulatory technologies to improve climate data reporting, access and analysis. This is a live panel for our event this afternoon. And it'll be moderated by Dean Ritz. Dean is the president of the Data Foundation's board of directors and also a member of the data coalition's advisory board. I'll also share, Dean is the author of a fantastic paper from the data foundation on machine readability, which you can find at datafoundation.org. And he happens to be an expert on a great number of issues including, data reporting, machine readability and some of the key issues that this panel will discuss. I will now

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

invite the panel and Dean up to the stage and Dean will introduce you to our panelists. Dean, take it away.

Dean Ritz: Great. Nice opportunity to get some great ideas out into the meet space. I'll be yet another person who's feeling a little nostalgic and glad to be out of the house and here amongst people. Well, here on the panel today, we have three panelists as you can count. We have Kirsten Dabo, who's the chief data officer at the Federal Energy Regulatory Commission. Seated next to her is Kevin Richards and he's vice president and head of US government relations for SAP. And right here, Christie Howell, who is a senior economist from the International Monetary Fund, the IMF. Well, we're just going to instead of having opening statements we decided we would have sort of customized questions for each and you will learn about their work through the answers to the questions. So way down there, Kirsten, if we could start with you, if you're all right with that.

Kirsten: Sure.

Dean Ritz: I do want to mention one thing, I want to give a shout out to the Federal Energy Regulatory Commission for its successful recent completion of a project for moving some of its reporting to semantic data. Which is something that is exciting to me because as we heard for those of us who were here this morning, the importance of eliminating the ambiguity and therefore increasing the trust in the data is a means of creating value. So I just wanted to express my appreciation for the work that you have done and are doing over at the federal energy regulatory commission. Now, to the question. Can you tell us about how a CDO addresses data governance and how data governance might play into these big questions that we face about ESG and climate?

Kirsten: Sure. Well, thanks so much for having me here today. So I guess obviously I'm the government person here on this panel, but CDOs per the evidence act are responsible for lifecycle data management, plus about 13 other functions spelled out in the foundations for evidence based policy making act, thanks to Nick. So we have a number of specific requirements that are spelled out in the evidence act that we need to meet. And the evidence act is unique in that it requires all agencies not just the 24 named CFO act agencies to comply, but all agencies. So that includes all of us small agencies, all of us independent regulatory agencies. So the CDO role is pretty new to us in the small agency space. So when talking about what that means for data governance, fundamentally data governance ensures that data is consistent, trustworthy and doesn't get misused.

So these governance responsibilities are core to effectively managing that data life cycle. Effective data governance and a managed data life cycle designates trusted data sources, it maintains a comprehensive data inventory, it defines data standards, determines policies and strategies for promoting enterprise data management activities, promotes data stewardship to facilitate data

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

sharing, collaboration, data quality. So the ability to answer big questions, and I often use a cake baking analogy here, how do we bake a cake if we don't know what ingredients are available, what recipes should be used to produce the right flavors? Would we eat the cake if we didn't know whether it contained allergens or toxins? So governance is very important to making sure we know we have. So really it's the same for answering the big questions with data and why governance is so critical to this process. Data governance helps us know what data we have and who stewards the data to ensure safety and quality of the data.

It identifies the appropriate standards, logic, transformations and business rules that the organization agrees to collectively. So that anyone looking at a data product understands what each of the data elements means. This ensures we don't misuse the data and this really reduces our risk. And to really touch on how our role is changing, I mean as I mentioned earlier, the CDO role is pretty new across the government. And I think very specifically to the small agencies. And we are really the newest to the C suite.

So we're really very much in the formative years. We're still defining our roles for our organizations, we're putting together our resource and budget justifications and we're really looking to gain traction on data management, data governance initiatives. Our organizations are still figuring out what data management data governance means to them and how to integrate in ways that are minimally disruptive to our ongoing work. And I know when I talk to my peer CEOs across government, it's a regular topic of conversation as to how our role is evolving. We want to ensure that we're positioning ourselves. As responsive as possible to the various needs of our organizations.

Dean Ritz: Yeah. You noted something which is really, I don't think is common knowledge, is that the CDO wasn't common in most of the agencies. So it suggests that is it rising awareness of the importance of data and that data isn't just a CIO task, but actually it requires a different level of attention and expertise because of the value that's in the data, not just operational.

Kirsten: Absolutely. I mean, we're very much in those learning years. So there's an education happening in our agencies as to why data is a separate function from CIO. I mean, even where CDOs report to CIOs, having an official CDO whose responsibility is really the life cycle of data management, that's an important distinction. It's not a technology implementation. I mean, just because data often lives in technology doesn't mean they're the same thing.

Dean Ritz: Right. Right. Very good. Well, thank you very much for the remarks. Kevin, let's go to you next because this rising awareness that she speaks of is something obviously it's a focus of a lot of your work. It's noted here that you attended COP26 recently. And so curious about your perspective on how these issues that are obviously the focus for a lot of people, are influencing data strategy, right?

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

The need to answer these questions to understand and it's a global need, how it's affecting data strategy if you can speak to that in your work.

Kevin Richards: Yeah. No, thank you Dean. And thank you to the data coalition for having me here today. It's a pleasure to be at an event in person. I think you can all see we have paper in our hands, I think we got used to have our laptops and our little cheat sheets so excuse the extra paper. But it's an important topic and representing SAP, it's the world's largest enterprise software company and data being at the heart of our business model, I want to make sure I have my data accurate and accessible for everyone. And I'm not sure if all of you are familiar with SAP, but we have a 500,000 customers across 26 different industry lines of business. So while we really did an event, our pulse is certainly on the digital economy. And when it comes to sustainability, SAP as a company is very focused on helping ourselves and our customers achieve the goal of zero waste, zero emissions and zero inequality.

And we really believe that data can shape the future of sustainable living. And we really want to help our customers be more sustainable. And our solutions really help businesses embed operational, experiential and financial data and insights to drive sustainability at scale. And we're committed to partnering with our friends in government on their sustainability efforts. We've embarked on an initiative called Climate 21, which we discussed at COP26 that focuses on promoting discussions around sustainability and helping our customers both in government and the private sector reduce their carbon footprint. In addition, SAP has been at this for a while as we've been the software industry leader of the Dow Jones sustainability index, is for the past 14 years. So as an industry leader, it was really our first time that we attended a COP-

PART 2 OF 6 ENDS [01:10:04]

Kevin Richards: ... was really our first time that we attended a COP, in COP26. I think what we found interesting is that business was really prevalent at the conference and the leaders at Glasgow, the conference recognized the critical role of businesses and driving change. It was the first COP our representatives attended. We listened. We were engaged.

The Glasgow Climate Pack to really give a mandate to countries to do more to strengthen their 2030 targets over the next year. But it also recognized the importance of international partnerships and the development and the deployment of technologies and technology transfer. What we really found in interesting, as we heard repetitively, the phrase, "You can't manage what you can't measure," in almost every forum in every stakeholder group.

This was repeated and reemphasized throughout the conference, which to us really explicitly linked to strengthening the role of strengthening data and digital solutions for climate action. Again, big presence by business. We certainly saw that there's a significant role for companies like SAP to play as a stakeholder, as

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

we run the operating systems of businesses and helping our customers to get grips around the climate industry.

First and foremost, the question of getting transparency and alignment with businesses' processes that drive the challenges that we have in the climate space. Consumers are really driving demand of businesses for transparency. Now during CEO calls, their shareholder calls, 44% of CEOs say sustainability is the number one topic. That's an increase over previous years.

In Q1 of 2021, the word "climate" or "sustainability" or a synonym close to that, was mentioned over 4,700 times on corporate calls, on shareholder calls, which we found interesting. That trend by businesses is going up, and we really helped them with their ESG strategy and understanding sustainability and the impact that it has on their operations and on their customers.

Dean Ritz: Very good. Thank you very much for that. Kristy, all right. That was interesting. When we talk about climate data in... We know there's an event coming up on the 21st where the SEC might be announcing some of their plans for collecting ESG data as part of public company reporting processes.

In that space, some people use the term non-financial data versus financial data. Some of us promote the term indirect financial data, because these act that we measure actually do have some impact on the financial performance of an entity. When we get to the International Monetary Fund, which is about finance, and we have Kristy here for that, we're interested in your work at the IMF and how that work assists the transparency demands by governments, by investors and by NGOs.

[inaudible 01:13:13] you tell us a little bit also about your background. The math nerd, I think, on the panel here.

Kristy Howell: Well, I'm an economist. I would say, we usually say we like to estimate things, but yes, a little bit of math background. First of all, I want to say thanks to The Data Coalition for having me here today. I'm a senior economist at the International Monetary Fund. But prior to that, I did spend 16 years in government at the Bureau of Economic Analysis in the Department of Commerce. I guess that's where my math background comes from, and more specifically macroeconomic statistics.

I do have a little bit of a different perspective when it comes to looking at these climate change and ESG issues, because at the IMF, we're focused on the macro economy. Whereas, a lot of the discussion is around what's being done, as Kevin was just describing at the firm level, but there is this interesting nexus when you think about where that macro and micro meets.

That's where a lot of our work is working with countries on how to improve data gaps around, for example, how economic activities are impacting the economy,

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

and really to understand that you need both the macro average statistics, but you also need the information about where these firms are located and what industries they're engaged and all of that. It really brings both those things together.

Two things I wanted to mention about the work that I'm doing at the IMF. First of all, one thing that's interesting is I joined the IMF last September, and they had just recently started to engaging more with their members on climate change. Some people hearing the IMF, which is an international organization focused on the macro economy and financial stability, might be surprised that they would be so engaged on climate change.

But of course, if you think about it, climate change is probably the single biggest potential impact on long term growth for economies. In that sense, the IMF sees it as what they call a macro critical issue, and so climate adaptation, mitigation policies, all of these things are going to be crucial towards ensuring a smooth transition to a new low carbon economy. This is something that is very important to the IMF as it works with its member countries.

But in the data space, which is where I live. I work in the statistics department there, the IMF developed a climate change indicators dashboard, and this is something that we launched last April. The idea of this was to really be a first step at trying to put something out there to address this emerging need for more information on climate and both how climate impacts the macro economy, but also how the economy is impacting environment. It's looking at both sides of that.

This was a dashboard that we developed relatively quickly and working a lot with other partners from other international organizations and other agencies. There's data from the US government, from NOAA and others, that's incorporated into our database.

Then another area that I've been working that the IMF is involved in is what's called the G20 Data Gaps Initiative. This is a project that's been going on coming out of the global financial crisis to identify data gaps and how we can work with the G20 economies to improve data in those areas. But we're now shifting towards the focus on a new data gaps initiative that's going to focus on climate change as well as other ESG issues. So things like financial inclusion and wellness, sustainability, all of those topics are going to be covered.

An important pillar of this is data access and data sharing, so that's going to be a big part of that initiative. The IMF has several data standard initiatives that have been done in the past, using things, encouraging the use of SDMX and so forth. All of that will be brought into that initiative.

Some of the outcomes here will be things like more timely information on greenhouse gas emissions, again, tying them to the economic activities that are

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

creating them, and then better measurement of physical and transition risks. So looking at how the financial sector specifically is going to be impacted by climate change and more information on expenditures by government on climate action.

I actually have in my notes, and you can see if you don't believe me, as statisticians like to say, if you don't measure it, you can't manage it.

Kevin Richards: [crosstalk 01:17:28]

Kristy Howell: So, Kevin stole my line, but absolutely. I think in the case of climate change, this could not be more true that policy makers are going to need the best information as they look to make good decisions about the future of our world.

Dean Ritz: Very good. Thank you to that. My mate, who does a lot of woodworking, she says you have to measure twice and cut once.

Kristy Howell: Yes, cut once. That's right.

Dean Ritz: I'm going to actually do a little quick follow on, because you used this phrase "data standards," and there's multiple meanings to the word "standards." I'm curious if the IMF is, in some of your, the role that a standard of a vocabulary plays versus the role of a syntax, which is used to express the vocabulary, or perhaps the syntax is used for collecting a data since you obviously acquire and aggregate data from many sources. I don't know if you work addresses some of those issues.

Kristy Howell: In terms of macroeconomic statistics, the IMF sets standards for how countries should report certain information. When we say standards in that sense, specifically my background is in, for example, the balance of payments area, and the IMF sets the standards for how things should be defined, what concepts should be covered, what's in scope when you're talking about the production boundary, as far as economic statistics.

But we also have what I mentioned, the data standard initiatives, and this is things like standardizing data and metadata and using SDMX and other common frameworks for reporting data. The IMF and other international organizations have worked together on several initiatives over the last decade or so to make more use of that type of standardized reporting for countries to report their data to those international organizations. Of course, also to disseminate it publicly so that other users can make use of that standardized information.

Dean Ritz: Part of your work is to facilitate that the data is, as data will say, data interoperability. The data has perhaps sufficient description agreement around its meaning, its structure, so that the analysis, the questions need to be answered, might actually be answered.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Kristy Howell: Exactly.

Dean Ritz: Kirsten, this brings the thought back to you since you're a chief data officer and you're responsible for data strategy at the agency. Can you tell us a little about how strategy takes an into account how other entities will consume the data? Because this is one thing, all agencies collect data. Data collection is a start of answering so many questions. So curious about how your strategy addresses, what the questions are, takes the questions into account as you develop your data strategy.

Kirsten: That's a big question, but maybe I'll talk through how we generate the strategy. I guess I'll speak first to, for a strategic plan, which is peppered with references of data throughout our strategic plan, but really specifically discusses new challenges and opportunities for FERC to leverage analytics to help identify and understand issues, assess implications of actions and value like commission effectiveness.

That phrase does come directly from our strategic plan, and then to be responsive to this within our data strategy, I work closely with our CIO. We built our data strategy into our information resources management strategic plan. That's a government term, but all agencies have to create those IRM strategic plans. So, we did create some goals for mission first and drive innovation.

Then we have key objectives under that where we're maturing data stewardship, which includes managing our information collections program and ensuring an effective data management program to include those internal data standardization and managing our data catalog. So, that gets to some of that sort of standardization piece that we were talking about.

Then other objectives under that drive innovation, we're looking to improve data driven-decision making, which includes implementing our data analytics capabilities against our roadmap and rolling out some data science training. The goal for our strategy is to ensure we're effectively bringing the data, the tools and the people together. We want to be able to reduce the time spent by data scientists and analysts finding data and tools, frankly, so that they can spend their time doing actual analysis to be able to answer those key analytic questions.

To my cake-baking analogy, if we have data scientists who are basically those master chefs, I want them to be master chefs. I don't want them to be spending 80% of their time being hunters and gatherers out on the field trying to find the data and then even using a camp stove, trying to figure out how to bring this together. I want to make sure that they have their ingredients, they have the right tools so that they can be answering those big questions.

Frankly, there's more of those big questions than we could even begin to count. That's where some of the important things like governance comes in. We want

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

to make sure that we're building models, and I can get a little bit geeky on things like Zachman's framework and the who, what, when, where, why, how, so that we're categorizing data and tagging data in ways where we can anticipate those big questions, because there's not one question we're trying to answer. There's numerous questions. But can we set up our data? Can we tag our data in ways that can try to anticipate some of those key questions?

To speak then to the standardization piece and our strategy, FERC does have a 115 information collections. We work collaboratively with our program offices to try to reduce the burden of those information collections, but also be innovative in our collection mechanisms. That's where things like our XBRL success has come from.

Then we have our data stewardship framework, and we recently to data stewardship campaign with goals to continue growing our data catalog and maturing data stewardship across the organization so that we have more shared awareness of our controlled vocabulary. We can document that and understand what we mean when we say certain things, because FERC has I think 13 program offices at this time.

Our office of public participation is a recent addition, but everyone speaks with a different vernacular and a different context, but making sure that we all collectively understand what we mean when we're saying certain things, particularly when we're dashboarding that, it's important. Those are conversations we have through our data governance body and why data stewardship is so important.

We need to know who our stewards are. We need to be able to bring that community together to have the effective conversation, so we can come to that at least internal standardization. When we talk about things like transforming our data, when we bring in our collections, some of our collections come in some type of standard format, like the XBRL. Other collections don't, just based on whatever scenario creates that data.

Sometimes we can't standardize at the collection point. So can we transform when we ingest it? Can we bring it through some type of ETL process? We can at least come to some type of consistency internally so we can be better prepared to answer those big questions. I think that's probably the best way to answer how we're looking at the data, how we're looking to standardize. That's the big picture view.

Dean Ritz: There's so many moving parts. We talked about, you have your data stories, you have your data vocabulary, which is written specification on paper. You have some of the data vocabularies, which are machine readable. You have the data coming in various structures or sometimes unstructured.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Kirsten: Lots of unstructured data too. That's almost a whole nother conversation, but yes.

Dean Ritz: The amount of screen scraping that goes on can probably light a whole country with the energy that's used. It's a technology community, so here we have an NGO, government agency. We have a community organizer over there and Richard at SAP.

Richard, could you give some advice on how these entities, and if you were to add business to the mix, can collaborate and start addressing, solving some of what is obviously a common problem. Earlier I talked about or noted that this kind of information is wanted obviously by government agencies that have role related responsibilities to do things like, for example, assure fair markets and energy transmission and energy costs.

You have NGOs that are interested in the impact of activities of these entities. You have investors who want to invest and make capital allocation decisions based on the promises of either the economic activity and in the confidence that the activities are being done in a manner that's not so destructive.

Every business isn't a legal business. We try not to have slavery, for example, because it's a way to make money, but it's something we do not permit. That's a crazy aside, so we'll stop right there. But Richard, as the community organizer, what advice could you give for the collaboration of the three types of entities?

Kevin Richards: Sure, Dean. I guess in terms of SAP, who like at Data is playing a real key role in the fight against climate change. We're helping to accurately measure the carbon footprint of agencies operations and assess the carbon footprint of products for instance. Now, there's really a deluge of climate data, and scientists are really scrambling to find ways to store, analyze and preserve unprecedented amounts of information about the effects of rising global temperatures.

With new data collection techniques and technologies continuing to advance, there's a real focus on reliable and readily available data. It's critically important to governments working to set policy and monitor compliance with international climate packs as well as to help local authorities and trying to help their communities adapt to unusual weather patterns and rising seas.

An interesting data point I'd like to relate to the group is that with the ever-rising numbers of sensors that are in at it in satellites and aircraft and ocean buoys and the like, there's more data all the time. Over the next decade, officials managing the main US repositories of climate related information expect their archive's total volume to grow from 83 petabytes today to more than 650 petabytes over the next 10 years.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Of course, one petabyte of digital memory can hold thousands of feature length movies, and about 650 petabytes can hold contents of the Library of Congress 30 times over. You can think about government and the private sector and the ability to leverage that data to address the challenges around climate. When you think about SAP, we really support government's climate efforts by embedding sustainability into its analytics and applications and provide customers with sources of more sustainable products and also provide visibility into waste flows to help agencies towards zero waste.

There's examples of governments around the world leveraging climate data in different ways. Singapore, for example, is monitoring emissions from diesel vehicles in real times. Hong Kong is using sensors and data to help aging government buildings like their Cultural Center become more energy efficient. Canada has turned to AI to produce chemical substances that accelerate the production of clean energy. It's also using AI and data to prepare the country's infrastructure for electric vehicles.

We've worked also with the government of Heidelberg, City Council of Heidelberg in Germany to streamline waste management. If a container exceeds a certain capacity, a request automatically goes out to create trucks to empty that waste disposal.

I think when it comes to this issue, the ability to have a public-private partnership and work together with governments to realize the true promise of data is very important. I might note, Dean, that in terms of the US government, I'd like to commend the administration and President Biden for the executive order on tackling the climate crisis at home and abroad, which really takes a whole of government approach and focuses on an initiative to deliver accessible and actionable information to individuals and communities that are being hit by floods, droughts, wildfires, coastal erosion, and other communities that are faced with intensifying climate impacts.

The focus that the administration has put on making climate data available to the public is important. It's something that SAP supports and everything from NOAA using AI to improve accessibility and accurate timely climate information to these different task forces that have been put together to expand and prove climate information and services to the public and federal mapping services and advancing the nation's geospatial capabilities to promote climate resilience are all very important factors to help the private sector work together with the government and to address these issues.

Dean Ritz:

Great. Thank you very much. I want to pick up on one of your points for you, Kristy. You mentioned this idea of, well there's an immense amount of data that already exists and it's growing. In the sessions earlier today, there was a lot of recognition that machine learning and artificial intelligence is the primary mechanism we're going to apply in order to scale, in order to understand the data.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Tens of millions of documents, you can't have everybody read them and figure out how to tag them. You have to find some way of automating it. You also talked about the Biden administration asking that this data be made public. So here we are, we have data from many sources, many formats. It's becoming a part of our commonwealth. It's a common asset for us.

You're a statistician and you're looking at all these different data sources, and I wonder if you have any anecdotes about things that were surprising, the connections that you might have made across these disparate data sets from your work at the IMF. Sorry to put you on the spot. It's a new question on the list.

Kristy Howell: I don't know if I have the best example for something that was surprising to me, but I do think that there are a lot of examples where these data mining, web scraping, other techniques are being used for maybe topics that people wouldn't necessarily think of.

One of the issues around climate change that is going to be part of the focus of this new initiative that I was talking about is looking at sustainable finance. You've got examples of central banks that are already doing a lot in this. Central banks, you don't necessarily think about them as being in the statistical world, but some of the macroeconomic statistics that are produced are actually coming from the statistics and research offices of central banks across the world.

They're having to use some of these information to pull information out of the company records, exactly what we've been talking about with whatever information filings companies these are putting out, whatever they have to say about climate in them. They're using web scrapping to pull that out and start to pull together their own information on what kind of impacts climate is going to have on these companies.

If you think about not only the immediate hazard impacts of climate, what companies are potentially exposed to floods and droughts and things like that, based on perhaps where they're located, but also what's going to happen to fossil fuel companies when their assets are stranded.

That's going down that perspective of what happens as consumers' interest shift, and maybe that changes a company's business model, and what are going to be the forward looking impacts on them. This is something that these central banks, as I mentioned, are already doing a lot of work on.

There's an example from the Bank of Spain that launched a project to use webs scraping for that, but also then looking at forward looking indicators. That's also something that we're covering in this initiative is looking at different climate scenarios that have been laid out by the Network for Greening the Financial System, based on these different scenarios that might play out as we hopefully get to net zero.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

What impact is that going to have on the individual financial firm? What impact is that going to have on the financial sector as a whole and the entire global financial system? I think maybe that didn't quite answer your question, but hopefully gave you an example, something unexpected that this data, that these tools could be used for.

Dean Ritz: What it I think it shows is that it takes domain expertise to and a lot of work in order to recognize new types of value out of the combinations of data. You have the luxury of being able to find whatever data might be of whatever use. That's something that the technology is facilitating. We're not just all trading CSV files forever, which is a nice thing and PDF files.

Kirsten, I wanted to ask if on your work and the technology, and well, not necessarily from your strategic plan, but are there certain technologies or methods that give you optimism? You go, "Oh, This is pretty cool. This can actually help us either scale the amount of data, deal with the interoperability of the data, increase the trust or accessibility of the data." If there's a...

Kirsten: Sure. I guess where we are in our life cycle. I started at FERC 2019, July of 2019, so it's been about two and a half years. A big focus that we've been working on is moving to the cloud and building a data analytics platform in the cloud. We obviously want to be able to benefit from the scalable compute capacity and improve our customer experience and really get to more of that self-service analytic capability.

We're really getting to a point where everyone is a data analyst to some degree, especially if you look at the younger generations coming out of college. Everyone seems to have more of that analytical capability. I even look at, I have triplet seven year olds, even looking at my seven year olds and how they understand computers and technology and that data, it blows my mind. Really, everyone is some type of data analyst.

I and my team won't have the capacity to be responsive to every analytic question asked. I need to ensure we've established the right methodology and that we're providing the right tools to make access to the data and the tools more streamlined and we automate where possible. We're implementing analytics governance processes to ensure that the of products created across the commission by analysts are as consistent and trustworthy as possible.

That's where looking at data governance and stewardship tools is an important part of that strategy, where we maintain a roadmap at FERC for our cloud platform. We have capabilities identified across the value chain and where we can automate it in the data governance space, I think is very important, because that will help us improve some of that standardization, some of the communication.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

It should no longer be word of mouth. It shouldn't be someone who knows the secret handshake or the particular person to contact to find the data to get the information. We want to have all of that centralized in a place where people can find it easily. They can quickly... They can have that single pane of glass. They can do their quick search to find out what data do I have related to this particular domain or this particular topic, spin it into a workspace, do the analysis they need.

If it creates value, if it's determined that this is going to be usable, what does that analytics governance process look like? We can then deploy it into a production space and have trust that the analysis, the transformations, have been done correctly. We all agree to what that means.

Then a specific recent success that we've had at FERC is in standing up our cloud platform. One of our first big efforts was we completed a reporting modernization project for all of our internal reporting and dashboards. This project was pretty successful, because we were able to integrate our governance process and the internal data standards process for consistency across the reports.

We basically had about 530 legacy PDF-style reports. We were able to consolidate that down to about 50 different dashboards, because there was so much overlap of data elements across the 530 different PDF reports, apply a consistent definition. We could model at data consistently, and now we're down to those 50 dashboards that are available to the appropriate program offices.

Some of those are made available through enterprise capabilities. Then we can also certify some of those data sets and dashboards and enable the self-service piece. So, that's where we are right now. We're continuing to grow that and looking at ways to scale that and continue bringing in more data, because there's so much, so much more that we can do. But at least, that was a huge success even just to get out of our legacy reporting environment and those 530 PDF-style reports.

Dean Ritz: So you've been busy?

Kirsten: Yeah.

Dean Ritz: Very good.

Kirsten: Plus the triplets.

Dean Ritz: That was training ground to then become the first CDO over at FERC, no doubt. Is there a gap that you would, since we have a little brain trust here and there are people in the internet watching this, is there a particular gap or problem you would like someone to solve that would assist you in your successful reign?

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

- Kirsten: Just speaking broadly across the data space, just to ensuring that we can continue to get access to clean, well-defined, well-documented stewarding governed data. That's probably my simple ground truth. It's a lot harder than it sounds probably, and as I mentioned earlier, we are very much at this learning phase. A lot of it's also just evangelizing the importance of why this matters, why we need to be talking about this, why having that clean, high-quality, well-defined, well-documented sewer data is so important.
- Dean Ritz: Very good. I think you get a lot of agreement on that. It's the holy grail, or in some sense, without having the data that we can trust without eliminating ambiguity, without knowing where it came from, having the governance behind it and the machinery [inaudible 01:40:00] way, it's an impediment to making major decisions on it.
- Kirsten: I don't want anyone to look at a product, a data product, and say, "That doesn't feel right to me."
- Dean Ritz: Look at the resulting data says, "The gut test doesn't pass here."
- Kirsten: Right. But if you can point to the governance process, the documentation say, but this what the data is showing us, that's how we get to that true data-driven decision making.
- Dean Ritz: Very good. Richard, I'm curious, what have you heard, because you speak across many agencies, might be a common theme of improvement with a gap they'd like to close?
- Kevin Richards: It's a good question, Dean. I think I talked a little bit about the volume of information, which is really vast, but I think one of the challenges in a lot of cases is the information is fundamentally unstructured. So having that data more structured and I think machine readable, and we talked a little bit about standards. I think it's important to have internationally recognized standards and ISO standards. But to me, that unstructured data is a key challenge that I think all agencies and departments really face.
- Dean Ritz: Very good. Can you believe I'm going to have the same question for you? What's the gap that you'd like to fill?
- Kristy Howell: Well, I think that a lot of the points have already been made about the need for standardization and harmonization. This is as I was describing earlier, something that we work on in the IMF in terms of macroeconomic statistics. But just more generally, there's this recognition that our classification systems that we have, and as a statistician, we have many.
- We have some for industries, economic activities, you name it, but maybe they need to be reevaluated to look at whether or not, again, looking back at this topic of climate, are they meeting our needs for that purpose? An example that

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

we have that we're already aware of is in a lot of the classifications, there's an energy industry, but it's not separately identifying renewables.

That's a simple change that needs to be made in order to make the data more usable for looking at climate issues. But those types of changes to classification systems are not always something that happens very quickly. I would say that's one area that wasn't already covered that we could think about as in terms of making some improvements.

Dean Ritz: Very good. I'm not surprised in this result. I'm sure hardly any of you are going to be surprised. The challenge is figuring out the classifications and then figuring out how to express those classifications in a machine readable way, so that they can be applied to data either algorithmically through machine learning, on PDF files or in more explicit mechanisms that we have.

I want to thank our panelists for being here in person and for answering all our questions here. Let's give a hand for Kirsten Dalboe and Kristy Howell and Kevin Richards. All right.

Speaker 2: Well, I'll say also thank you to our moderator, Dean.

Dean Ritz: [inaudible 01:43:10]

Speaker 2: Thank you all. This has been fantastic.

Dean Ritz: Very [inaudible 01:43:12]

Speaker 2: Well, we are going to turn to the next part of our program, and I would like to welcome our afternoon keynote speaker. Hugh Halpern is the director of the government publishing office. He's actually a great leader in the topic of many of the themes that we've been talking about this afternoon.

He also carries with him a great bio and history that I'm just going to highlight a few things before I cede the floor to him. In his background, he's been characterized as one of the 50 most powerful congressional staffers in his history, a new power player in his history and is a long time congressional staffer prior to his appointment as the head of the GPO.

More recently, he's been leading efforts to really modernize and improve some of the customer experiences at an agency that's important to many people in the American public for reasons that he's going to talk more about. With all of that said, I'll say, Hugh, thank you for joining us today, and I will turn the floor to you.

Hugh Halpern: Thanks so much. You good? Thank you, guys. Thanks so much. I'm really happy to be here today and thanks to Nick and the Data Coalition for having me.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Actually, if we could flip over to my slides.

Speaker 3: [inaudible 01:44:43]

Hugh Halpern: While we're doing that, let me tell you a little bit about me. I spent a little bit more than 30 years on Capitol Hill, started as my local congressman's driver, and when I retired-

PART 3 OF 6 ENDS [01:45:04]

Hugh Halpern: ... started as my local congressman's driver, and when I retired in January 2019, I was the head of floor operations for speaker Paul Ryan. And a lot of my time on the Hill, probably more than half of my career, I was in positions where I was responsible for document production. And during that time and with that responsibility, I got to know GPO really well as a customer, and frankly, there were a lot of times, I wasn't all that happy with what I was getting out of GPO or what they were demanding of me to turn over work product that I thought should be pretty easy. So I retired and I was trying to figure out what I was going to do when I grow up and some folks came to me and said, "We're really glad you're free because GPO needs a new head." And lo and behold, probably eight, nine months later, I found myself as the 28th director of the Government Publishing Office.

For those of you who are not familiar with what GPO does, GPO is a captive service provider to the entire federal government. We operate as a business, so I have profit and loss. I look at a financial statement every single month, and hopefully those numbers are in the black. But we produce and provide print and publishing services to all three branches of government. We manufacture everything from the federal register, congressional record, congressional bills and a host of other things to the US passport. And we're roughly just shy of about a \$1 billion in revenue each year.

So what I'd like to talk to you today about is what that process looks like, what GPO's role with Congress in producing legislative documents looks like, and hopefully how we're going to make that better here in the near future. All right. So as we walk into this discussion, one of the key things to remember is that one of the decisions that Congress made about 200 and some odd years ago was that they were really going to be focused on how these documents looked in print. So that is the key metaphor you've got to keep in mind, and that decision in the very early days of this republic really fashioned everything that came after. Because as you all are familiar, when you have a product, you build systems around that product and then when something changes, it's very hard to change. So now to give you an idea, that is a concurrent resolution from 1868. It was produced in handset type at 732 North Capitol Street. The same site the GPO has been on since its opening days in the civil war. And that particular image is available online through the congressional serial set, which is hosted on gov info, which is GPO's trusted digital repository.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

You can compare that with a concurrent resolution from a year and change ago. It looks a little bit different, it's probably a little bit cleaner. This is set digitally. Has line numbers. But the key thing to remember is that we have chosen to express law in terms of typography. And when we decided to do that, that meant that all of the other decisions that go along with that, how do we change law, how do we publish it, all of that keys off of the typography. Now, for the first 175 years of the Republic, that's all well and good because it doesn't really matter... data wasn't the thing. But as you got into the 1980s and 1990s and the information underlying that typography became important, that's when we really started discovering problems.

So in the 1980s, 1990s, GPO developed an electronic typesetting system. We actually started using electronic typesetting in the '60s and '70s, but that was all mainframe-based and not something that we could replicate elsewhere. But in the 1980s, we developed a PC-based product and developed this system. So what you're looking at is a text editor, DOS-based text editor called XyWrite, which has long been swept into the dust bin of history, and what you see on that screen is GPO's proprietary typesetting language. We call that locator code; the slang term for it is bell codes. So those bullets that run down the left-hand side of the screen, back in the day, those were all bell characters, and I can really date some people in here probably, but if you remember control G on the old terminals, like an Osborne or something like, there's a little bell character and that's where that came from. So everything in GPO's locator format is keyed to a three...

You look down there at the bottom, there's a bell eye 21, and that is the character that was used to begin paragraphs in our typesetting system. Still used today. Locator codes are still used today, and I'll explain both as an intermediate step and sort of in the raw form. So we used XyWrite for a really long time. I used it myself when I was on the Hill. It is really ill behaved as once you get past about Windows 95, so we upgraded. We moved to Text Pad, and that gave us such huge features as syntax highlighting and locator code hints on the side, but other than that, it was really basically the same thing that we had been doing for a really long time.

Now in the late 1990s, early 2000s, Congress really took a leap ahead. And one thing I should have mentioned, in that late 1980s period, there were two offices on the Hill that decided that they were going to work with GPO to improve both the speed and accuracy of drafting of congressional documents, and that's the House and the Senate offices of legislative counsel. Those are the attorneys who actually draft the legislative language, and again, they are working in typography. The thing to remember is that if you're amending something in the US code, you're going to say, strike the third section of Section 1, 2, 3, 4. And if something is supposed to be in bold, say a section heading, that has to be represented in bold. Similarly, if you're just striking or replacing words, you have to show where that is. All of the instructions are to an unseen clerk.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

So what we had was those offices were using GPO's systems because we found that it was actually just easier for them to learn this system, draft in this system and then produce the output that we could then consume on the other end. Now, in the late 1990s, early 2000s, we made sort of a sea change. For drafting legislation we shifted over to an XML-based editor. And so that was the first time where Congress really started describing things as what they were rather than how they looked. So it was no longer this is all caps, bold. It's now that's a section heading. This is a product called XMetaL. It is the base XML editor used by both the House and the Senate. This is the House version. The Senate actual uses version later than this one, but this is the instance I have running on my desk, and I turned tags on just so you could see that we actually do have a format that uses tags.

So this XML formulation has grown into what we call USLM, United States Legislative Markup. And that platform is something that we are working very closely, GPO, the house, the Senate, the office of the law revision council, the people who are the keepers of the US code, we are all working together to start building out USLM so that we can account for all of the legislative formats, for all of the formats for the ancillary documents, so committee reports, the congressional record, and then even into the regulatory space as we try and incorporate things with the federal register and things like that. But this is basically the application that's used today to edit.

Now, what happens when you go to print? That's where things get fun. So you're an attorney and office of leg council. You've just finished writing the bill that will change the world, and you have to get it to print. So you hit Control-P. Well, what happens? Well, we take that good USLM structured data, and there's a little script that runs, and it changes that file back into GPO's locator code. So we're translating backwards. Great. Then there's a little piece of software called MicroComp, and that's GPO's composition engine, which has been in existence since about 1983 and we continue to use it today. We are holding it together with Band-aids and bailing wire, and it is sometimes a little finicky, but you run it through MicroComp and out the other side comes a raw PostScript file. Then Adobe Distiller pulls that raw PostScript file and generates a press-ready PDF at the end. And if everything works well, you're golden. But if it doesn't, you're kind of screwed because it has really bad error handling. It has all of the kinds of problems that you could imagine. If you've got that number of dependencies combined with a product that is that old.

I have had to tell Speakers of the House on two occasions that we could not move forward with whatever their particular priority was because we could not get it to print. Those were unpleasant conversations. So we need to replace this. We need something new. And the problem is that fixing systems this old and this complex is actually really, really hard. So we've had to continue using MicroComp as GPO owns this print process throughout the last several years. And I can tell you again, as a customer, somebody who was on the Hill, who every time Microsoft would send out patches to Windows, well at that 7, you

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

would hold your breath and hope that MicroComp would still work. But I would always ask, I'm like, "When are we replacing this? When are we doing the something more modern?" It didn't matter when I asked, it was always two years away. Always two years away. Well, the good news is we're no longer two years away. We're much sooner than that. Keeping my fingers crossed.

XPub is the future of composition with GPO and XPub is our next generation composition engine, but it is really sort of the hub of GPO's operations for the foreseeable future. XPub began life as what we call the CSR, the composition system replacement. Thankfully, we did better marketing than that. So we came up with XPub and XPub is based on the XPP commercial XML publishing system. It's widely used in a variety of commercial print settings today, and is really robust and is a huge leap forward. And eventually this will sort of become the hub of our operations at GPO. It will take over that print function for our congressional customers, and it will really facilitate the development of a whole slew of new applications for all of our federal customers.

So how does printing work after XPub? You start with good XML file, it goes into XPub and you get a PDF on the other side, press ready, ready to go. It's pretty simple. There's a lot going on behind the scenes there, but that's really how it functions. There are not a whole lot more dependencies other than that. So it's that ability to take that XML and translate that, image it and produce that PDF on the other side, that is really key to improving what we've got going on these days.

So you're going to say, well, what's that mean for me? That's what everybody says. How many people here, and I realize I've got a thinner audience than I might have virtually, but how many people here have ever tried to copy and paste text out of GPO and Congress.gov's text display and then started swearing a lot? Okay. I want to apologize on behalf of Congress, behalf of GPO and everybody else, because you know why? That text display is terrible, and I usually don't use words as family friendly as terrible. There's weird punctuation in that display. There is stuff that's in all caps that is actually upper and lower case, but small caps. Each line has a carriage return at the end. So if you want an actual paragraph that you can use and manipulate, you got to go through and strip all that stuff out. It is terrible.

Okay. Here's my gift to you. When we are able to deploy XPub with our congressional customers, you will get a responsive HTML text display. It will look good on your tablet. It will look good on your phone. It will look good on your web browser and you can copy and paste like any other text document. You can put it into a Word document. You can put it into a Google doc. You can collaborate with it. You can track changes. You can do whatever you want, but it's really easy after XPub. And that is one of the great features we're going to be able to deliver for folks.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

On the authoring side, there are even bigger benefits. So let me talk a little bit about what our current workflow looks like. I'm going to use my favorite document, the document I have been responsible for probably more than anything else is the lowly committee report in the House. So I'm probably personally responsible for someplace in the neighborhood of 900 to 1000 committee reports over all of my different committee engagements. So there is usually a portion that is legislative in nature that is prepared by the office of legislative council and is produced in either Locator or XML, so we get a PDF. So you get that paper copy and PDF. Then the committee is going to author the guts of that report. We're generally going to do that in Word. [inaudible 02:03:47] of Microsoft shop, we're going to do that Word. Then we're going to take that in a bunch of other pieces and we're going to take those printouts, put a rubber band around them, send them over to GPO and cross our fingers and hope for the best.

What happens in GPO is those PDFs that were prepared by the office of leg council, we can pull those files directly from them. So we've got that XML file coming over us, but that Word file that the committee prepared, we don't really have a good way to ingest that. So what happens is we still will get that electronic file usually either via FTP or on physical media, thumb drive or a CD. We'll take that file, we'll strip out all the formatting, put it back to a text file and then, and I am not making this up, there is a person, they markup specialist in our proof room, and they will take the printout of that text file and they have a gigantic tray of rubber stamps. And those rubber stamps all correspond to those locator codes I showed you earlier. And they will go through and in the appropriate places, they will use those rubber stamps to mark the formatting locator codes where they go. Then another person will take that piece of paper, pull up the text file and actually key in those locator codes. It actually works pretty well. It requires a lot of people and it is really expensive, but it works most of the time. But it's also really inefficient.

So one of the great things that XPub will do is it will accept input from a variety of different sources. Let's see. Looks like my remote is not... oh, there we go. No, there we go. So we can pull in that good XML file or, if you prepare a document using a template that we can provide you where you're using Word styles, we can translate those styles into the XML required to compose that properly, and that is a huge step forward. And frankly, as I'm looking at the future as the head of an organization where we're looking for growth opportunities, building solutions based around the tools that folks are using now, so around Word and enabling you to get good copy out is actually a huge step forward.

So then out of XPub it can generate that PDF. If it's bringing in stuff in Word, it can generate that good USLM file or it can produce a responsive HTML text display. That is a huge step forward from where we are, and something that I think over time will kind of revolutionize that way the Congress produces its documents. And frankly, that's something we would like to replicate with other

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

agencies, whether it's OMB, we've got the President's budget coming out here at some point, and that's, that's a document that we produced using our existing typesetting tools, so MicroComp and the rest. But there are lots of other opportunities as well.

So we also want to make this a lighter weight experience for our customers. So right now, if you're running XMetaL and MicroComp on your desk, that requires two different sets of people to support that. So somebody from the House, I'm a House guy, so somebody from the House to support the XMetaL installation, and then somebody from GPO to support the MicroComp installation, and that's all running locally on your machine. With this, we would like to be in the position where we can take that Word file, have you upload that to a GPO website and it spins around and then can generate all those files for you. Good XML, good HTML, and a press ready PDF out the other side, ready to go. So you can use the tools you've already got installed on your desk, and you don't need anything special.

If you're interested in XPub, if you're interested in looking at this stuff, there's a lot of information on GPO's GitHub repository, so please take a look there and then also has samples of the responsive HTML display. But there's also actually a lot of other information there related to legislative transparency in using legislative docs.

So with that, I'd like to thank you all for taking this little trip with me down the down the rabbit hole to learn about how Congress prints. And if we've got some time for questions, I'd love to take them.

- Nick Hart: Yeah, we have just a couple of minutes for maybe two questions. So if there are folks in the room that might have a couple of questions, and while you're raising your hands, we have a mic runner here. I'll just say you're very generous in not using the same colorful language in describing some of the copy paste problems that I experienced when I was at OMB. My choice of terms was much more colorful in the errors that I experienced. So I'm happy to-
- Hugh Halpern: I'm trying to keep your family friendly designation.
- Nick Hart: All right. Are there any questions for you? Thought I saw a hand over here? Go ahead.
- Tobias Schroede...: Seeing that no one else has questions, I'll just make a quick comment. I'm Tobias Schroeder. I'm the director for the eRulemaking program. First I want to commend you on your staff, Lisa LaPlant and Matt [Landgraffer 02:10:28]. Very collaborative, very knowledgeable, and I've enjoyed working with them. But the whole premise of everything that you've talked about is what happens when you hit Control-P, but I also want to emphasize how supportive I am of this whole effort. I think it's fantastic. But in its USLM format, before it ever gets published for somebody to read, computers can now be used to inter-operate

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

with all those data for a lot of value added and complimentary benefits. So I think it's really fantastic that you're evolving in this way.

Hugh Halpern: Absolutely. And you're spot on with our team. Lisa and all of the folks who have sort of seen the development of USLM have been a huge asset, not only to GPO, but I think to those of us who believe that information, once it hits that public sphere, should be easily accessible. And you're absolutely right. I think USLM has tremendous benefits, and one of the things that I'm a big believer in is that there are a lot of other people who can think of novel applications way better than I can. So a good example of how USLM can operate is actually a current issue, certainly in the House. There's discussion... right now, if you wanted to look up how a member voted on the floor, it's pretty easy. The clerk maintains a database and it's easily searchable, but not so much for committee votes. If you're trying to look that information up in a committee report, half the time what you're going to see is an image of a tally sheet. And while that means that's pretty accurate data, that's the exact same tally sheet that the clerk used when they were in the room and they were marking off that tally, it's not terribly usable.

But using some of the data elements in USLM, if you were to build that into a database, you can use the bio guide indicator so that you can identify that individual member of Congress. So you can see how he or she voted across their entire career, across whatever committees they might be on. You can identify that person, and sometimes that's hard when they're Mr. Smith of Iowa or something like that. And there have been a lot of Mr. Smiths over the years. But things like that, applications like that are really where we see some potential over the next several years. So no, that's a great point and one of the huge benefits of USLM.

Nick Hart: All right. One final short question.

Alex: All right. So one of the challenges I've seen is when there's citations in the law back to legislation from the 1950s and otherwise, it can be difficult to sort of automatically parse back to that. And so I'm wondering what sort of efforts you're making in handling the PDF problem that's been so much discussed, digitizing PDFs and putting them into a more structured format?

Hugh Halpern: So great question. This actually goes back to my original premise, which is a couple hundred years ago, we made a lot of bad decisions when it came to how we organize our laws in Congress. And in fact, it has less to do with making one bad decision, but then making a whole series of decisions that made sense at the time, but didn't really contribute to having a unified system of laws and regulations in this country. Other countries do it a lot better, a lot states do it a lot better, but we are encumbered with these systems we built.

So the short answer is we've done a fairly decent job of migrating most of the law into USLM. So there's the US code, of course, and that is entirely produced

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

in USLM now. Actually, that was the first product we produced with XPub, and we were able to shave, I want to say seven months off the production time, for the last full printing of the US code. But again, unlike a lot of other countries, unlike a lot of the states, the United States has a number of different codes. We've got the US code, we've got the internal revenue code, and then you have a whole bunch of standalone laws. So one of the projects that we've been assisting the House and Senate offices of leg counsel with is moving those compilations of laws, those compilations of those individual laws into USLM, and I think that's going to give you a much better base to work from. And so some of those citations to those individual one-off things are going to be a lot easier in the future.

That said there are still going to be things, appropriations law theoretically disappears I after a year, but there are sometimes permanent provisions in there. The good news is those bills are produced using USLM today, but citing back to them may be a little bit more difficult. But these are all problems we're thinking about and trying to resolve. It's less a question of actually going back and digitizing. Frankly, in the case of some of the standalone things, it was establishing ownership, making sure that the United States was the owner of that information. So that's something where we continue to work on.

Nick Hart: Well, this is a great place to leave us. Let me thank you for your leadership in this space, but also to your entire team for all the work that is happening here. This is a very exciting set of progress in momentum for the activities that are underway. So thank you for joining us today.

Hugh Halpern: Thank you.

Nick Hart: And we look forward to this all happening in, I think you said less than two years, if I caught that right.

Hugh Halpern: So finger crossed, we hope to be deploying XPub for bills, resolutions, and amendments, hopefully by the end of this calendar year. I believe that we will be ready. It will be dependent on our customer, when they want to deploy that, but we are hopeful that will, within this next two-year time span you will see the fruits of that result.

Nick Hart: All right. Well, please join me in thanking Hugh Halpern.

And this is also a great segue to our next panel discussion. I'd like to invite Jason Fichtner, who's vice president and senior economist at the Bipartisan Policy Center, and our other in-person speakers and virtual speakers who will be joining us for the final panel discussion of today on improving citizen engagement and transparency in rulemaking and administrative processes using reg tech solutions. The discussion you just heard is a very natural segue to talk a little bit less about the legislative items and more about the executive. So Jason, I will turn it over to you to introduce us to the panel. Thanks for joining.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Jason Fichtner: Thank you. Appreciate it. Let me get... Quick look, make sure we're good to go. Perfect. Everyone's there. So hello, everyone. Thank you for sticking around today. I appreciate it. My name is again, Jason Fichtner. I'm vice president and chief economist at the Bipartisan Policy Center, where Nick mentioned he's still currently a fellow. He used to be employed there, and as soon as I showed up, he left. Now, I realize correlation's not causation, but you have to wonder sometimes, so it kind of hurts.

I also have a lot of experience in government. I spent almost a decade at the United States Congress working for the joint economic committee, where I was trying to cut and paste documents as well off my computer screen with very little luck, and I did use colorful language. And I also had several positions at the Social Security Administration, including chief economist and the principal deputy commissioner. So I also did a lot of work with rulemaking and transparency and how to help citizens engage with government and what we're getting at and accumulating all the discussion today is evidence-based policy and how we use rulemaking and how we use data to inform people. And this panel's improving citizen engagement and transparency in rulemaking using reg tech solutions.

And it's the citizen engagement part that's also very important in this because unless you have citizen engagement, unless you have citizens who understand how government works, unless they have a voice in how their government works, you don't have legitimacy with policy. And we've seen way too often what happens when you no longer have legitimacy in your government. People don't follow the rules. They don't believe that there's a fair shake in the rules. So part of this conversation we're going to have today is how do we make sure that we're using technology properly to have citizen engagement? And we have a wonderful panel today of four people.

Before I turn it over, just a quick sort of housekeeping, especially for our panelists so they know. You can see us. We have a confidence monitor, so can see you. So we're not going to turn our head around our-

PART 4 OF 6 ENDS [02:20:04]

Jason Fichtner: We have a confidence monitor so we can see you. So we're not going to turn our head around our shoulders, but if you have questions, you can do the thumbs up, you can raise your virtual hand, we'll see it, which is good, and then we can stop and make sure you have time to ask questions or make comments. Also keep in mind that there's about a second delay, not in the video, but in the audio as microphones cut in and off. So might turn over to you, take one second before you start speaking, let the mics change over and we'll go from there. So with that, again, this panel is going to discuss how RegTech solutions and approaches can streamline compliance as well as offer insight into implementation and impact with federal rules and support transparent

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

regulation decision making, that again very importantly, encourages the American public to meaningfully engage with their government.

I am going to ask each of our panelists to come up or virtually sit down and virtually come up and to sort of give an introduction about who they are, what they are. So we'll start with the question and then we'll go to a moderate discussion. I know everyone come up. Virginia I'll have you start first, but sort of an intro question to start is, could you begin by telling us a little bit about yourself, introduce yourself, your work and your perspective on sort of these broad issues dealing with how we engage citizens using RegTech. Oh, you okay?

Virginia Huth:

All right. So little bit about myself and where I started from, it's actually been full circle. I started out as a regulatory analyst at OIRA, Office of Information and Regulatory Affairs back at OMB in the 1990s. And then I got a little into technology, worked for IBM for about five years as a consultant. And then I took on leadership positions in IT and contracting and in budgeting in various agencies in the federal government. And then in 2011, I moved to GSA and I moved back to regulations. I took over the GSA acquisition regulations portfolio, and then two years ago, I took over the eRulemaking Program Management Office and another program when it moved over from EPA to GSA.

And I'll add that the program just moved to the technology transformation service at GSA, so we've seen a lot of change in the last couple years. Right now we're excited because we have the opportunity to modernize. We just got some money in the President's Budget. And I think with TTS around us, we're really excited about the opportunities we have there, not just to move off an antiquated system, again, it's 20 years old monolithic architecture, but also to implement some exciting data analytic strategies to really support RegTech, to really bring us up to where we need to be in the modern world. And we'll be looking at... We have a proof of concept I'm going to talk to you more about that brought a sort of a trifecta of machine, readable language, no code software and natural language processing to bear on a couple solutions.

Before I go into that, I do want to give credit, Tobias Schweiger, in the audience here. He is really the heart and soul of erulemaking. He had the vision for this proof of concept. He had the tenacity to make it happen when we got all of our funding zeroed out and was able to partner with the centers for Medicaid and Medicare services. So really a lot of credit, much of the credit, almost all the credit goes to Tobias for bringing us up to where we are today, so thank you Tobias.

So stepping back for a sec, and I know this is a really a very tech-savvy audience, so I hope I don't offend anyone, but one of the things we've discovered is that PolicyTech is probably more nuanced and more contextual specific than a lot of other text that is non machine readable. So it makes it really difficult for policy makers to really wade through a lot of challenging information and data, and especially for typical approaches to natural language processing, to be able to

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

manage through that data. So again, at best natural language processing allows you to look for patterns, but it doesn't provide the context. Ultimately, we understand that machines don't read the way humans read. We provide context in a natural way. So again, sort of a silly example and probably not needed for this audience, but I enjoyed putting it together. What if we were to do a search on the effective climate change on sea life, and we got a hit for grilled salmon on a restaurant menu. Totally irrelevant, but the word patterns are there.

So GSA approach again, has been to focus on three things, that natural language processing, machine learning and no code software, to build a technological approach that does enable machine readable tasks that gives us the ability to provide context. And this way our machine knows the difference between overheated Octopi on the Atlantic sea bed and grilled Atlantic octopus on a bed of greens. Okay. It's just a silly example, but I thought it might be entertaining.

So that being said, we just finished a proof of concept with CMS to explore a couple very specific use cases. How to enable improved public engagement in the regulatory process, and also how to enable the rule makers to more efficiently effective process those comments. We examined how those regulators could identify topics that are out of date, as well as areas of regulations that are contradictory or duplicative. And I would say there are many other examples for how this technology could be used in terms of assessing DEIA impacts, assessing cost benefits of rules, both retrospectively and proactively, and ultimately towards the benefits of speeding up the regulatory process, reducing unintended consequences and promoting social equity and all of this directly supports the January 21st, 2021 Presidential letter on modernizing the regulatory review process.

I'll just conclude with, as we observe the rise of tyranny in some other countries, I am more than ever proud of the way that the US government approach to the regulatory process supports transparency, accountability, integrity. I truly believe that the provisions in the Administrative Procedure Act in 1946 are fundamental to our democratic values of our democracy. And I'm also proud of the way our program at GSA is looking to enhance those aspects of accountability and transparency to really support better regulatory outcomes. Thank you.

Jason Fichtner: Thanks, Virginia.

Susan, I'm going to turn to you next, please. Would you introduce yourself and sort of the same thing and tell us who you are a little bit about your work and then some perspectives on these broad issues like Virginia commented on.

Susan Dudley: Yeah, sure. Thank you, Jason, and thank you also to Nick for inviting me. So I am the director of the George Washington University Regulatory Study Center, and I'm also a professor in the Triburg School of Public Policy and Public Administration, where I served on Nick Harts dissertation committee, which is

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

one of my... I'll take some credit for Nick. I also teach a regulatory comment clinic there, teaching students how to engage in rule making. I served as administrator of the Office Of Information and Regulatory Affairs, which Virginia mentioned back in 2007 to 2009, which was really in the early stages of making information available online and erulemaking and so much has happened since then. So now I run this regulatory study center, which focuses on improving regulatory policy and practice. So we, my colleagues and I evaluate regulations from different angles, the process by which they're developed, public engagement, regulatory impact analysis, and the merits of individual regulations. We'll actually file comments using erulemaking, using regulations.gov to write comments from a public interest perspective.

And we apply different tools, the faculty and students of the center from the Law School, Econ Department, Public Health, Political Science, as well as Public Policy and Public Administration. And increasingly we're using data analytics tools. And when I say me, I mean the young people on my team, and I just admire what they do. So if you're interested in the Reg part of RegTech, I would invite you to visit our website, it's regulatorystudies.columbia.gwu.edu. Sign up for our Regulation Digest and to get notifications of events. We're actually having an event next month on agile regulation working with OECD. And I think there may be enough overlap in that people here would be interested in that. So back to you, Jason.

Jason Fichtner: Thank you, Susan. Kirsten, same sort of question, brief introduction, please, and tell us about your work and some general comments.

Kirsten Gullics...: Great. Thank you, Jason. Good afternoon, everyone. And thank you for having me and thank you for inviting me. My name is Kirsten Gullicson and I am in a nonpartisan support role in the Office of the Clerk in the House of Representatives. I have been in the Clerk's Office for 24 years, and prior to that, I worked for my home state congressman. And I have received some of those phone calls from Hugh Hefner when he couldn't get that legislation printed. I support that XMetal editor. And I also am part of our working group, our leg branch XML working group, to set those standards for that United States Legislative Markup, that USLM that he was talking about.

And so I really wanted to just share with you that standard setting is critical, but it's an often dry topic, but there is a small dedicated group of us from the House, the Senate, GPO and the Library of Congress and others who are doing just that. And in standard setting is most often across our organizations and agencies. It's difficult work, adding that digital layer or replacing, adding a digital layer to your paper process, or even replacing the paper with a digital layer, but it can be done.

And as Hugh said, they have played a key role in managing our generation to Schema, again called USLM United States Legislative Markup. In the USLM, XML Schema builds upon the Akoma Ntoso international standard for representing

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

legislation, executive and judicial documents in XML. And I was part of that work to get Akoma Ntoso approved through the OASIS Standard. The USLM was first made available for the US Code through our Office of Law Revision Council here in the US House, and USLM characterized as a second generation XML because it provides more flexibility, than for example, our first generation build DTD or what the FR and the CFR XML is being used. So GPO does translate a legacy as GML print format into FR, CFR XML.

And what I want to share with the group is that in parallel with our effort to model the US LM for our enrolled bills, public laws and our statutes at large, GPO and OFR completed a pilot project to model and convert into USLM some of the subsets of the Code of Federal Regulations and Federal Registry. So a major outcome of that pilot project was a significant update to the USLM XML Schema, so that documents in the Federal Registry and the Code of Federal Regulations could be printed and published using the USLM format. So I'm really proud of that work that we did together and in parallel. And some of those meetings where we had all those organizations together and people together were pleasant and good, and we got that USLM Schema to a place where we could prove that it could be used for the Code of Federal Regulations.

Again, cross organizational work can be difficult, standard setting can be difficult. Again, adding that digital layer to our centzole old paper process for law making is challenging, but again, it can be done. And the benefits of using USLM in the underlying Schema Akoma Ntoso provide the flexibility that is needed so we can build systems that are interoperable, and provide that high value not only to the public employees who create our law making documents, including the bills, resolutions, and regulations, but for that public who need and want to participate in all levels of our process. So I'm really excited to be here and to hear everyone's work that they're doing and contribute our experience of trying to meet that citizen engagement from this side of the law making process. Well, you all down there are working with regulations. So thank you.

Jason Fichtner: Thank you. Thank you, Kirsten. I love your colorful background, that's really cool to have for a virtual environment. [crosstalk 02:34:01] I appreciate your patience. Sorry, go ahead.

Kirsten Gullics...: Oh, no, I was just saying, thank you.

Jason Fichtner: Okay. Reeve, sorry. Thanks for your patience. Please, the floor's yours, sir.

Reeve Bull: Wonderful. Thanks Jason, man. And delighted to be here. So I'm Reeve Bull, I'm the research director at the Administrative Conference of the United States. I'm a lawyer by background, I was in private practice prior to coming over to ACUS. And I think what that means for our purposes is you're going to get a lot of dilettantism. I think from my end, I've probably know the least about the technology than probably anybody in the room. And my agency, ACUS, is not

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

sort of technology focused, but what we do a lot of is work in the regulatory space. So what I'd like to sort of briefly outline, and what I hope I can contribute in the discussion is sort of the perspective from the stakeholders that engage with this process and experts in the regulatory space. We at ACUS have been involved in the rulemaking, for quite some time. When we reopened our doors in 2010, obviously that was sort of one of the most important changes that had happened in the rulemaking space in a long time. So it's something that over the last decade or so, we've really had a significant focus on.

And I think as Virginia and others have said, rulemaking was sort of effectively a paper based process that was then translated to an online environment. And I think for the most part it's worked reasonably well, but the system's over 20 years old now, and there are definitely some strains that have emerged. That's why I'm especially excited about the work that Virginia and Tobias and others are doing at GSA to really modernize the system. I think there's enormous potential here.

If I could just sort of highlight a few of the things we at ACUS have looked at that I think are relevant to the topic, so one is you've got regulations.gov, it's I think a relatively straightforward user interface, but increasingly, a lot of people are much more accustomed use of social media more than anything else. That's sort of a Web 1.0 or 1.5, or whatever metaphor you want to use, but it's not necessarily how people always engage. So one thing we looked at is sort of how to use social media to enhance citizen engagement, at least points them to this mechanism participating. Another sort of basic thing we looked at, but was just making sure that all the content gets available online. You do, in some instances still have people submitting paper comments. So ensuring that those are all available on regs.gov for people to access.

One thing we looked at recently, and that I worked on pretty extensively as a consultant, in addition to my research director role was this so-called phenomenon of mass computer generated and fake comments where you have certain rule makings that get massive numbers of comments. And increasingly those are being generated by computers, by bots, rather than by actual human beings. And that's created some real challenges for the system. And we looked at some ways to address those issues and also for the agency to be more transparent in terms of what it's doing.

And then finally, one thing that we're currently looking at an ongoing ACUS project is the use of AI, particularly to look at existing regulatory tax. And this is not just in the federal government, it can also be state governments or international governments process it, and try to get a sense of where regulations might be needlessly duplicative, where there might be errors, where there might be opportunities for streamlining. So that's an ongoing study that we're doing in ACUS.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

So, I think these are all fascinating topics. I think that technology creates a lot of challenges, but I think it also creates some major opportunities. And that's why I'm very excited for this panel, because I think there's a lot of great work that's being done and a lot of applications that can really make agencies both make the process work better for themselves, but also engage the public in a much more meaningful way.

Jason Fichtner: Listening to your conversations and opening remarks, and then being here today, listening to the other panelists, you realize there are obviously a lot of challenges and there's a lot of gaps. So one of the things I want to ask all of you, and Susan, I'll ask you to go first and then we'll go around the panel. But thinking about these gaps, what do you consider the biggest challenges that are facing the regulatory ecosystems use of technology today? And maybe some areas for greatest improvement, because sometimes we don't even think about the idea of, can you copy and paste from a website, for example. But I'm sure there's a lot more things that are complicated we need to talk about. Susan.

Susan Dudley: Yeah. And probably my thoughts on that, Jason, are going to be a little more low tech than the others. So the old practice since 1946, the Administrative Procedure Act that I think Virginia mentioned and Hugh mentioned earlier, is you develop a regulatory proposal and you base it on models and assumptions of how the world works, and then you seek public comment and then you issue the final rule and then you enforce the final rule. And I think we're rightfully proud as Virginia mentioned, of how the public comment process is such a long standing and important part of our regulatory development process. And it has become more sophisticated since 1946, the analyses are more complicated and sophisticated and the preambles are lengthier and rulemaking really has improved opportunities for the public to get involved. But we still are constrained by that linear process, which is not viable for a rapidly changing and evolving landscape.

So a lot of attention now is being paid to agile or adaptive regulation in which agencies embrace a learning agenda and are testing and evaluating. And that needs a continuous feedback loop versus that one shot opportunity for public engagement. We do it once before the rules in effect when we really don't know what its impacts will be. So there are huge uncertainties for some of the challenges we need to tackle now, but I think those uncertainties mean real opportunities for regulators who are willing to use data to experiment and learn.

And let me just take a couple more minutes and talk about national academy of sciences project that I was involved in developing a consensus report on how the Department of Energy sets their fuel economy standards. So the statutory goal that DOE has is to achieve the maximum improvement in energy efficiency that technology feasible and economically justified.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

So to set standards, DOE gathers lots of data. In fact, they even hire contractors to go and buy appliances off the store shelves, break them apart and figure out how much each element costs. So it's very heavy on that on models and on assumptions, but there are very large uncertainties in those ex-ante estimates and there's large variability across populations yet there's very little ex-post evaluation of the validity of those ex-ante estimates. And I don't just want to pick on DOE I think that's something that we see in regulation across the board. So the National Academies Committee that I was on, recommendations focused on both quantifying and presenting uncertainty better, but also collecting better data.

I can talk more about uncertainty if people are interested, but I think more relevant for this is regarding the additional data. This is a quick quote, "True improvements are likely to come from gathering and using new data, rather from grafting more sophisticated economic models onto the analysis core. So observations on how producers and consumers actually are behaving in response to the standard, measure of the Institute performance of the products, data for conducting ex post analysis that either validates those assumptions made in the prior standards or suggests where maybe the biggest uncertainties were, and we can drill down and improve that, and evaluate the implications of the prior forecasts, inaccuracies and mistakes for future standard setting." So I'll stop there. But I think that ex-post analysis and analysis in real-time is something that new users of data could really improve, not only existing regulation, but future development of regulation.

Jason Fichtner: And Virginia, I'm going to turn to you and ask you what your thoughts are on the biggest challenges, where you think the gaps are right now.

Virginia Huth: With respect to Susan's comment or anyway?

Jason Fichtner: Open. You can take Susan's comments if you want, or whatever your thought.

Virginia Huth: Well, I'd say I think the gap to help us achieve our goal is that we need to be able to have more standardized metadata standards, in order to really enable more crosscutting analysis across different regulatory domains. And then even with our partners at GPO and at the US Congress. And I know Tobias, we've had conversations with our partners down there as well, and that's ways off, but this is where data standards are really like the rivets that hold it all together. But I think once we have access to that data, I think this is exactly what we can do. We can do better predictive analysis about the outcomes of regulations so that we can avoid unintended consequences. We can better assess the DEIA impacts, we could assess the cost benefits, we could do more research.

And back to your point, Susan, I love the idea, there's been discussion about this. It's a little outside my domain, but some agencies are talking about more discussions in the pre-rulemaking stage. A lot more public back and forth, whether it's social media or just discussions, because right now you basically

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

issue the proposed rule making and it's boom, there you are, comments. It's like, boom, boom, it's really limiting in the way the process is constructed. So I think there's tons of potential compliance in another area that we've barely touched on, but there's room for a whole nother panel on this. And I hope you get the chance to talk about all the applications for data in improving regulatory outcomes.

Jason Fichtner: Reeve, do you any additional thoughts on this topic?

Reeve Bull: Yeah, absolutely. So I guess sort of the first thing I'd say is I strongly agree with what both Susan and Virginia has said. I think that I'm not sure I have a whole lot to say that sort of innovative beyond that, but I think that they lay it out very nicely. And I think that's really an important way of looking at this. And I think is indeed perhaps the biggest gap is if you look at where the public involvement currently takes place, it's sort of in the middle of the process, if you will. It's when the agency has a notice of proposed rulemaking, where the agency already through to some extent, what approach it wants to take and then just soliciting the public's views on that. But there are steps proceeding that point and then there are steps following that point.

So there's a whole extensive process by which the agency goes about deciding where it wants to regulate or how it wants to regulate before it even gets to the NPRM. And then once the rule has been finalized, as Susan mentioned, there's a whole host of information that I think could be enormously valuable in terms of assessing how the regulation is working, were the initial estimates accurate, was the regulatory impact analysis on point, or does reality diverge from the predictions in certain respects?

And I think the way that we've currently structured, it really limits us in terms of really only getting useful outside information or predominantly getting useful outside information and that at middle step and the notice of proposed rulemaking. So I think that a critical aspect of this is sort of thinking where can we engage, the public and experts and stakeholders more meaningfully when crafting new regulations, and then same question with respect to existing regulations. Especially because particularly in this tech space, there's an enormous amount of expertise that exists in the private sector that really the notice and comment process was designed to leverage that expertise.

And I think we could do a much more effective job of that, engaging the public and experts and private sector entities, particularly in that sort of early step of the process, where do we want to regulate, and then the back end of the process are the regulations working? And the two of course are connected to each other because your answers to the ex-post analysis then governs, do we modify existing regulations or do we adopt new regulations or rescind existing regulations? So it's really a whole life cycle. And I think public input can be very, very valuable at all of the steps. And I think sort of the existing process is

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

somewhat constraining in that sense. So that is what I would see as sort of the gap trying to extend the process to the whole cycle.

Susan Dudley: Yeah.

Jason Fichtner: Kirsten, I want to give you a quick chance to respond to that and also, but also maybe tell us a little bit more when you're given the response of the gap, because what we're hearing today too, is some of these things that we're trying to improve are still like a few years away. So one of the things you were talking about with the USLM Schema is can we do something with that now? Can we do something today with tools as well? And how does that fit in?

Kirsten Gullics...: Yeah. Yeah. Thank you. I hear so much, which is very parallel to what the challenges we face in Congress is that as Hugh said, we don't have a single unified code. And so we're going to different databases of the law in trying to figure out how does that bill proposal impact current law. And that law again is in a couple different places, in the statutes at large, in the US Code, in the statute. So some of the lawyers in the room, or some of the policy folks who might have heard positive and non-positive law. And so in one of our projects where we're using USLM, which I think is relevant to this discussion is we're trying to do comparisons in trying to do comparisons from one document bill version to another bill version, but also do the comparison to the current law so that we can serve up a user experience where folks can understand the changes in context, right.

And so that's one thing, understanding the changes in context, but then there's that temporalness. So just like in regulations, Congress has a temporalness. We know that the law changes sometimes our effective dates in the law are not set at the same. So it's very similar to a proposed regulation, a comment period, a final regulation. And then what's what I heard Susan say is how do you make that so it's more agile so that you can go in and modify a particular piece of the regulation to be more relevant today in a quicker path.

So USLM and in the same standards of having the proposed regulation and the final regulation and steps in between can allow you to build tools to see all the changes in context. So you know what the current regulation is today, you can see the PR posed changes and then coming up, see what the final changes are in all three stages, that's a challenge, but that standard settings, those things that Virginia also talked about in using natural language processing and her examples of how do we figure it out in our unique context of regulations, we're doing in the same thing, because we know when we look at these documents, we want something different than word track changes and USLM can put all of that into that context theme. So I hope that was helpful, that temporalness

Jason Fichtner: Zoom and Teams and everything else that we've been like doing these, watching the chats and asking questions. And I love being in person because we can go back to old school where people can slip me notes as opposed to this is the old

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

school chat where someone writes a note. So this is a good way to use the note process and being in person. And this sort of leads into the next question, about you wanted to make a comment on the common advisor tool, which goes into how we sort of use technology.

Virginia Huth: Yeah. Count me out.

Jason Fichtner: No, no. I like, but I think it's good to be back in person. I like having to watch a chat, we can just do this back and forth. It's great.

Virginia Huth: I know there's so much to say, but I have to brag a little bit about a particular aspect of this proof of concept that we conducted recently. And again, it's just an example of the ways we can engage the public in different ways, there's probably others. But what we did with this particular proof of concept was set up a sort of a comment advisor. So a member of the public, a commenter could come in, look at the regulation and perhaps answer some screening questions about their interest, their profile. I'll use the more generic example to make it easier. Maybe you were looking at a Medicare rule and you might want to... Are you a patient? Are you a doctor? Are you an insurer? Are you in a hospital? Whatever.

And so certain things might come to you and you could then comment more specifically on those. Of course, you can comment on anything, but this helps target to you the parts of the regulation that are most relevant. You could also, if you were somewhat knowledgeable, pull up other regulations that you wanted to compare and then be able to do a direct contrast to what's similar or different, and why you think they need to be reconciled.

The system saves your profile, if you will, and it doesn't have to be associated with your identity, it's just a persona. And this helps with crowdsourcing so that the rule makers might say, oh, this is what all the doctors are saying, this is what all the patients are saying. And so there's a way to sort of aggregate it. And then ultimately all this information can be racked and stacked so that the rule makers have a more efficient and effective way of going through all the voluminous comments. And the comments are then pulled out instead of like getting one big letter with 10 different things in it, they're all put in their proper buckets. And so again, it's a way to more effectively engage with the public for those meaningful comments and substantive comments, and also enabling the rule makers to have access to that. And again, there's all sorts of different things we could do, but that's what we tested in our proof of concept.

Jason Fichtner: Thank you. And just so that all the panelists know we were a little bit behind on schedule, so I don't want to anyone's time from the panel discussions. We'll wait another 10 or 15 minutes and then we'll call our panel just so the panelists know we still have some time, don't get nervous.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

But Virginia, thanks. That's actually a really good segue into the something I wanted to ask the group, and Reeve, I'm going to start with you on this, because your organization spent a lot of time and energy focused on transparency and citizen engagement. And that's something I mentioned in my upbringing remarks about, if you don't have this kind of engagement, you don't have transparency, you don't have legitimacy. And I think it's important we spend a few minutes just talking about how these tech solutions really could advance engagement. And so if you think about like where should citizens be involved and how we can make it better for them? What do you think some key areas are? And then I'll ask the rest of the panel to comment as well.

Reeve Bull: Wonderful. Thanks Jason. So yeah, happy to offer my thoughts on that. And let me say what Virginia just said really resonates with me a lot. That's actually, I think fantastic, the work that GSA is doing in terms of...

PART 5 OF 6 ENDS [02:55:04]

Reeve Bull: ... is doing in terms of gauging the type of input or the tools to seek input based on the type of question the agency is asking. So let me provide a little bit of background of some work we did that I think really shows why that's such a critical issue and why it's so key to the legitimacy of the rule making process more generally.

So, I mentioned a study we did, I think it was last year, on mass computer generated and fraudulent comments. And, one of the major findings is the public really tends to think of the rule making process as something akin to a vote, basically. They file comments and say either I agree with this or I disagree with this, save the whales or whatever. And, it's generally not what the agency is looking for. Under the APA, under the process that we created for rule making, the goal is to get information from the public. And, it doesn't necessarily have to be technical information. It could be, how does this rule affect me personally? Or what would be some of the downstream consequences of a rule that an agency wouldn't anticipate? It doesn't necessarily have to be science or economics, it could just be your own personal experience with the rule.

And, that's where I think the transparency is really paramount in terms of the agency being clear on what type of information it is that it's seeking. And, there are a number of different models. A lot of agencies are already doing this. They're already including targeted inquiries in terms of the types of information they're looking for. European Union, actually, I think does a very good job of this. They're [inaudible 02:56:53] when issuing a proposed regulation, usually have a set of targeted questions.

And then I think it's also, there's really no replacement for the agency just being very candid with the public. What we're looking for is primarily information that explains what the effect of the rule would be. We're not looking for a vote, one

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

way or the other. And a handful of agencies have sort of said that explicitly, in terms of their comment policy. And I think that's good. I think that it makes it clear what the purpose of the process is, and I think it saves frustration on both ends.

I think with respect to the public, there's often this misperception, and you particularly saw it in connection with, I think most profoundly the FCC net neutrality rule making, where there was something like, I believe 27 billion comments that were filed, and the vast majority of them were either, this is a great idea, or this is a horrible idea. It really didn't have a whole lot of technical input, one way or the other. And there were, John Oliver and others actually covered that particular role, and I think that kind of fed in this sort of narrative that it's a vote.

But it also creates frustration for the agencies, in that then they have a massive number of comments that they have to process. So I think that's where the agencies being clear, in terms of what the purpose of the process is, and then trying to guide the public to the greatest extent possible. This is the sort of input we're looking for, this is the type of information we might find most valuable, is very, very important in terms of managing expectations on both sides, and I think ultimately maintaining the legitimacy of the system.

Jason Fichtner: Point we have to mention, because I think when the agency gets comments, they sort of know. I remember when I was, when my staff walked in one day [inaudible 02:58:45] security administration said, "We got the comments back," "How many?" "Thousands" I'm like, "I'm not going to read them all" "Don't worry, sir, we're going to categorize them, but this is the citizen engagement, you want the government to read it and make sure you have the information." Susan, I want to turn to you and maybe have a quick discussion about what you think could be the best way for citizen engagement. Because you talked having like retrospective review and that's on the agency side to do better, but how do we get the citizens involved as [Reeve 02:59:10] was mentioning?

Susan Dudley: Yeah, I guess I'd like to reinforce what Reeve and Virginia just said, but also go back to Reeve's earlier comment that there are a lot of stages during the rulemaking process to get involved. And I think in addition to my hope that agencies will do better evaluation. So that's engaging affected parties by understanding how the rule is actually working, but also in advance. So I think earlier engagement in the rule making process is another, an important way to engage the public and be more transparent and more accountable. We miss, where there's always that gap [inaudible 02:59:56] you just said.

Jason Fichtner: So, that was my fault this time then. So Kirsten, where would you say we should focus for civic engagement? And then I'll do a wrap up question for all of us.

Kirsten Gullics...: Yeah. I think I agree with all the comments and I also, I think Virginia's point about really understanding the semantics behind the comments and the

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

commenter in the value... There's like weight, there's weight to the comments. So you have to, how can you build tools to weight the comment so that the rule makers know which ones really to listen to and really are helpful in refining the proposed rule. And then how do you respond? Because you have to respond to all, like you said, you have to respond to all of them. We have... We, as citizens expect to be listened to and we as government employees need to listen to them, but how do you that in natural language processing? And some of that AI can surface that up. It takes some work, but I think it will improve the efficiencies and speed up that process and meet what Susan's talking about, being agile so that we can respond to just how fast everything is happening.

Jason Fichtner: Final question for everybody, I'll go around the room. We'll do reverse order for how we started, which means pretend you're going to get the last word as opposed to the first word in. Reeve, you're going to get to go first. One of the things that was interesting about today too, is we saw we had some chief data officers here and the chief data officers are kind of new to the C-suite. So, if you could give one recommendation to chief data officers in government agencies regarding RegTech or data solutions for regulations, what would you encourage them to do to foster their work and foster everything we've been talking about here today for civic engagement, proper use of tools, data, et cetera? Reeve, you get to go first.

Reeve Bull: Wonderful. So, yes, I guess my main recommendation, and I think it sort of draws on what we've been discussing today is really to sort of keep the big picture in mind. I think that, and in particular sort of keep the stakeholder in mind, keep the public in mind that when we're talking about the rule making process, it's open, it's transparent and members of the public are heavily engaged. And so it's important to make sure that the data is available, but it's also important to think about the end users. So making it available in a usable way, in a way that sort of accords with what people are accustomed to in terms of non-government websites that they may interact with. And then also... But also giving some thought to the technical aspects of it and making it available in a readable way, in a way that facilitates private sector applications, where there are a lot of companies, a lot of nonprofits and not a lot of other entities that I think could really do some interesting things and can actually provide a lot of input to agencies.

That's very, very valuable, if they can get access to the data in a way that's usable and facilitates these sorts of applications. So I think that would be my overall recommendation is to sort of keep the big picture in mind and particularly keep the end user in mind and what will be most valuable to them.

Jason Fichtner: Thank you. Kirsten, you're sitting in a room with a bunch of chief data officers. What do you want to tell them?

Kirsten Gullics...: One key lesson that we know from our USLM work is get all the players in the room and think as if you are one agency, think of it as if you could do anything in

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

the world and tackle the problem from that perspective first, and then identify your institutional challenges or your roles and responsibilities. Because if you can give your tech people, your human centered design folks and your longtime subject matter experts who aren't always necessarily going to come to the table with that human center design, that some of our, I think Susan talked about the younger folk and some of the newer people who want to help us build tools, if you give them the freedom to think like a one central unit and then tackle your institutional challenges or your silos, and say, "Oh, well, we might be able to get it done." You're going to have better outcomes.

And, that's really the hard cross organizational work is to get everyone in the room to trust them, to do that big attitude. And then journey mapping, walk through all your work, walk through the journey mapping, there's value now in that new process of getting emotions out and how your public feels and your internal staff feel and really understand the workflow for, the problem you're trying to solve.

Jason Fichtner: Thank you. That's really good advice. Susan.

Kirsten Gullics...: Thank you. Thank You.

Susan Dudley: I guess, for the CDO plan for ex post evaluation, help agencies plan the very beginning of the rule making, what data needs will there be and how will they be used, but I'm going to offer a broader piece of advice. And that is humility, that bigger data should not lead... shouldn't necessarily mean a bigger purview for government decisions. For example, in the energy efficiency standards, again, connected devices and the internet of things may mean that consumers can see their energy savings in real time from the more efficient appliances, and it may negate the need for these standards altogether. So there are really exciting opportunities for big data to address problems that make government not necessary in some areas.

Jason Fichtner: That's very good. Thank you, Susan. [Jenny 03:06:09], you get to bring us home.

Virginia Huth: All right. So I have two pieces of advice. The first one is we need to converge on some data standards so that we can have the interoperability across domains with GPO, with the Congress, with other entities, potentially with States. And dare I say, even international entities, thinking about my colleague over here from the IMF. So one, some standards so we can really use this data across all sorts of domains and improve the accessibility. Two, it's great to have data standards, but you got to implement them in real systems. And I think there needs to be a partnership with the CIOs and the CIO council. For example, rule making, I suppose, we're lucky we have a 20 year old system, it's practically a tear down. So we get to build from scratch and we'll be building in standards, but a lot of systems they're archaic. If, trying to retroactively do this, it's going to be impossible. But as we rebuild systems and think about investments, we need to make data standards core to that investment decision.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Jason Fichtner: Thank you. And because we are running over time and I wanted to get us back on, we're going to skip the Q and A, but I appreciate everyone staying around today. I appreciate the panelists all being here and the hybrid environment, while not perfect, I think worked pretty well. So please give a round of applause for all the panelists today. Thank you.

Nick Hart: All right. Well, thank you all so much. This was a great panel. Great discussion. Jason, good to see you. Virginia, thanks so much for being with us. And to all of our virtual panelists, really appreciate your insights. We'll free you from the stage and we're now going to welcome our next fireside chat. Our next speaker, Mike Willis. And while they're queuing him up, I will just say for those who were online for the opening remarks that I offered, the discussion about humility, that Susan [Devly 03:08:11] just mentioned, she's also a former white house staffer.

So we did not prep her for that, but it's just another connection to the fact that when you work there, you learn this lesson about humility. So it's truly a crosscutting theme. Mike, good to see you. Thanks for joining us for the closing session of our RegTech Data Summit, for this year. For those who are familiar with Mike Willis' work, I'll just introduce him very briefly. He's currently the Associate Director of the Office of Data Science and Innovation in the Division of Economic and Risk analysis at the SEC. We've invited him to join us also as the Chair of the Regulatory Oversight Committee. And Mike, can you hear us okay?

Mike Willis: Can hear you great, Nick.

Nick Hart: All right. Good. Thanks for being with us. So the audience has just heard in the last little bit from experts in the Design of Regulations, including a former administrator, Susan Devly from [inaudible 03:09:11] at the white house, encouraging strategies for public engagement, participatory processes, the use of technology and data. And I guess I'm curious as you are listening into the discussion in the virtual green room, your career has taken you from regulatory agencies in the private sector, you've worked with federal agencies, what's your sense of how well we're really positioned to make some of these transitional steps today administratively, as we're thinking about the customers, the American people, and actually participating in these processes by better using technology? What's your perspective on the same question?

Mike Willis: Nick, thanks for having me today in this important discussion, but let me start with my SEC disclaimer. The SEC disclaimer's responsibility for any public or private statement by any employer or commissioner. This discussion expresses my views and does not necessarily reflect those of the commission, other commissioners or other members of the staff. Okay. So got that out of the way. I think we're very well positioned and I suggest that the insights on the process transformations may be the most useful way to engage the participants so that they fully understand the capabilities and features of these new technologies and standards. In other words, I think we should answer the question. Why

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

should I engage, what's in it for me to do so? I mean, we remember... Nick... A cell phone. Why do we need a cell phone? I have a phone at the house.

I have one at the office. I don't need another phone. People cannot conceptualize what that meant initially. And so if we look at this process of how do we engage in the request for comment process, the current method today among agencies is to engage the market participants. What you find is a mountain of unstructured letters that requires a small lawyer armory to process. It's a highly manual process, is well out of step in today's digital age of survey monkey and all these structured languages and standards. Could we make that more structured and standardized, the whole process? So that enables the more effective, and not only communication to the agencies, but analysis by market participants. How's that information available to the market participants so they can easily analyze it? So maybe to just wrap up, I want to just speak to AI.

I do think it is something that we can apply in the comment process, but I don't think we should use it to structure the unstructured input. That's not what we should do. We should not use the AI to initially structure the unstructured elephant comment letters. Rather, I think we should use it to analyze the individual structured bite size elephant pieces. And yes, I'm referring to unstructured comment letters as elephants, in this analogy. Absolutely. And the reason I'm doing that is because everyone knows the best way to eat an elephant. And, that's also how AI might most effectively be used to analyze comment letter feedback, just one bite at a time.

Nick Hart: This imagery of trying to figure out the best way to eat an elephant, so. This chair, the ROC, the Regulatory Oversight Committee that you've recently been tapped for has a lot of responsibilities. And among them, you work on an issue that many of us are familiar with the LEI, the Legal Entity Identifier, derivative identifiers. So this is a concept that many in the room here are familiar with, a form of data standard. Tell us a little bit about this. I think a lot of the folks who are tuned in virtually and here in the room are familiar with this, but this is also a new concept for many in our audience. Tell us why they should care, why does this matter?

Mike Willis: Well, I think in short, lower cost, streamlined processes and better analysis, that's what's in it for you. The Regulatory Oversight is a group of about 65 government agencies from around the world, we're providing consensus based oversight on the identifiers that Nick just mentioned. I guess I'll just stick to LEI as an example. This was developed as a reaction to the '08 financial crisis, and then the underlying use of proprietary identifiers, that resulted in manual transaction and clearing processes and manual counterpart risk assessments.

Did I mention, Nick, that this created manual processes, were sort of back to the theme of the comment letters as well? When you have standards like the LEI, it provides a common, unique identifier, it facilitates more automated clearing,

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

more automated risk assessments. And I think that's the key punchline, standards facilitate automation and streamlining of processes, which results in lower cost and more effective analysis. So that's why I'm promoting the idea of the LEI, why I'm here today, and I appreciate it. But I'd also mention that this is for any standardized identifiers, not just the LEI and the derivative identifiers, things like the FIGI, the Financial Instrument Global Identifier, which is useful for financial instruments, these standardized identifiers, every one of them enables those same characteristics of streamlined processes, lower cost and quicker analysis.

Nick Hart: Topic that gets a lot of emphasis when we talk about financial services. And I guess I'm wondering if you have examples of the business case outside of financial markets. Are there examples of where this might be relevant or useful in other contexts or other business cases that you track?

Mike Willis: Absolutely. I may give you half a dozen real quick. First one is just legal structures. Some businesses have obtained the LEI, to just lay out their organizational structure, their hierarchical chart. So if you're a business, a question to ask is, where do you manage your legal entity structures? Second is compliance. Nick, I recall, a data foundation paper a few years ago, that identified 50 legal entity identifiers on the federal agencies. So imagine if you will, a compliance process where a single identifier could be used by a commercial entity in dealing with all federal agencies, all state agencies, and literally all agencies around the world. Third, would be analysis. This is a very basic idea. When you go to analyze different disparate data sets, the first thing you have to do is map them together. And the LEI provides that common thread that will facilitate that mapping and make you able to do the analysis much quicker.

Fourth is KYC and OB, this is clearly more of a financial services sector issue, because in that sector, if you don't know who your customer is, you could be facing significant anti-money laundering penalties, but I would suggest that's a useful habit for any business that wants to fully understand the entities that they want to do business with. Fifth, is cross border activity. And the recent supply chain constraints have really brought this to light. And I would want to point out that the US customs and border patrol pilot of their global business identifier, which combines the LEI with the global location number to allow the tracking of cross bordering partners and transactions. That's a good example. I'd also note, just to give people idea of the scope of international activity, is the People's Bank of China, and the Reserve Bank of India are also mandating the LEI for their cross border activity.

So, that's two very large countries that are jumping into this boat. And last, I'll mention is the World Trade Organization and the International Chamber of Commerce just released their standards toolkit for cross border paperless trade. And that also speaks directly to the LEI and the GLM, the Global Location Number. And then last I'll mention, this one's probably a little out of the box,

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

that's cyber security. And this is around enhancements to the traditional firewall based protocols that are enabled via a Zero Trust Architecture, which uses digital identities. Now those identities can be tied to an agency specific identifier, or that can be tied to an international identifier, and there is the connection to the LEI.

Nick Hart: All right. So cyber is an interesting one and it's actually a topic that has not come up much in our dialogue today. I'm wondering if maybe you can explain and maybe expand on that just a bit. What's the cyber connection here?

Mike Willis: Sure, absolutely. We know this is an issue for all market participants. And I think for everybody sitting in the audience, just ask yourself, how many usernames and passwords do you have? Or put another way, how many of your kids are using your username and password for your Netflix account? This is just a persistent, in our face, issue all the time. So let me connect the dots between cybersecurity and LEI. In May of last year, the White House issued a cybersecurity executive order that mandated the federal agencies adopt the Zero Trust Architecture. In June of this year, OMB followed with a federal strategy to move the US federal government towards a Zero Trust Architecture. Now Zero Trust Architecture's a big word, but very simply put it's a computing environment in which no one is trusted and resource access is based upon your digital identity credentials, you as an individual.

So maybe a few words in how that works, and I'm going to use an analogy here of your house. So your house is like a perimeter has walls around it. That's a perimeter based model. So if somebody breaches the front door, then they can walk in and steal anything in your house. But in a Zero Trust Architecture, not only do you have that perimeter, but everything in the house is encrypted and only those trusted individuals have access to those ideas, those artifacts, whatever they are. So if somebody breaches your front door in a ZTA environment, the jewelry box and the cat litter box both look the same and everything there all looks the same. So where does that third party start? And since they're not trusted, they really can't see anything. They can't access anything. So this is the theme of Zero Trust Architecture.

So in thinking how government agencies [inaudible 03:19:10] with their Zero Trust Architecture, this includes [inaudible 03:19:13] of how to identify the entities, to which digital credentials are assigned. Is this a continuation of the traditional agency specific model? Or should we think of about an internet standards based supply chain approach? Should agency specific identifiers be used that would continue this one to one, like the SEC and the CIK, not useful someplace else, but that's the one to one model. A supply chain model like the LEI would enable one to many. So not only could the registrar use it with the SEC, they could use it elsewhere. Now, just to flip this around for a second, this is the question that really helped me understand and that is, how will the user of a digital certificate validate or verify the digital credential? Well, I think it would be useful to have the entity responsible for issuing the digital credential

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

somehow wrapped into the digital credential. Is the issuer from a large federal agency, a state department of motor vehicles, or is it from ABC hacker company, Inc?

In considering the inclusion of the issuer within the digital credential, one must consider how that issuer is identified. By a name? I don't think that works, we all know. Or some form of identifier. Now we're back to the real value of the LEI as it offers that international standardized unambiguous identifier for legal entities issuing those digital credentials. Now, to the extent that an issuer might use an identifier that would present another level of opacity for any file [inaudible 03:20:49] trying to validate. SEC were to use the CIK, it's just going to create more opacity in the marketplace for people trying to use it. Now, all that said, the punchline here is really it's a supply chain approach and just as your prior panel said, international standards are really beneficial in this context. And so as agencies think about this, it would be useful to think about how these digital identities not only work with their regulated entities, but how it would work by the regulated entities using them in their respective supply chains.

Nick Hart: All right. Well, thanks for connecting the dots so clearly for us on that one. We promised to be brief with your time, and maybe I'll just let you have a final word here. As we're closing out our day of our RegTech summit, are there any final thoughts or last words that you'd offer to our audience?

Mike Willis: Yes, in less than 60 seconds. We're all leveraging the internet for business and compliance purposes. And that platform works based largely on standards. Now, data and identifier standards are entering in this discussion through either the open government data act that you mentioned earlier, Nick, through the Zero Trust Architecture implementation, or other analytical demands. Now history tells us that standards, lower cost, streamline processes, enable greater throughputs and other benefits. So as the agencies think about the Zero Trust implementation, they could leverage the international standards such as the LEI and the W3C digital verifiable credential, as a supply chain approach. And folks can learn more about that on the LEI digital strategy page. So more broadly as you look at the future, it may be very useful to consider these standards for data identifier and the credentials that we just talked about. Nick, I really appreciate you letting me share some time and thoughts today with the group. It's been a real pleasure to see you again, hope to see you in person soon.

Nick Hart: All right, Mike, thanks so much for joining us and we look forward to seeing you soon too. I want to just take a minute to close us out for the day. Let's give Mike a round of applause.

And truly to all of our speakers and participants through the day. Thank you so much for joining us, whether virtual or online, we've had over a hundred in-person attendees, kind of rotating in and out through the day. Hundreds RSVP to participate online for our virtual experience. And clearly what you've heard is that there's a lot of potential for improvements [inaudible 03:23:26] both

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

government and society. We've heard about examples of things that are underway, potential for imagination of what is possible. Importantly, and this is an area that we like to emphasize, the value proposition of what is possible and what government should be doing or could be doing as it's thinking about improving its value to customers or the American people. Today, we covered a lot of ground and I hope you're seeing how a lot of these themes are actually connected and that it didn't seem like a dispersed agenda.

Underlying a lot of these themes are actually some very basic concepts, the value and the need for data governance, processes, leaders, data quality, data standards. And these are things that we spend a lot of time and energy talking about. It's sort of fascinating the number of people that recognize throughout both the public forum this morning and our RegTech comments this afternoon. Often it's not a technology problem. That's kind of amazing, especially when we're working with government, it's not a technology problem. A lot of the technologies that we're describing exist, it's a process and a problem of finding the technologies and getting them in the hands of the right people, in the right places, the right agencies. Sometimes there are technology challenges that we do need to solve for. We need to recognize that and also align the resources and the capabilities where necessary.

So let's focus on those challenges where they're needed. We know there's a lot of policy solutions that are out there, a lot of policy problems that we still have to address. Data Coalition, still working on many of those. I can now mention this because it's late in the day. One of the activities that we are actively working on in the RegTech space, is a piece of legislation called the financial transparency act, passed the house recently and is pending in the United States Senate. Hopefully we'll see some activity imminently that will push it across the finish line. Ongoing implementation of some of the laws that we described earlier in the day. And so much more when it comes to disclosure modernization and the customer experience. We hope that all of you will continue to join us in this endeavor. If you'd like to learn more about joining the Data Coalition as an organization, or as a member, you can check out our website, datacoalition.org, or speak with a member of our team.

I want to take a moment to thank all of our sponsors who have participated in this endeavor today, especially Donnelley Financial Solutions, the EDM Council, The Bipartisan Policy Center, and XBRL US the last three who joined us as special policy partners. Now, if you're still here in person, I want to invite you to our cocktail reception, which will be in the east gallery all the way at the end of the hallway here. And if you're joining us virtually, we hope to see you in person soon. That concludes today's event. So thank you all so much for joining and we'll see you soon.

PART 6 OF 6 ENDS [03:26:27]

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Nick Hart: Today's a really exciting day for us because this is our first in-person hybrid gathering in over two years. And that is not an insignificant milestone for our organization that is really built on collaboration and gatherings. So truly thank you for those who are here in person. Thank you for those who are joining us online.

Today begins as a full day of activities and dialogue about data. And we're starting with this public forum on accelerating AI in the public sector. We'll continue later this afternoon with our RegTech '22 Data Summit. Together these two events are designed to cover a pretty wide net of dialogue. Contemporaneous important set of topics that we need to have, that we must have, frankly, at this moment in our country's conversations about how we're better using data to address important topics. The last two years have shown, and I think many of us have seen, that data are integral to many of the things that are happening in our daily lives, and the things or issues that our country is tackling on a daily basis. Literally the future successes and failures of things that are happening across our country, whether it's related to public health, human services, transportation, agriculture, financial services, or defense. We're going to talk about all those things today in one way or another.

So I'll just be honest. I'm really happy to have you all here today, whether you're in person or online. So on behalf of the members of the Data Coalition, thank you for joining us. It's going to be an educational day. We're going to have opportunities for networking, for engagement, and thinking about the future of our country's data policies. We're going to have many more people joining us at different points in the day. So for those of you that are in person, know that we have invited folks who are going to be joining us for different parts of the dialogue. We expect that different parts of the community will be coming in and out. And that's okay. This is going to be an eclectic conversation. We have over 35 data leaders who will be joining us on stage for parts of this conversation in our different sessions, seven different panel sessions, and the event is going to highlight really different corners of the data community. So we're going to have a broad conversation.

We're exploring AI at the beginning. That's not an accident. It's because successful development of AI is one of the potential areas that a well developed and robust data ecosystem is key for, we think, in the next several years. AI holds enormous promise for the future of our government. And at this stage, we should also expect that government will be better, more responsible, and ethical users of AI in coming days, weeks, and months. And it's really an imminent set of expectations that we must all hold for our government that AI is implemented with care, with transparency and the expertise to improve AI systems.

There's a lot of initiatives happening in government today that have happened with a lot of bipartisan support. The AI in Government Act, the National AI Initiative Acts, a congressional resolution that was bipartisan a couple of years

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

ago, the Biden administration's National Research Resource Task Force, and the list goes on and on. Many of you have heard and some of you are participating in these initiatives.

Needless say there is a lot happening on AI and government today. And that's what we're here to talk about. We're going to hear from experts this morning, who will offer some suggestions, some solutions, some ideas, and maybe even some advice about how to navigate the nuances of the topics that government should be thinking about around bias and ethics and transparency. And hopefully our government will take some of this to heart. Because there are a lot of important themes that we should be thinking about as we're implementing AI strategies in the years ahead. Our work here at the Data Coalition is really focused on identifying problems, but also solutions. So we don't want to just focus on where the challenges are. We want to focus on how we can identify and implement solutions working in partnership with our government. So today offers us this unique opportunity to, excuse me, to discuss policies for making high quality government data more accessible and useful, to facilitate the use of emerging technologies to improve government services.

Before we get to all of that, I'm going to offer just a few logistics notes and then we'll get started. So for many of you, this could be your first in-person event that you are attending in quite a long time. Here at our conference facility, obviously you all showed your vaccination cards when you walked in the door. So thank you for doing that. I will ask for those who are willing to continue to mask throughout the day, unless you're on stage, eating or drinking. DC has lifted its indoor mask requirement, but we do know many are still adapting to levels of comfort. So for those of you, when you checked in, we had you put a level of comfort tag on your name badge: red, yellow, green. We do ask that you just respect those who have adapted different levels of color. That was our quick system for identifying how we could signal that.

There will be food and beverages available for those who are in person. And we encourage you to use those as available. If you're looking for the wifi passwords, they're here on the edges of the room. And if you're joining us on social media, we encourage you to use the hashtag #RegTech22 at any point during the day. If you're online, joining us virtually in our large audience, if you need assistance, use the chat box and we'll try to get to you as quickly as possible with any assistance that you may need.

With that, if you have any other questions, reach out to a member of our team or a member of the staff here on site, and we will do our best to answer any questions that you have. Alright, let's begin our discussion of the public forum. So without further ado, I'd like to invite our first speaker who's joining us today. Who's the director of the federal AI implementation, AI Center of Excellence at GSA, with opening remarks from Neil Chaudhry. Neil, please join me on stage.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Neil Chaudhry: Thank you. Thank you so much for allowing me this opportunity to come and talk to this wonderful group of people. I've been cooped up like many of you in my house. I didn't even realize people existed. Right? So thank you for the opportunity. And yeah, please forgive any quirks or things. I haven't been in public in a while. So with that being said, I'm here really in... I have a day job, but what I'm here is as a representative of ACT-IAC. I'm the Emergent Technology Community of Interest Government Chair. And what I'd like to point out is that government and our idea of providing service to the public only works in partnership with industry and forums, such as the Data Coalition. If it's not a coalition, it doesn't work, right?

Government doesn't have all the ideas, best ideas. Neither does industry. We work together I just admire the Data Coalition so much for everything that they have accomplished through partnership with both industry and with government and with Congress. And it's a model that I hope that the rest of the government continues to work with. And so pivoting back to where I am in the Emerging Technology Community of Interest in ATARC. One of the things that we focus quite on, obviously it's more focused on bringing industry and government together in the acquisition space in the partnership of delivery of things.

So in the Emerging Technology Community of Interest, one of the things we always think about is the next generation of technology. What are we doing about that? So for instance, Quantum. We are starting to talk about Quantum right now because by the time Quantum becomes commodity, becomes something that industry can deliver, government should have a very good idea of how to buy it. So it's the same thing with artificial intelligence and other emerging technologies. And especially with artificial intelligence, we have known it's coming. Artificial intelligence has been maturing. And other areas like data and the governance and things around it are still maturing, evolving. But this idea of how we buy services, artificial intelligence services that are driven by data, it's something we're still struggling with.

And that's one of the things that really I worry about quite a bit is that from a government perspective, we want to provide good service. And what happens is we have to define what that means so that industry can provide it in a reasonable way. And then we can just deliver better. From a user experience perspective from all these other perspectives, whether it's efficiency, effectiveness, they all work well.

So what we have been talking about in the ACT-IAC Emerging Technology Community of Interest is how do we share information about emerging technology across government. And this is where the data comes in is the data is there for the most part. Somebody has the data, whether it's in a structured form or an unstructured form. And by unstructured, I mean it's in somebody's mind, right? They know how to buy it, but nobody else knows they know how to buy it. So that's where we start with this idea of what do national or use case

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

and solutions libraries look like in terms of emerging technology, in terms of artificial intelligence. So, what we talk about this idea that if somebody has from a government agency has bought artificial intelligence services, they learn from experience. They have learned the ins and outs. They have learned the taxonomy. They have learned what it means for data rights. They have learned what it means just from experience about intellectual property rights, what a model means, what are you buying a service capability?

All those ins and outs of putting together a good solution. And what we need to do is that if somebody from another agency wants to execute a data, driven, AI driven project it, how do they do it? Do they start from scratch? Or do they build upon the learned experience of somebody else who has done it well in government? And that's where we're focused on. And that's where we're really focused on partnerships across industry and across the government. How do you increase this knowledge sharing on how to best put together a delivery package that can get executed really well? And this is, it comes from partnerships with government, industry, academia, and also forums just like this where we all share information. And it's this idea about how do we share this information with the least amount of bureaucratic friction? How do we make this information sharing seamless while still respecting property rights, privacy rights, and other rights that exist out there. And data rights, for example. So that's where we talk about this common place where stakeholders can share, search, learn about emerging technologies across the federal government.

And it's all about reducing the time it takes to deliver that capability to the American public. That's what this is about is reducing the time to deliver a capability. And this is something that as I talk to different stakeholders and different constituencies about, everybody says yes, Neil, this is a good idea. It makes sense. Sharing of information, sharing of best practices is a good idea. And then the question comes along, why aren't we moving, why haven't we gotten further along in this idea? And that's where it takes a solid belief. It takes active engagement from government and from industry to work together to make this happen. And this opening these communication channels are one way. And once again, I'm so grateful to the Data Coalition to allow me to come up here and stand and talk about this to their stakeholders who may not be aware about the fact that we are trying to get this coalition of the willing together to share successful use cases and solutions across government to expedite delivery.

And a lot of that is also based on data AI, the emerging technology stuff that we haven't bought frequently in the past, but we want to get so good at buying it it almost becomes a commodity exercise. And that's the vision. And really what it is about also protecting our industrial base, our competitiveness in this emerging technology space.

So what overall, what the concept is, it's really a national use case solutions library. It's meant to be a library. And just like a library it's meant to be free.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Obviously you'll have controls in there, just like you can't go in the library and walk out with a book. There are some controls in there because ultimately, like I'd said before, we want to protect the property rights and the intellectual property rights of industry. And also the privacy rights of the public that expects it. And also the proprietary nature of the way government works so that people aren't gaming the system. So it's a really complex equation. But it's meant to be a seamless, free way to assess whether or not a solution that if you're thinking about a problem and you say, Hey, I would like to think about solving X problem in my agency using an emerging technology. This is a seamless way to say, Hey, somebody else in government solve this parametrically similar problem. Let me go talk to them.

And this is right, the key with number three is providing those government points of contact for other government representatives. So you say, Hey, let's... I want to put together a machine learning solution. I've never done. Nobody in my agency has done it. Is there anybody else who has done it? And you say, yeah, there are a bunch of different people who have done it. Here's a phone number. Here's a point of contact. Let me talk to them. What is it that I should be aware about? Oh, data rights. Oh, great. Okay. Now I know how to structure my RFP or my RFI out there so that I can get better outcomes from industry. Industry knows what they need to be providing. So that's really where it boils down to is connecting different people in government and industry on solutions that have already worked for capabilities using emerging technology.

And then this is, we work on a white paper. It's already on our ACT-IAC website. Feel free to take a look at it when you can. And what it is this was a collaboration between government industry and academia, this white paper. So we're very proud of the effort that we put together in a very collaborative way. And so this is a work in progress. So just because we put the white paper out there doesn't mean that we have stopped iterating. So if you, after reading this white paper, you have comments, you have suggestions, feel free to reach out to us and we will work to make sure the next iteration of this paper addresses all those issues.

And really at, at the core of what we are talking about is this idea of reducing bureaucratic friction. Updating FAR, regulation part seven, which talks about market research and acquisition strategy. How do we put functions in there that act as carrots and sticks so that one of the first things that comes to mind for an acquisition professional is, I need to go and do some quick research with other government agencies to see if they have implemented a parametrically similar solution.

We are recommending that we have a dedicated, designated agency representative, a senior official in the agency that's responsible for managing such to library. And we have put a placeholder up there based on our comments as that CTO is the most likely person in the organization that would fit that bill. But once again, it's not being prescriptive. Each agency can designate whoever

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

they feel the best person. And, depending on the agency and their culture, it might be the CVO. So it's this idea about a senior official in that agency that says, yes, I am the accountable officer for this library that I share across government.

And the third thing is how do we work on seamless use case submission? Because if the process to submit a use case to your library is so painful, you'll find ways to avoid doing it. So how do we do that? And this is where we talk about emerging technologies and data in this physician heal thyself model is that, what does it look like? So a program manager and could take all his proprietary or his or her acquisition documents, dump them in a folder, and then we use natural language processing. We use other image recognition, other forms of AI to automatically build the library. So they don't have to spend hours tagging and metatagging things. It's seamless. You say, just give me access to your drive and I'll take care of it. That's where leveraging technology comes in place.

And the last thing is about agency sponsorship. What does this look like? And this is a part of the model that we haven't refined yet. It's working. Is this idea, what does it look like? Does each agency want to maintain their own library? Is there a collective library across the federal government maintained by organizations like the NYE, or JAC, or DAU? We haven't worked that out. And we purposely left that piece out just because that's not a decision for us to make. It's got to be a consensus of government leaders on how they want to proceed forward.

So with that, we had a quick flow chart, and this is sort of where we're looking at is identifying. How do we identify solutions? And these are the typical questions you say, how do I access another federal representative that has these answers? What technologies are out there? Who else has done this before? Where should I start? How can I accelerate the adoption within my agency? Or if I don't even have answers to those, it's like, where do I even start? Who do I contact to get knowledgeable about the subject that I'm going to spend hundreds of millions of dollars of government funds to buy? So it's all these, and this is where we talk about use case libraries.

And, once again, it's not meant to be a one size fits all. It's meant to provoke thought and its idea about, did it work? What were the underlying dependencies that made it work? Who were the stakeholders that helped make it work? Who are your points of contact? What are some of the things that you should consider in terms of data rights, in terms of intellectual property rights to make this work? Do you need private data, proprietary government data, and public data to make this work? And if you do, what sort of permissions or roadblocks do you need to address?

And then same thing with interagency agreements. For a solution like this to work, what were the underlying interagency agreements, the data sharing

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

agreements, the model sharing agreements, what are all those agreements, the funding that made this solution work? And that's really where we're talking about in FAR part seven, how do we update FAR part seven to address some of these concerns?

And once again, final. Please, you can reach out to us. You can obviously, you can't click right there, but ACT-IAC, the NUCSL website, it's right there. We do want to hear from you. We do understand. We are really looking at learning from others. The Data Coalition, the way the Data Coalition put together a coalition to address an emerging need that was evident, partnerships between government, and then legislating it correctly, and all the laws that have passed, we are hoping to take some of your best practices, learn from you in order to make this idea work. So pending that, I think I'm out of time, right? So please, I know I got very excited. It's been two years. So yes, I'm going to gently walk off stage and I'll be off... I'll be here for a little while in case you do have questions or concerns or comments. More than happy to talk. Thank you. And I am green badged.

Nick Hart:

Thank you. Alright. Neil, thank you so much. And also a good segue to say for all of our speakers this morning, if you're here in person, feel free to reach out to them during our networking breaks or at any time you can grab them out in the forum. So let's transition now to the public part of our forum. I'm going to invite our first panel for the in-person folks to come ahead, go ahead and come up the stage, Julia Lane, Brandon Sepulvado, and we'll have two speakers virtually. If you want to come, go ahead and join me right here. Sean Griffin, and Haniyeh Mahmoudian. And we will kick it over to you in just a second.

For our speakers, we are going to ask you to limit your remarks to six minutes. If you're in person, you'll see a clock in the back that will have five minutes. You'll have a minute after that. If you're virtual, you'll see a clock on the screen for you. And for the audience, you are invited to send questions at anytime to info@datafoundation.org, and we'll curate them for you. And if your virtual, you can send them through the platform and we'll try to answer as many questions as we can in the Q&Aa portion. And we're not going to get to all of them, so we will provide your questions directly to the speakers also as a follow up action item. So with all of that said, Julia, I would like to invite you to come to the podium and go first. Julia, take it away.

Julia Lane:

Great. Okay. Thank you so much. It's good to be here. And as they said, good to be here in person. So I'm going to talk a little bit about some practical ways in which we can use AI to get information about the way in which public data are used. So that sounds like a very straightforward thing. It's actually very hard to do. So I'm going to give you a very, a pretty high level overview of what's involved.

So the reason this came up is the Commission on Evidence-Based Policymaking, in which many of you are aware, pushed pretty hard for data, better ways of

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

using data to inform policymaking. The challenge that we faced as we were trying to do some work to inform the decision making of the commission was there are really only three pieces that need to be addressed. One is, how do you put the...bring the data together? The second is, how do you build the capacity to work with the data? And then the third thing is, how do you find out who else has worked with the data, so you can build a knowledge infrastructure? As you know, there's many challenges with that. We certainly found that in the Obama administration, when they created...you had the Open Data Act. Essentially what happened was agencies just provided a vomit of data out that you had no idea how it was used, what information there was. And that really limited the usefulness of that activity.

So what we thought as we started to address that third challenge that the college initiative, we built the Administrative Data Research Facility. We built trading programs. How you going to find out how data are being used? Well, the only way you can really find out how data are being used is to read 80 million publications, right? And 50 million federal documents. So how you going to do that? Well, it turns out that a lot of agencies, as they're responding to title two of the Evidence Act are actually trying to do that. They have human beings manually reading documents to find out how their data are being used because no one references data in the citations.

Now you could whine a lot and try and get people to put references in citations. We've been trying to do that for about 15 years, and it's about as good as me telling my kid to clean up his room. Not going to do it. What you can do is you can put a sticker on the fridge and say, if you clean up your room, you get a gold sticker. That's going to work a bit better from an incentive point of view.

So what we want to do, and what we have been doing is building an AI infrastructure that can automatically read 80 million public research publications and 40 or 50 million government documents to find from the semantic context how a dataset is referenced. So when you read a publication, you as a human being can see how it's cited from the words, from the way in which they're reported, pull that out, put it on a data inventory a lot like amazon.com. Imagine an amazon.com for data. You give people credit for data use, you put their names and rights on, for example, data.gov. You say, these are the experts in this topic in this area. And it's like anything where you get upvoted when you do things well. So that's what we did. We built a Kaggle competition. We worked with a bunch of programmatic and statistical agencies. So we also worked with my colleague on the National AI Task Force. Dan Stanzione. I should have said I'm on the National AI Task Force. And part of what we are doing there is to use AI to improve the search and discoverability for public dataset. So here's the idea: automatically build those GSA data.gov data inventories, document the usage using AI natural language processing methods, populate that with usage statistics, expert...

PART 1 OF 6 ENDS [00:30:04]

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Julia: ... Populate that with usage statistics, experts that we're also building podcasts, where they can have five minute information segment and that's effectively building an Amazon.com for government data that automates and provides the value information of data.gov Oh, sorry of the data that's being used in the public sector. If you want to find out more about it, you can email me.

We worked with the former federal CIO, Suzette Kent who was my boss when I was at the White House at OMB and you can Google Show Us the Data and the college initiative and it provides a whole lot of information on that and then you can see what other work we're doing in the area. But the basic idea is automated, usage based, data inventories using AI natural language process and machine learning tools. Okay. Thanks.

Brandon: Great, Okay. Hi everyone, thank you for coming today and thank you again to the Data Coalition for hosting today's event. Today, I want to draw attention to the problem of ensuring equity when we integrate and analyze many different types of data that are becoming available. So more specifically, I'll argue that, one, we need better methods to ensure AI systems don't harm already disadvantaged groups and two that, once better methods are developed policy discussions and community consensus are still a vital part of the process.

So as a brief example, consider linking a bunch of data sets for which one of the only common pieces of information is a person's name. This might seem like a record linkage problem, and it is but it quickly turns into an equity problem because we know for instance, that certain groups are likely to have much higher error rates than others.

In certain places for instance, with marriage, women regularly change their names in others they don't, still in others they adopt variants of their spouse's name. Also last names have not historically functioned everywhere, the same way that they do in the west and record linkage disambiguation algorithms frequently perform much worse on for instance, Eastern Asian names. When we consider the intersection of these factors so for instance, women in Eastern Asian countries matters are only further complicated.

So the context for which I'm going to say is that, data are more available than ever. This isn't new or profound but I do want to bring attention to it because it means that certain statistical and data science methods come to the forefront to ensure equitable AI systems. So to give you a bit of background on me, I'm a senior data scientist and sociologist at Newark, the University of Chicago. My own work centers on the R and D process from everything like developing knowledge graphs, to help synthetic biologists, to helping funders of health and science research answer questions about the policy implications of their funding activities.

A lot of my work pertains to helping decision makers understand how their programs impact the STEM workforce. And I like to use the STEM career

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

trajectories as an illustrative example throughout my talk today. There are many sources of data on such career trajectories. We often need to understand things like educational history, publications, patenting and the broader impacts of STEM for instance, on policy and the news and online discourse. These data exist within the context of automated systems.

So for example, several initiatives and many companies automatically ingest data from patent offices all over the world and peer review journals, they perform disambiguation to figure out who are the distinct individuals and the data, which are the distinct organizations and then they link the data together. Many researchers then analyze the data and companies rely upon these data to provide insights back to the US government.

However, we know that there are systematic quality issues across the different original data sources and that the issues become even greater if we don't take these problems into consideration during linkage and analysis. Further, the issues disproportionately impact populations that are already disadvantaged in STEM, for example, because there are no consistent identifiers as I said earlier across different data sources, figuring out things like the distinct individuals, frequently rely upon names but again the current state of the art performs poorly for many groups.

So we need methods upon which policy makers can rely when making decisions on important topics like, how do we encourage more diverse representation across all types of STEM careers? So I now want to make two points about accelerating AI in the public sector, inspired by my work with the R and D enterprise, but by no means constrained to AI in this specific context.

So first we need better methods to make transparent uncertainty and errors that are propagated within AI systems and that disproportionately impact certain populations. So record linkage is a well established field of statistics and recent research in the field takes to understand how errors in one linkage can impact downstream analysis or even additional linkages.

However, there aren't any kind of [inaudible 00:35:43] standard practices at this point for market linkage and machine learning offers methods for understanding biases and models for specific analysis. So one common method is to examine prediction in errors by subgroups or within natural language processing to look at linguistics associations with certain demographic groups. However, these are typically static analyses and aren't geared toward tracing through the different steps of automated systems concretely where biases are introduced and how they impact downstream tasks.

So the state of our tools to assess such biases by priority populations as such that we have tools but we still have a lot of work to refine them. And then the second point is that once we better develop these tools, policy questions remain. So assuming we develop tools to let us know things like, the linkage of

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

one data source with other existing sources contributes to disproportionate low estimate of scholarly impact for racial minorities, how do we determine what are acceptable thresholds? And at what point remediation within these AI systems must occur? These are policy and normative questions, not strictly mathematical ones. As a colleague aptly put these aren't problems that a new algorithm or kind of a need new neural network is going to solve. They require discussion and community consensus.

So, my six minutes are almost up and I'll end here by kind of reiterating my two key takeaways today. We need better tools to understand equity within dynamic AI systems. And two of these tools will entail additional policy discussions. So AI holds great promise for the public sector but only if we can develop appropriate tooling and common standards to ensure equity. Thank you.

Nick Hart: All right. Thank you Julia. Thank you Sean. We're now going to turn to two virtual speakers oh, I'm sorry Brandon. We're now going to turn to Sean, who is our first virtual speaker. Sean, go ahead.

Sean: Well, thank you so much and apologies for not being in there in person. We had a last minute request to meet with the mayor of New York City and I couldn't pass the opportunity because ultimately what we're looking to do at Disaster Tech and our core principle as a company, we are a public benefit company, we're actually a Certified B Corp, is to accelerate the adoption of advanced technologies including artificial intelligence, for public sector use to benefit the public like helping to serve those in the public.

Myself, I formally served at White House staff for National Security Council. I was director for incident management integration policy and going through the Louisiana floods of 2016. And in that, I heard there was a word document and a PDF, right? So we were doing back of the napkin analysis on the extent of the flooding problems in Louisiana and how do we engage with Congress on a supplemental bill of a billion, 2 billion, \$3 billion for community development block grant funding.

I came to the realization that nobody was innovating in a space. I could look down to the hall at the intelligence director or the defense director, at my colleagues. And they're using these remote sensing capabilities, AI, these data science tools that we weren't using in disaster management and emergency management. So I left government to found this company to resolve that challenge, the challenges that I experienced.

So we have a number of design principles at the company and the first is to co-design with the user and who we building this for. If we're not taking an agile software approach to bring the public sector, the public servants themselves into the design loop, then we're not going to make informed decisions or the

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

best informed decisions on the use of these technologies for their end use, right? They're the ones who are going to be taking advantage of the tools.

The second thing that we're doing is we're working with the academic industry. In fact, we co-founded the first national science foundation, AI Institute alongside Google Nvidia and IBM with about seven universities across the country anchored with the university of Oklahoma and I'm co-leading with IBM, the research to operations group where we are ensuring that the research for artificial intelligence and I should mention that the center of the AI Institute is researched for trustworthy AI, for climate weather and coast's lush oceanography applications.

In order to ensure that, that research again gets to the best usability is that we have the operators in the loop, the practitioners, the policy makers, the folks who are going to be the adopters of the technology once it moves through the research, development, testing and evaluation phases into operations ensuring that there's a transition and there's a path to commercialization.

We don't want to see billions of dollars of research in development funding, where it dies on the vine and gets locked up. And then the last thing I wanted to discuss today was the key point in that National Science Foundation AI Institute, is the concept of trustworthiness, right? Because I believe for AI applications to be used anywhere in public sector or in private sector as well, that it needs to be trustworthy, ethical, and transparent. And certainly as it talks... Previous speaker talked about things like equity but we have to make sure that these tools can be trusted and are highly ethical.

There is certain private companies who have made decisions, unethical decisions on the implementation of AI and when those AI tools have been advanced into the public sector for dual use capabilities, it might have been used say in military application, but then when it was adopted in the civilian sector for say community policing, those AI tools had adverse effects on targeting communities of color or vulnerable populations and so we can't allow these AI tools to be advanced into the public sector, being used by public servants where it has the adverse effect. It's not helping the public, it's hurting them, it's hurting their ability to serve.

And so the last thing I wanted to touch upon is technical literacy, right? One of the design principles at our company is also to increase risk literacy. Well, in order for government agencies to make procurement decisions, they have to understand what they're buying. If public servants don't have the upskilling, we went through this and we are continuing to go through this with cybersecurity professionals. We know there's a huge gap on cybersecurity, which is leaving us vulnerable, not just here at home but overseas with near peer adversaries that we're up against and one of the key milestones or underpinnings of advancing AI and the various legislation around it is advancing that adoption because of our near peer adversary threat.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

But again, if policy makers and public servants don't have the skillset to understand what they're buying, they're going to make bad decisions. And so we have to ensure that both with legislation as well as with appropriations that OMB and president's budget ensure that not just a chief data officer, but from the very entry point for entry, bring new talent into the government ensuring that they have the skills, they are ready for this fourth Industrial Revolution and ensuring that AI is not only benefiting the public but that the public servants can take high advantage of them. So I'll pause my comments there and thank you very much for the opportunity again, would love to be in person, but I appreciate the opportunity virtually.

Nick Hart: All right. Thank you very much, Sean. I'll turn it over now to Haniyeh Who is also joining virtually. All right, over to you.

Haniyeh: Thank you and thank you Data Coalition for putting this panel together. As other speakers have already mentioned in recent years, we've seen significant advancement in AI technologies which poses a great potential to improve the quality of our lives as well as providing the competitive edge the US needs in public sector.

But with such advancements, especially when it's used in government, it is crucial to ensure these systems are reliable, effective, build-on use ethically when it comes to gaining the public's trust. And this means implementing ethical and responsible AI frameworks for development and deployment of these systems to preserve Civil Liberty Rights as well as, as others mentioned ensuring fairness and also establishing a robust governance system with having accountability embedded into the process.

But we've seen such efforts, especially when it comes to ethics happening in Intelligence Community as well as department of defense, putting out principle frameworks for AI ethics and guidelines for their organizations. But when we want to incorporate these frameworks to the use cases that we have and engendering trust in the process, we need to understand and acknowledge that trust in AI system is a multidimensional problem.

The first element of this problem is people. There are many stakeholders involved in the process and it's important to understand what trust means for each stakeholders and what are their needs, other speakers touched on education, how we can leverage these stakeholders by providing educational support for them. So depending on their roles and responsibilities, they would be able to bring their elements of trust into the process.

In the second piece of this side is actually the whole process thinking about what are the risks that are associated with any AI system that you're working on, having your thorough risk assessments system in place to evaluate and understand the risks, understand the impacts that these AI systems have on individuals, groups, society or even environment. And we've seen this

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

undertaking substantial work, creating these risk frameworks for AI systems. So it's crucial to also include elements of governance. As I mentioned before, as part of these it was mentioned before and who is accessing the data, what's been worked on, any sort of governance processes that could be included as part of this assessment.

And the last piece that we need to understand is the technology itself, how we should be designing this technology to be able to support people, people who are building these systems, what are the tools that they need? Whether if it's explainable AI, whether if it's tools that they would be able to evaluate the data or evaluate the model's behavior to ensure it's equitable, to ensure the data is representative of the population that they're going to apply these AI system to. Or whether if it's around governance piece, how we should be monitoring these systems behavior over time, how we should be thinking about all those risk elements that we might have identified throughout the process, what technology can support how we can create tools to identify those risks, how we would be managing them, but if it's alerting a human so they would be able to intervene if AI is behaving in an unexpected way or tools that we would be able to evaluate a system and understand, oh, we have a concern here there might be evidence of bias and how we can possibly mediate that through the technology.

So these are the ways that we can think about how the technology can support the people and process aspect of it but at the same time we touched on education. The AI Initiative Act already mentioned and highlighted the support that's required for education piece of it. But thinking about this whole process of people, process and technology altogether, in order to be able to engender trust in the system, we need to work on all these three pieces together. And with that, I would like to thank the panel and answer any questions that you would have.

Nick Hart:

Well, thank you all so much for sharing your perspectives and insights. I'm going to admit in full disclosure, we're about 15 minutes behind our scheduled times and that's my fault. So we have about 10 questions that have come in for the panelists and we're going to commit to getting you all those questions but I'm going to ask you all just one simple question and an effort to get us back on schedule, which is that if each of you had to offer just one recommendation based on your respective remarks to government today, what would it be in the space of your respective comments and Julia before you answer, I'm going to say that one of my staff members said, "This is awesome" as you were offering your remarks so I apologize we didn't let you see your slides but we will make them available as a follow up item. So I'm going to give for the folks in person, give you the mic to go first and then we'll turn to the virtual speakers.

Julia:

Well I think again speaking from multiple perspectives but from someone who cares a great deal about making the best use of public data, which is a crying need, thinking about the ways of using AI to describe how data are used and for

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

what purpose and who are the experts so that you can build a community of replicable work around the data. So my recommendation, and I'm sorry that our colleague from GSA is not here but my recommendation would be that data.gov, incorporate automatically generated information about usage for every public dataset that is on data.gov and the machine learning models that were developed from the Kaggle competition and that are documented on Show Us the Data, are available and public use and we're happy and the whole group is happy to work with them, including Texas Advanced Computing Centers.

Nick Hart: All right, great idea. We'll pass that on don't worry. Brandon?

brandon: I think my recommendation would be that despite the incredible promise that AI offers for the public sector, I think we need to focus first on producing AI that does no harm, right? And that's actually not easy because of unintended consequences and so one of the ways that we can do that is develop, to work with multiple stakeholders as some of the other panelists have said to develop better tools to understand, not only what people expect from AI but also to monitor as it's implemented, how it's impacting the different potential stakeholders.

Nick Hart: All right. And we'll go to our virtual speakers next, Sean.

Sean: Well, thank you for the question so what I would say is, first of all, that these government agencies have to understand that machine learning is just the tip of the iceberg and everything underneath that water below the water level on the iceberg is data and that includes high performance computing.

We have these incredible labs, national labs, when I was at department of energy we used to fund many of them as well as universities who have these computing clusters with GPUs and CPU infrastructure that most people in the public have no access to, right? And so if you want us to be able to advance, this community to advance the use of artificial intelligence technologies for public benefit then we really have to democratize not just the data and the models, we have to democratize the use of HPC infrastructure, high performance computing infrastructure, or else we're going to be outsourcing it to Google, Amazon, AWS, and Azure which is great for those companies. But we've already invested in a lot of high performance computing infrastructure and I would like to see if I had the one wish, one ask is to provide better access equitable access to those already government funded resources that have been there for decades, so thank you.

Nick Hart: All right. I'm going to provide no commentary about the companies that you've just mentioned who may also be in the room. All right and to our final speaker, any recommendations?

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

- Haniyeh: Absolutely. I just want to echo what all the other panelists have already mentioned but in addition to that, I would say, as we say all of the panelists has already mentioned that there's a great potential here but at the same time we need to make sure for every work that we are doing, when it comes to AI, having a thorough understanding of its impact, being able to bring different stakeholders in the table and understand from each of those perspectives, what are the risk? What are the impacts that this system can have? And before putting these systems in production, putting it in deployment, being able to [inaudible 00:55:14] some of those risks that might appear and also have ways to manage those risks.
- Nick Hart: To all of our speakers in the first panel, thank you so much for sharing your insights. I want to go ahead and encourage everyone to give you a round of applause and invite our second panel to the stage. This is an all in person panel that includes Tasha Austin from Deloitte, Alexis Bonnell, from Google, Peggy Tsai from BigID and Mike Anderson from Informatica. Our first speaker will be Tasha Austin, Tasha take it away.
- Tasha Austin: Thank you so much and I had to check in on Alexis after the Google comment was made, she's an amazing leader at Google and soon as that comment was made, I was like I just want to let you know, we still need you today. Good morning. I'm Tasha Austin as Nick said, I'm a Principal in Deloitte Government & Public Services sector. I lead a lot of our work that we do in AI and data analytics and I also serve as the director of Deloitte's Artificial Intelligence Institute for government. I mean, I'll talk a little bit in a few minutes about some of the work that we're doing there, and also lead some of our national DEI strategic initiatives for Deloitte, which has been an interesting journey when we think about the past two years and the things that have happened in the communities, more of our racially and ethnically diverse communities, we've even been leveraging AI and thinking about how to use responsible AI to drive more equitable outcomes in some of our diverse communities so that's near and dear to my heart as well, but I am pleased to join the conversation today.
- The one thing about going behind amazing leaders that have already spoke is that you get to come back in and say the same thing that they said, or at least amplify some of what they said so please forgive me in advance because I think some of the things that Haniyeh and even Brandon focused on with talent and some of the same things that I would want to hit on but I would just say that certainly just sort of amplifies the need when we actually are all thinking alike about the things that are needed in this space.
- And so what I thought we would talk about is equity and how we need to be considering equity, in a development of AI as we look to accelerate that in a public sector and really I just want to talk about some of the things that we're doing at Deloitte to really push the conversation and amplify some of the things that we need to continue to think about on the topics of equity.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

And we really want to focus on how we can also act, be intentional about some of the things that we do based on how shared obligation of considering equity when we're developing these AI technologies. And so, as I mentioned, I serve as a director of our Deloitte Artificial Intelligence Institute for government and we think of that as an innovation hub and what we offer through our leadership on different areas that are top of mind for many of us, you've heard us talk about some of those topics a day, whether it's trustworthy AI, whether it's AI inequity, whether it's AI for good. We also talk about talent upskilling, reskilling. So we offer lots of perspectives, white papers, articles, conversations with clients and other stakeholders on those topics that matter.

We also talk about ethics and governance strategies and we also talk about specific use cases as I really love the first conversation that Haniyeh had with us about the library I think that is so much needed and we are even thinking about that in Deloitte as far as how we can build a use case library around topics that are important to our clients, whether it's mortgage lending and looking at those practices and how to use AI and going back and rethinking where does equity play into that? When we think about some of the practices that may not have been fair back in history, back in the day and how we're trying to change the dynamics around that today. We think about the talent and recruiting processes and how AI can accelerate that and present more opportunities to use AI in that space, but also how it could be harmful.

When we're thinking about recruiting, if AI is misused in the talent space, we also focus on how AI can help advance needs and support interventions that are needed in the public services space and so, we continue to offer perspectives from some of those use cases and many others. One of the other things that we talked about, and I think Haniyeh hit on this as well as, we're focused on building an ecosystem from bringing government together to industry, to Academia and we're doing a lot of stuff with Academia. There's such a huge opportunity there, I mean I'm going to spend a little bit of time talking about that. But we really think about the ecosystem as we are focused on advancing equity and inclusive principles more broadly and part of the dialogue that we continue to push is the need to ensure that the AI community reaches beyond technologists.

Other disciplines matter when we're talking about bringing diversity to the table to really solve problems with the use of AI. So whether that's my humanities majors, whether that was on my business majors, whether those was...

PART 2 OF 6 ENDS [01:00:04]

Tasha Austin: That's my humanities majors, whether those are my business majors, whether those are some of my other social sciences majors, we really need all of them at the table to really address equity and to consider equity as we're building out these technologies. So let me share a little bit about the work we're doing in academia. And Nick told me that there was an option to come in person or

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

hybrid. I'm an introvert. I would've rather stayed at home, but he knew I lived in DC, so there was really no excuse for me not to be here, but coming in person comes with so many opportunities because I've been able to network right away, coming in the door. Julia's already put me to work. Now, I'm trying to backtrack the introduction that I gave to Julia, because once I said some things of her, she's like, "I have a job for you."

But anyway, I am very thankful for that. And it really aligns to something that I'm about to tee up with some of the work that we're doing in academia. We are focused on bringing industry expertise into the classrooms. I'm a professor. I work in the community college space. I didn't tell Julia that because I figured she'd asked me to do more based on our first request. But some of the things that we're focused on, even as I go and lecture in other universities, was just recently doing some work with Georgetown and talking about bias mitigation and remediation strategies based on all the different biases that we need to think about when we're building AI products and solutions. We're also in the classrooms. Some of the schools, believe it or not, AI fluency is just not a thing yet. AI literacy is just not a thing yet. And so we go into the classrooms to teach the students that. Even in the academia space, you'd be surprised, but there's plenty of opportunity to continue to do that.

One of the other areas that we're focused on is advancing our research and development efforts in the areas of health equity, education equity. Academia's getting really smart in their thinking about bringing diverse perspectives together to advance those types of research opportunities. And recently, I have been working with UVA and a set of historically black colleges and universities, and Deloitte sponsored the capstone programs and some of the other programming that they did in this space, to do some research and advance it, really good research in the equity space. It could certainly be leveraged to inform how we serve our clients today in a public sector. And so the more we can get ahead of that and advance that and get into the classrooms and partner with academia, the better off we'll be when we think about how we accelerate AI in a public sector, but how we also have a equity mindset in doing that the right way.

And so that's some of the other work that we're doing. And we're also being really smart about taking academia to market with us. Some of the work that we're doing for our clients requires some really deep knowledge, and we have lots of amazing leaders, academic researchers in the academia space. And so we've been teaming with them and bringing them along as small business partners on some of our contracts to solve some of these problems that our clients are looking for us to address when it comes to leveraging AI in a public sector.

One of the other areas that we have been focused on when we think about future talent is going back into the communities. You hit on STEM talent so beautifully, Brandon, and I would just say there are communities that just don't

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

get exposure to a lot of the emerging technologies today. And so Deloitte has been building partnerships with nonprofit organizations and figuring out how do we leverage the relationships we have... How do we leverage all of the knowledge that we have and the technologies we have built and bring those back into the grade schools, particularly in those areas where they're under resourced in their community, and they just would never get exposure to these technologies? But that's important when we think about the needs to have diverse talent around the table as we're building future technologies. And I think this is just critical step that we all must take to realize broader equity.

And so, as I just wrap up, I will just say we can be intentional and even more intentional about equity considerations in AI development. And everything I'll mention, you've already heard. You've heard me talk about the need for having talent around the table, from making sure that we have a greater diversity of people in the ideation phase to even a continuous monitoring phase after we've launched or implemented an AI technology. Data diversity, you've already heard people talk about the need for better representation of data and how we have to be mindful of that as we're looking to solve problems. Ethics and governance, I don't need to say much more about it. I think everybody's hit on that so far in their conversations, and trustworthy AI continue needs to be top of mind. And then just the fact of implementation of the legislation that continues to come out. We connect with congressional stakeholders. We talk about some of the policies that they're shaping. We offer perspectives in that space as well, but we also have to help our clients understand how to implement those technologies as they're looking to... those technologies, as well as prioritizing areas like accountability and transparency as they're building out some of the development across their organization.

We also just put emphasis, and when we're talking with our clients, on aligning AI with social values, such as fairness, despite some of the possible trade-offs for efficiency and profitability, and I will just say that I would love to continue this conversation with all of you. What Deloitte is doing in the next month is hosting our first annual AI symposium for government later on in April. Thanks to The Data Foundation and many other Alliance partners, Google, Nvidia, Snowflake who will also be sponsoring and co-facilitating a lot of the sessions that we'll have in that symposium. We're really looking to bring together a group of senior executives and visionaries across government and industry at all stages of AI adoption to really push conversations forward in the advancement and the use of AI in a public sector, while also considering equity in mind. So those are my remarks. Thank you.

Nick: It's your call.

Tasha Austin: I told Nick I had to leave, but-

Nick: So we're behind and I know Tasha has a hard stop, but I'll also just acknowledge, Tasha is a member of The Data Coalition's advisory board. So while you're

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

hearing a lot from Data Coalition members on this panel, Tasha, thanks so much for your service also to The Data Coalition. Let's turn next to Alexis.

Alexis Bonnell:

So like many others, this is my first time out in the wild. And so to honor that, I'm wearing the shoes that you see me. It's not just a googly thing. It's a celebration of being with you all. But part of what I wanted to share today, there's so many great data insights, but ultimately, data is about people. And so I wanted to share a little bit about what it is that we're learning around data leadership. And this comes from a very personal place for me. Before coming to Google, I was one of the federal chief innovation officers. And if you know anything about really great innovation, you know that innovation without evidence is meaningless. And you know that you can't have evidence without data. So I was a data champion without even knowing that I was a data champion, but one of the things I think that's really important to that is taking that a step farther.

So if you have data, and data drives evidence, evidence drives insight, and this is the really critical one. Insight drives the rationale for change. And we've been through a lot of change. So one of the things I think that's really important for us to understand is what does it mean for leaders in a data-driven age in how they have to lead differently and the idea that change may come more frequently because of that volume of data, because of that insight? And the other thing I think we confront in the public sector... And I'll speak for myself as a public servant previously... and that is our preparedness for the rate of change. I think everyone in here has some type of digital transformation or data transformation or organizational transformation program underway right now, and that's because there are a lot of things that we need to change. A lot of things we're being prompted by, but I want us to think about the idea of our comfortability with the rate of change is directly equal to the flexibility that's required in our leadership and culture.

For many of you, you may have been that person who brought that piece of data that was pushing change and you came up against elements of your agency that wasn't ready for that, wasn't ready for what that meant. And I think we have to be eyes wide open in realizing that we have to navigate culture and change as data starts to give us different insight. And so, ultimately, what that means is data drives change, and more data requires more allies. And AI is one of those really critical allies for us to be able to navigate this amount of data. So whether that is cyber, whether that is spam, whether that is navigating curbside, we all know that we're already experiencing AI as an ally. And we talked a lot, and we'll continue to talk today, about intentionality.

But what I actually want to share with you is one of the coolest parts of my job is that I get to now serve public servants and go into the rooms when they want to have the conversation around "We want to reimagine the DMV. We want to reimagine medical education. We want to reimagine our mission in a data age, in a digital age and what that looks like." And one of the things that I want to

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

share with you today is, what are the traits and behaviors of those leaders who are really leaning in to data fluency, to change driven by data? And one of the ways I can summarize it is that it's really seeing this appetite to move away from a recognition of what might be a time capsule culture, looking at our organizations or even our data in this moment of time, and instead, being really intentional about how we create a change engine culture, how we accept the fact that evolution is part of our leadership journey.

And so I want to highlight some of the things that I have noticed and recognize and really respected about some of the public servant leaders that I think are leaning forward. I think, number one, is that they're recognizing that really the role of public servant is changing. And what I mean by that is that every public servant is now an information steward. If you think about it, no matter what your job is in government, your job function is to take in information, make the best decision you can, apply that to the resources you assign, and then take in more information to figure out if you did that as effectively as possible. And so if you think about it, every role is a data information steward role. So I think it's important that we think about that in a cultural sense. It's not just the data scientist job. It's not just this job. We all now have a different relationship with information.

The second is these leaders' ability to expect change and plan for it. Again, looking to build in flexibilities and adaptabilities because they know that flow of data that they didn't have six months ago that they might have in a year from now, especially on things like social equity and others, is going to make a game-changing difference in how they leverage their program. A third, and this is a really critical cultural one, is the ability to develop comfort in ambiguity while seeking insight. And what I mean by that is the ability to take a decision with data you're confronted, even if it doesn't feel perfect. And if you think about the last two years having to navigate change, we were talking about real-time policy in response to COVID and other things. So this ability to, again, gain confidence.

The fifth was the idea of seeking evolution, but not change for change sake. So sometimes people look at change, transformation, et cetera, and say, "Okay, this is just another round." But when we think about AI, we've talked a lot about the idea of how are you intentional about that? And thinking about governance, you may have a very different governance structure if what you're trying to do with your AI is to figure out whether a screw is loose on a windmill versus whether or not someone should have access to particular benefits. So being intentional and really thinking through that.

And I only have a few more. The other one that I think is really critical... And I think a lot of times, as leaders... I know for myself... I thought I did this well. And I realize now maybe I didn't do it as well as I would like, but that is defaulting to information transparency and openness. And one of the things that's been really interesting is how many organizations have come to us and said, "How do

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

we, in essence, do google it for our inside information so that our internal teams can actually find information easier?"

One of the last ones, and I think one of the most important when we think about culture, is to value and foster curiosity, making it easy for people all across the organization to find data, to look at it, and to be curious is ultimately what drives change. When we think about investment, one of the things these leaders are thinking about is the fact that they're making decisions right now that have generational impact, and they do not want to be the person who created the next legacy system. They also don't want to be the person who sunk \$500 million in a platform that wasn't carbon neutral. There are real generational considerations about the decision we're making around data and how they set up our organizations for the future.

One of the final ones is they seem to have recalibrated their relationship with risk, and I think this is actually really important. The more information, the more speed we're navigating, the more we have to understand and be intentional about risk. And what's interesting, oftentimes, in large organizations, risk equals threat. Risk equals threat, balanced with opportunity. And so that ability to be able to navigate one.

And the final one that I would say when it comes to the AI story is that, for many of us, it's talking about going to the cloud, getting some of these basics. And I want to, hopefully, inspire us that getting to the cloud, having that flexibility, having that volume of scale, that's the starting point, not the end point. And the final thing that I would say is, ultimately, what we've really learned, and what I've learned from these public leaders and we have to remember about AI, is AI is a learning technology. And that means the day that you launch it, it should not be the same a year from now. It shouldn't be answering the exact same questions a year from now. It should be giving us insight. It should be that ally in curiosity. So I want to thank all of you. Many of you I've gotten to work with before, but most importantly, the public servants who are teaching us about what great AI and great data looks like and giving us confidence in the future of what our government can do. Thanks.

Nick: All right, Alexis, thanks so much. Next up, Mike Anderson.

Michael Anderson...: Nick, I know you're serious about the six minutes, so I prepared for this like I'm testifying before Congress. Yep, so I have a written statement to move forward here. So, good morning, and thank you for the opportunity to provide public comments today. The data protection and privacy topic is certainly an important one in today's data-driven world. And we can take a discussion on it in several different in directions. I'll spend my six minutes this morning addressing how data used to train AI can be better protected using modern and automated technologies. I'll proceed on the assumption that all of us here already understand the importance of data privacy and protection in general. With existing private regulations, such as a GDPR, CCPR, HIPAA, and so on, the failure

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

of organizations to comply with data privacy laws can be extremely costly due to fines and really, if you think about it, the loss of reputation as well for your organization. For public sector institutions specifically, more citizen private data losses, on top of those that already fill the media headlines, will not improve constituent trust in their governmental institutions. Agencies already face an uphill battle on that front.

So what can we do? Well, sensitive data used for AI or any analytics capabilities, streams from a myriad of devices and sources and is being used primarily within cloud data lakes due to the efficiency and scaling capability of cloud environments for algorithm AI training. Lots of personal data is included in this streaming to the cloud. However, with increased data volume and use comes increased risk from insider neglect of abuse and exposure from adversarial attacks.

Masking data during the prep process to feed AI algorithms is an ideal approach for a few reasons. In most cases, you simply mask out the sensitive or private data that isn't required for the AI training. When you do mask data that's sensitive, it maintains context and referential integrity. As masking is data centric, it's portable. So whether you run analytics on-prem or in the cloud, that protection can travel with the data. In Informatica's history of working with data masking on AI or analytics' use cases, masking solves data privacy issues in most cases. However, in more complex environments, a reliable data governance framework for building and maintaining privacy trust is often required. A data privacy governance solution helps organizations improve privacy readiness, meet compliancy goals, and makes data safer for use by addressing data center requirements of today's AI initiatives.

This can be accomplished through a consistent framework approach that defines regulated data and policies and aligns across people, processes, and systems, locates, classifies, understands personal and sensitive data across the organization, links employee and other user identities to sensitive data, analyzes privacy risk and determines remediation priorities, protects personal data and centrally manages privacy consent and rights to comply with privacy regulations and policy standards. And the framework should demonstrate compliance with regulations and appropriate use policies through on-demand reports and visualizations and with continuous risk analysis.

To accomplish this framework approach, identifying critical data privacy governance capabilities is helpful. So if you're looking to improve how data used to train AI can be better protected, the following six capabilities should be on your requirements list. Let me run through them. Number one, defining and managing data governance policies. The right tools should define, document, and measure business and technology policies and responsibilities, processes, and data terms. With the governance task framework and visual workflows, organizations can identify business user stakeholders and align them with the data and processes they own. Number two, discovering, classifying, and

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

understanding personal and sensitive data with embedded AI. This enables an enterprise-wide view and automated analysis of personal data assets and metadata. Organizations must be able to rapidly locate, classify, and understand their data environments across multi-cloud, relational, and file storage systems for both structured and unstructured data.

Number three, mapping identities to personal and sensitive data. Building a data subject registry and linking it to personal and sensitive data enables organizations to quickly access data that belongs to identities, such as patient, citizen, employee, et cetera. Number four, analyzing data risk and establishing protection plans. The ability to analyze personal data risk and provide customizable risk impact models is essential to helping organizations prioritize remediation and effectively take action. Risk to private and sensitive data should be continuously measured and recorded to provide key risk indicators for privacy and protection programs supporting AI initiatives. Five, linking and managing identities to control data access. AI initiative practitioners must also be able to discover what sensitive data is held for subjects. And this is possible with embedded software intelligence in a tool to automatically discover associated PII across an environment.

Finally, number six, many disciplines utilizing AI and automation to enforce data rights with consent management is essential. By utilizing a collaborative environment to centrally manage data privacy preferences and consent to enforce data rights, a master data management capability enables consent management through workflow processes, the creation of a single view of the data subject, provisioning of record history, and enabling audit for data privacy verification support. Trust my six-minute overview is helpful to those working on AI initiatives. Informatica supports 5,000 global customers, and many use the capabilities I discuss today. Again, Mike Anderson, chief strategist with Informatica. Thank you for your time this morning. Nick.

Nick: All right, next up, Peggy.

Peggy Tsai: Good morning. Hello, everyone. For the past two years, I've done webinars talking to my Zoom camera, so it's really nice to be able to expand my peripheral view. So my name is Peggy Tsai. I'm the chief data officer at BigID. Today, I'll be talking about data protection and privacy and how artificial intelligence can be used to discover, catalog, and classify personal data. So like many of my panelists today talked about artificial intelligence, you have to trust the data in order to trust your artificial intelligence results. And it all starts with good data governance practices. For the past 18 years, I worked in the financial services industry, where I helped design and operationalized data governance programs. And in my last two years at BigID, I've spoken to many chief data officers who are responsible for privacy operations, which means that they need to leverage their existing data governance programs to find sensitive data and ensure privacy compliance.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Today, my customers are global companies in financial services, telecom, retail, and they own dispersed data ecosystems, where their data is siloed and fragmented and the size of their data is in petabytes. And just for some context, the Library of Congress has about 20 petabytes of data. So we're talking about large scales of data, and the complexity really requires a lot of new technologies to help with the monitoring of real-time data, as well as unstructured data in forms of emails, paper documents, and other information that's shared in your SharePoint.

In recent years, data in motion and unstructured data are really two untapped types of data that's traditionally ignored when it comes to data management. And that's because identification of personal information and personally-identifiable information cannot be governed through traditional technologies. Now, with AI and machine learning, we can scale it up really quickly and govern data that's previously been too known, too large, and unknown, what I like to call dark data. For those unfamiliar with my company, BigID, our technology focus on data discovery to find, connect, and understand disjointed data sets often at large data scales. Imagine being able to automatically create an understanding that looks like a spider web of data connections based on a person or topic, being able to find common data usages, identify outdated data sources, and really understand the interdependencies of data in the data value chain. These insights and recommendations for data is where AI is really powerful and really useful for today.

We anticipate the usage of AI machine learning to automatically build and support the inventorying, impact analysis, and reporting personal data and metadata for the government as well. In contrast, without using AI, and I have done this in the past, I've seen the efforts be really a manual exercise that's time-consuming to complete and really highly dependent on manual resources to tag and identify that metadata. We all know that data changes really quickly and often without any notice. Therefore, it's even more imperative that we use artificial intelligence to help monitor and automate the data discovery that is critical to the foundation of any successful privacy compliance program.

Let me elaborate a little bit more on the transformative impact that AI and machine learning can have on a privacy governance framework. First, BigID's data discovery capabilities, accurately categorizes data, including datasets and documents using patterns, natural language processing, and deep learning to label the data for proper usage amongst data consumers and partners. This supports building out a foundation that pulls together and documents all the information that is available and can be data shared. Secondly, by using the latest data science techniques, classification has allowed BigID to quickly label the data based on its metadata and attributes. BigID has a custom-built, global classifiers to tag, label, and understand all the databases and documents. That means that if there's sensitive information that's saved on a file or stored in a database or pushed through your data pipelines, that can be alerted as personal

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

information and then policies and proper workflows can be automatically initiated.

The third component of data discovery is correlation. This is one of BigID's patented discovery methods to connect similar and related data across different data ecosystems tied to a person or entity. It is very commonly used when it comes to data subject access request or DESARs. Correlation has proven useful for our customers to automatically detect information tied to a person and does not require pre-knowledge of where that actual data is. The fourth technique is actually called clustering, and that really helps with understanding duplications amongst data, especially when it comes to unstructured and structured. And it helps with remediation efforts when identifying multiple copies of data.

So many of our customers ask us about privacy compliance for different state and in their industry, different regulations. And as a chief privacy officer, I recommend building out a program that is scalable and reasonable for all upcoming and expansive regulations, which means you cannot hard code rules to find your data. You need to use AI to learn and discover your personal data faster and give you the insights and recommendations to take action on a much larger scale. I want to be clear that artificial intelligence is not a magic bullet. It needs process and people, but we're seeing it used more effectively to accelerate and scale up privacy governance programs faster and more efficiently than ever before. BigID is part of that data strategy amongst organizations to build and execute on that data governance. We strongly advocate for automated data discovery using multiple discovery techniques to build that strong data foundation. So, with that said, I'd like to thank The Data Coalition and Data Foundation for putting this event together and having me here today.

Nick: Well, I will say, I really appreciate the emphasis across your collective remarks about people because it is probably not said enough when we talk about artificial intelligence. Having said that, out of respect of all of your time, I'm going to continue to try to get us back on schedule. So I would like to also ask this panel a single question for a quick-hit response. And we will commit to also getting you all the questions that have come in from the audience, so you can see and maybe have some continued dialogue. So, in that spirit, let's do a hypothetical. You're over in the West Wing, you've run into the president on the elevator, and you have until you get from the bottom floor of the West Wing to the top floor. I don't know why that's the ride, but let's just say it is. It's a short elevator ride. What's your recommendation to the president about where you're going to take AI, what you need to make the field more successful in, let's say, the next year? I'm going to pass you the mic. There's a mic down there. Let's start with Tasha, and we'll go down the line.

Tasha Austin: Hold on. Can you hear me? I definitely told Nick not to give me a hard question. I've been working on a proposal until 2 o'clock this morning, and he still did that. I think I would just go back to what I said about the need for more resources in some of the underrepresented communities when we think about

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

diversification and a talent that's needed around the tables to appropriately build these models. I think I would say to our leader that we need to be focused on pivoting those resource

PART 3 OF 6 ENDS [01:30:04]

Speaker 1: A leader that we need to be focused on pivoting those resources into those communities and technologies in those communities so that we don't feel the pressure in industry in another spaces. Obviously we're going to help deliver and bring those things into the classrooms when we think about the digital divide, but obviously we also want to look to see where can additional funding and resources come from to help close that gap in the classrooms. And we still have a lot of schools that are struggling in that area.

Nick Hart: That's great. Alexis.

Alexis: This is a hard one. I think ultimately it would come down for me to the idea of really this is an absolutely amazing moment in history where we are putting unprecedented resources to really thinking about who we want to be in the future. And so I think I would look for ways to say, how can we use this moment to be intentional about the organization, the government, et cetera, that we want to be and make sure that the data and the technology are enabling that. I think the second thing I would do is find a way to make sure that no conversation that's happening in government right now that gets stopped because someone says, but what about privacy? But what about security? There are answers and you've heard a ton of different ways to solve that now. And so I think finding a way for us to drive into the curiosity and the application of the solutions to those questions, instead of just asking those questions would be really awesome.

Speaker 2: I'd like to really echo a lot of what Alexis has said about data and the fact that there's so much potential now in the availability of data and it's really being able to harness all that. And I know with the Federal Data Strategy, it's really about executing on that and putting the right mindset and leaders in place that know how to do data management and data governance and focusing on the really potential and just being able to start using it, testing it out and really leveraging the benefits that I've seen AI has in the private sector.

Speaker 3: Nick, I'm glad you gave me the opportunity to practice what I'm going to be saying to the president this afternoon when I'm in the West Wing. So just really quickly this is what I would say. I would say, the framework of people, processes, and technology applies here. You've got to get the right people in place in government service that can make a difference, especially at the leadership level, or you're not going to be able to build a data driven organization and government. Number one, processes flow from number one, three technology, for the most part it's ready to go. But you got to get the people and processes in place, the people's the hardest part. But I encourage

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

any government people here today are listening, the technology's ready to go. Just execute, take the steps now to move forward. Thanks.

Nick Hart: Well, this genuinely that I could ask you all questions all day long, but we do have to keep going in our program. So thank you all very much for sharing your insights and perspectives. For a programming note, we are technically going to skip our planned break. If you're in person or in the audience virtually and you need to take a break, I encourage you to do so. But we are going to continue rolling right now into our planned panel number three. This is an all virtual panel and I would like to welcome to our virtual stage all four of the speakers, Alex Howard, Darren Menachemson, Darren Wray, and Tori Kim, and we'll take each of you in that order. So first up, I'd like to welcome Alex Howard. Alex, go ahead.

Alex Howard: Okay. Can you hear me now? Rock. So thank you so much to Nick and to everyone at the Data Coalition and the Data Foundation for host team this public forum. As you may know, today is National Freedom Information Day chosen in honor of President James Madison's birthday, commonly hailed as the father of open government in the United States. The idea that in order for a public to be self-governing, we need to have information about our government in order to hold it accountable. And it's wonderful that this forum is happening in Sunshine Week when we celebrate the public's right to know, because that gives us a wonderful opportunity to talk about how government transparency and accountability and disclosure can accelerate artificial intelligence while protecting privacy, security, and human rights as I think we've heard about already this morning.

As you know, once law becomes encoded in technology, code becomes law. Code itself can govern how people live, work, play, how we know and access information. And it's important I think that we talk about accelerating artificial intelligence in the public sector with a larger context of the world that we live in. It cannot come at the expense of human rights, civil liberties, the public's right to know all of which are central to democratic societies. We know that artificial intelligence is going to be part of our everyday life going forward across places, but public sector algorithms have a special importance because people don't have a choice, but to use them. They don't have a choice, but to be governed by them. This is going to be increasingly true in the private sector as well.

And for that reason, we need to make sure that from making unemployment decisions, to getting loans, to parole hearings, to education, to work, the code that's going to govern how we live, work, play, and learn is open to us or that it's at least auditable to our representatives and to regulators. They must be auditable to ensure that existing inequalities and injustice is not codified in biased data that is collected and put out in a rush to modernize, that in a rush to capitalize in the digital gold mine. And we all know the height that has surrounded big data now for over a decade, I used to be part of O'Reilly Media.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

And the strategy of getting to market quickly by getting as much data as you can has been executed upon in the private sector.

We all know that the public sector has tremendous amounts of data that can be used to train algorithms, that can be used to target systems, that can be used to predict and anticipate problems and get out ahead of them. But if we don't create systems for democratic accountability from the beginning with openness as the default, I think we're going to run into trouble. And I think we can take inspiration from other places on this. Open data and open source code can reveal and check algorithmic bias and racial gender or religious discrimination in public services, public accommodations, and public access to information. And over the last five years, we've seen other nations enact laws and regulations that focus on these issues, from France to the Netherlands, to New Zealand, to the United Kingdom.

In France particularly, the Digital Republic law mandates transparency of government used algorithms. Public agencies are required to publicly list any algorithmic tools they use and publish their rules. Now with the Data Coalition and Foundation know all about data inventory is they know that data.gov is supposed to work on this. But imagine if Congress ordered federal agencies to do the same thing they do for data inventories, but with code inventories and algorithms at code.gov. Imagine the Office of Management and Budget now with a confirmed director pushing this issue. Imagine an explicit extension of the Freedom of Information Act to code and metadata with an expectation of proactive disclosure.

Imagine there being an adversarial opportunity for us to get code out of the public sector that's being used to make decisions that's not currently accessible to the public. Imagine investments in the human and technical capacity of the SCC, FEC and FTC to audit AI in the public sector in the places where you have new actors that are shaping the civic discourse of this moment, but are leaving people behind offline. Imagine for a moment a democratic vision for artificial intelligence in the public sector that centers on human rights, not profit, that centers on the needs of the public to know in order to be self-governing, instead of the authoritarian models we've seen from other countries that are focused on coercion, control, secrecy, opacity, and secrecy that elevate the corporate needs of different entities and societies over the public's needs.

As you may know, the Federal Data Strategy was the number one commitment in the fourth national action plan for the open government partnership. I'd ask you, how many of you have heard of it? Please ask your colleagues in government when the General Services Administration and the White House will begin co-creating the fifth national action plan for open government? And that this is important to you, push for commitments on artificial intelligence and democracy, push for algorithmic transparency and accountability, push for audits, push for the recognition that in a democracy we need to be able to know

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

what our government is doing on our behalf, and that includes the digital sphere as well.

So I'd ask you in this Sunshine Week, please commit to pushing our government of by and for the people to collaborate with the people in developing legislation and rules that govern the use of AI in the public sector and create expectations for it not to be misused in the private sector, codifying our bill of rights into the technologies that we develop and use every day. If we don't do this, we risk seeing our democracy erode further as OPEC algorithms and unaccountable private actors move quickly to create a new standard of living and work that does not put democracy at its heart that does not center on human rights or civil liberties, but instead puts profits above patriotism.

I thank you for this opportunity to talk with you again. I wish I could be there in person and I look forward to what everyone else has to say.

Nick Hart: I know Alex is also running an event of his own in just minutes. So thanks for joining us, Alex. I'd like to next welcome Darren from ThinkPlace.

Darren Menachem...: Thank you. I'm Australian and one of my country's most significant contributions to global regulatory practice is a model that's known as responsive regulation, and it was developed towards the end of the 20th century. And that since its development, it's been adopted across Australian government and many OECD countries around the world. So for those of you who aren't regulatory wanks and I assume that there are a few of you here today, responsive regulation says that government needs proportionate responses to non-compliance. So if someone's trying to do the right thing or lawful thing but isn't quite managing it for one reason or another, you don't immediately break out the lash and start flying. Instead you think about how you can use softer tools like education or maybe it's a more intensive support or incentives, or maybe even some certain words of caution to help and motivate people to move towards more compliant behavior.

Likewise, if a person's willfully trying to break regulatory rules, when where softer tools haven't worked, you escalate your response moving to harder tools like demand letters or financial penalties, and you trying to move people back towards compliance. So this thinking was born in efforts to drive better compliance with regulation and law, but it's also been applied in service delivery. So if a person's showing signs of need, signs of vulnerability, government agencies will usually have a way of identifying this and then provide increasing levels of service and support proportionate to that person's need, whether it's helping people with a mix of social disadvantages who lose employment to get them back into work or supporting people at risk of serious chronic disease to help them develop better health behaviors.

So when responsive regulation and its younger sibling, responsive servicing came onto the scene, data exploitation was in its infancy, data [inaudible]

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

01:42:24] was a term that hadn't been coined yet and analytics were pretty basic. Artificial intelligence was the stuff Hollywood, but things have changed though, haven't they? So in the last decade or so governments have started to fully appreciate the value of the data assets and the value of joining up data to get a better understanding of citizens or customers or whatever we call people in our particular neck of the government woods. Analytics have become more and more sophisticated, genuine AI, ML projects are starting to roll out in policy domains from law enforcement to social service, to taxation, to education. So it's everywhere, it's becoming more prevalent. And with the right data and the right analytics, we can make powerful guesses about people based on their demographics, their behaviors, their history and intervene in some way potentially before risk eventuate. So if we don't manage to nip the risk in the bud to optimize its response to tailor to that person.

And we can do that in a way that has unimaginable precision or would've been unimaginable only a few years ago. So this isn't a bad thing, in fact, the potential to use technology to achieve humane policy outcomes is huge. It's not necessarily a bad thing, but it could be. There's been a huge investment in data and AI or analytical technology in general, but the investment in ethical capability and ethical capacity inside government in tackling big questions about how we should and how we shouldn't use this technology often haven't kept pace. There's this architecture of ethical risk and government needs to confront it. It's in areas like automated decision making in areas like inclusion, data right and protections, and in unintended consequences when things become more automated. And in the future where humans might play a smaller day to day role in regulatory governance and service delivery, that these risks are really heightened.

So think of things like poor design of algorithmic logic resulting in errors, or just the wrong application of law or policy or administrative rules. Think about virgin introduction of bias because of how data has been collected over time, potentially over many years, or just how it's interpreted when artificial intelligence is being trained. And this can result in negative discrimination just plain on unfairness in decisions and intervention, and not just negative discrimination, automated positive discrimination can have pretty dire consequences as well. If government is predicatively trying to help somebody, but in doing so closes out other options that might have been better for their case or makes a person feel stigmatized, that can also cause genuine harm.

We have risks around explainability, we have risks around unfairness in how people are being serviced, just because of barriers to access, whether it's their ability or affordability, or maybe they're just underrepresented in the data holdings of government and therefore they're being excluded from the algorithm. So there's a longer list of these and I don't have time to cover, but obviously that list would give a lot of you a heebie-jeebies and rightfully so. And inoculating government against these types of risk, I think is something that we need to spend a lot of time and focus on. It means thinking quite differently

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

about what it means to be a government, to operate in a world where big data and AI are now intrinsically part of how things work. If agencies are rolling out systems and doing so in the absence of the right types of ethics monitoring say, it may be a month or even years before a severe ethics issue is identified.

So inadvertently those types of issues can be perpetuated. And we've seen how poor algorithmic decision making and poor analytics has already dramatically affected the lives of people in the last few years in profoundly life changing ways. We've seen case studies of financial ruin, case studies of suicide, case studies of entrenchment of disadvantage, and no doubt we're going to see many more over the next couple of years. And maybe the thing that worries me most is that the role of digital ethicists is not a common role in government. We are trying to push these very powerful technologies through trying to transform government, but we don't necessarily have the ethical or cloud in size to keep us from creating the steps. So the question I posed to people today is, have we created the right ethical capability needed to deliver in the promise of AI in a way that is just fair, safe, inclusive, and dignified in such a profoundly complex space? Thank you very much for giving me a bit of your time and your reckoning today.

Nick Hart: All right. Thank you, Darren. Our next speaker will be Tori Kim from eBrevia. Tori, over to you.

Tori Kim: Thanks so much. Good morning, everyone. My name is Tori Kim and I am a Corporate Attorney and Project Manager with eBrevia, which is DFIN contract analytics solution. I joined the DFIN team about two years ago prior to which I was a practicing attorney at a New York law firm, where I gained a lot of perspective into the process of reviewing large volumes of documents and all of the pain points that come along with that. I spent many hours pouring over contracts in the context of due diligence and searching through hundreds of different pages to try to find certain key obligations within those documents. And my frustrations with this pretty inefficient manual contract review process are really what led me to have a curiosity about technology in this space, and trying to understand the ways in which AI and automation could really potentially help to improve this process for any reviewers of contracts or other documents.

As I went through this journey and this exploration, I started to feel like AI was a bit of a black box, and coming from industry like the legal industry, just pretty traditional, I had some hesitations about what AI could do for me. The first hesitation that I had was whether I could actually trust the output that an algorithm was generating for me. Would the AI be accurate? And how would I be able to ensure that the data the system was pulling out of my documents was actually the information that I was interested in? My second hesitation was surrounding the complexity of AI. As someone who is completely non-technical, I wasn't so sure that I had the background to really understand, let alone train AI

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

solution to do the things that I needed it to do. On top of that, didn't AI require a big investment of time and a background in data science to be able to use?

In thinking about these different hesitations, I also reflected on how I might be able to settle these anxieties that I was having. What would make AI feel accessible to people like me, non-technical users who have certain subject matter expertise in an area? With respect to my concern on the trustworthiness of AI, one thing that I think that could really help is if I was able to actively participate in the process of training the AI. If I could somehow take my domain expertise and imbue that into an AI model, be able to tailor the parameters of what the AI was doing to my specific needs, that would allow me to better ensure that the results aligned well with my expectations. Another thing that might help is if I was able to track a particular model from a bunch of different angles. So having visibility into the data that I had used to train that model, the flexibility to make changes to that training data in order to improve or tweak that model in different ways. And also being able to monitor things like accuracy metrics at every stage of the process.

With respect to my concerns on complexity, the thing that I think would give me the most comfort is having any engagement with AI being as user friendly as possible. As a non-technical person, words like machine learning are a little bit scary and can be intimidating. So to the extent that a UI design would be able to make that process more digestible and easier for me to navigate, allowing me to train an AI model without having a data scientist on hand to help me would be a really huge value for me. Also to the extent that the process of actually training the model was relatively straightforward and quick to do would also give me confidence in the ease of the process, as well as its scalability. I'm excited to be a part of eBrevia because it addresses these barriers to adoption in all of the ways that I just talked about. And in doing so, it makes AI accessible for everyone.

At a high level, what eBrevia does is takes AI to parse through documents and extract key data from them. But importantly, what it provides as an opportunity for reviewers to do is really sit in the driver's seat and be able to customize an AI's learning based on their own knowledge without any technical expertise required. So they're given that control without all of the complexity. With an eBrevia, you're able to train a model with clicks rather than with code and reviewers can go in and consistently iterate and improve upon those models with access to all of the training data and all of the accuracy metrics that they need in order to succeed.

On top of all of this, the system learns very quickly and it's completely document agnostic, so it can span across use cases. Within the public sector, this can manifest in a bunch of different applications, whether that is reviewing contracts for compliance with a certain government regulation or taking a look at a bunch of loan documentation to improve the supervision of bank by regulators. Ultimately by using eBrevia, reviewers are able to solve for those

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

core inefficiencies that I encountered when I was reviewing a lot of contracts. And it allows the AI and humans to collaborate in a way that enables contract review to be a lot more efficient and effective.

Thank you so much. Look forward to your questions.

Nick Hart: Well, Darren, Tori, Alex, thanks so much for your comments. We have quite a few questions that have come in from the audience, and I'll maybe just open up by asking one to all of you, because this issue of auditability is cut across as a theme, the reliability of some of the information that's coming in. And I guess I'm wondering if you have any, and this is to all three of you, recommendations for those who are in government as we're thinking about this concept of auditability. And Alex, you've referenced it first, so we'll start with you. Are there obvious things that we should be recommending to government agencies as they're exploring AI applications when it comes to this concept? What should they do, auditability? Alex, over to you.

Alex Howard: I quite like open by default. The work of the Data Foundation and Data Coalition into making federal spending open by default, I think is a great example. You can go to USAspending.gov right now and see what we're spending money on. You can go to data.gov and see lots of data about other things. There's no reason that we shouldn't be able to go to code.gov and see the public sector algorithms that can be open be open. Now, I think there's a really nuanced conversation that's going on around the world and we shouldn't divorce ourselves from it. This is something that France has been dealing with now since 2016 when it enacted its law. I'd say, go talk to colleagues there, go see how difficult it is. Go look at New York City's efforts to create auditability in public sector algorithms and create a registry of them.

This is not something that's brand new. And I think that we need to be careful about it in the sense that it's easy to see how algorithms could be gained if they're made fully transparent, in certain contexts that could be problematic. But if you see, for instance, recidivism algorithm used to deny people parole, and it's doing so in a way that is not racially just, then I think it makes sense to suspend it and to call for audits by civil rights entities and government. And we're going to see that happen again and again and again, and I think that the thing to do is to look for openness by default, and then to empower regulators to look inside of the black box without demanding it be unspooled. Asking Google to put out its algorithm is probably not going to work, but requiring them to let the FTC in to see how it's actually showing people information or services I think could be quite relevant.

Nick Hart: Interesting. All right, Darren, any thoughts on this question?

Darren Menachem...: Yeah. Just to agree with what Alexander has said. I want to just focus on a particular part of the audited regime, which I think sometimes goes underserved, which is looking at ways you can understand what the human

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

experience of being regulated by or serviced by an AI algorithmic based capability and understanding what unintended consequences might be happening below the surface of code and qualitative reporting. If people are feeling like they don't understand why things are happening to them, if people are feeling like they are having an undignified experience, they aren't being respected, they aren't able to assert their rights, that what's happening to them is unfair. Sometimes your only way you are able to understand that is to actually go out and speak to real people with real experiences and bring that more qualitative side of the evidence base back into the governance conversation. So when we're thinking about auditability, I guess the question that we should all be asking is are we tapping into the human experience, getting below the data into what's actually happening in the lives of people?

Nick Hart: All right. Tori, any perspective on the same question?

Tori Kim: Sure. Yeah. I'm happy to speak to this, maybe more in the context of a tool like eBrevia and contract analytics and how that kind of reliability element plays into it. I think it goes back to what I was talking about before, when this concept of having the individual, the person who's tasked, whether that's a regulator or other reviewer with this documentation, with this data to actually actively participate in the process of putting that algorithm together, the training that goes into it. I think that is really invaluable to some of that transparency and understanding how things are working beneath the hood when concepts like AI and machine learning can be a bit of a black box for people.

Nick Hart: Just say, I appreciate that all of you have mentioned the term transparency and multiple points during your remarks. And given that it is also Sunshine Week here in the United States, it's a very poignant time to also be having this conversation. In the United States, we have a law called the Foundations for Evidence-Based Policymaking Act and Alex, you alluded to this. There is a mandate about open data. So maybe this is a conversation that we can carry forward in terms of having some recommendations around how we connect some of the dots in the policy mechanisms. Well, I want to thank all of our panelists for this discussion. I'll just acknowledge that we did lose the virtual feed for one of the panelists while we were online here. So there are four speakers, we only heard from three of them and we'll see if we can get the fourth one back before we break at the end of the day here.

So thanks to our three speakers here and we're going to keep moving now. Our next speaker will be here in person with us. We'll have remarks from Julius Chang, who's the Director of Innovation at GrantSolutions, HHS's Office of Grants. Julius, over to you. Welcome.

Julius Chang: All right, good morning, everyone. My name is Julius Chang, I'm the Director of Innovation at grantsolutions.gov. And I'm really excited to be able to talk to you today about some AI focus issues that we are working on within the public sector that really are trying to address a lot of things that you guys talked about

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

today. Empowering our people to do more or less, ensuring promoted equity across government services. But before I dive into those things real quick, just a brief highlight about GrantSolutions, who we are. We are a govern-wide shared service that provides grants management technology...

PART 4 OF 6 ENDS [02:00:04]

Julius:

... Shared service that provides grants management technology, and consulting solutions across the federal government. We are operated out of the department of health, but we serve almost every area of government. We are partnering with 10 plus [inaudible 02:00:15] departments and independent agencies. Last year at F121, we helped agencies issue over 100 billion dollars in grant awards across 2000 different grant programs.

Our mission at Grant Solutions is... It's grants focused. It's uplifting the lives of the American public through increased efficiency of grant services and increased effectiveness of grant funding. We execute on that mission by providing innovative technology solutions, to focus on providing software tools that are easy to use and intuitive. These days, looking at incorporating AI into our solution stack.

That way, we can shift federal grantors, and recipients, from doing the mundane, paper pushing and administrative transparency work, which is important because it ensures that we're being compliant with all laws and regulations regarding government funding, but it allows them to focus their intention on higher value activities, such as monitoring outputs and increasing in the outcomes of the grant funding. Then communities in need get the help that is intended with this grant funding.

With that said, I'll go into two initiatives that we've been working on the last couple years. I think you'll see themes across both initiatives that we are trying to do more with AI, to automate mundane, repetitive tasks, and to help... The focus is not to take over the process, but to help focus grantors on edge cases, edge scenarios, problematic situations, with the intent that projects that aren't going well or that we see more risks, are they're not stopped or not denied funding.

It allows grantors to focus on those projects, and help correct whatever challenges they may have. That way, those communities that depend on only a couple small organizations for healthcare education or infrastructure needs, they can still give those services, because now grantors can work more closely with those underperforming organizations to help them do a better job. Once again, the goal at the end of the day is to help everyone get access to public services. If we can direct agencies to spend their staff hours on helping problematic ones, then that benefits the public as a whole.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

The first tool I want to talk about, is something that we have just released to... That we're pushing for a big adopt across the federal government today. It's called, Recipient Data Insights. It is our risk management solution that aggregates risk data from a number of already public government services, or government data sources, into one single repository. It allows agencies, agency staff, to conduct their pre-award risk assessments in a quick, efficient and timely fashion. For those that don't know, risk assessment is a required part anytime the government gives out money to an organization.

They need to do a couple... That risk assessment is not clearly defined across agencies. But generally in working with our community of partners, we found a couple common things. One, there's a whole host of routine compliance checks. Yes, no, questions. Are they registered in [inaudible 02:03:38].gov? Are they delinquent in any federal debt? Are they on any federal department lists? Those types of things. There are maybe 10 or so different sources that a grant specialist may have to check.

That process is very, very mundane and very time intensive. They have to go to government site A, do a search, repeat the process at websites B through F. Before the [inaudible 02:04:02] this tool, it would take a grand specialist on the order of over an hour to do those basic compliance checks. But these are all already located in government databases. Why can't AI do that for them, right? Now we pull it all in one place, and we have our AI bots run those checks for them. We have little intuitive popups that highlight, is everything good? Great. You can move on to the next one.

If not, if there's an issue with... Maybe the registration has expired in [inaudible 02:04:29].gov, it flags it so that way they know this is one area where I need to work on a little bit quicker. This first wave of just pulling data together has already reaped benefits for those agencies that have decided to employ it. It now takes them less than 15 minutes to do the same number of client checks. Just for scale, we're trying to push this out as broadly as possible for the government. Last year, we issued over a 100 thousand different grant actions through grantsolutions.gov.

If you multiply that by the hour saved per grant action, it's 100 thousand person hours that could be focused in the future on actually fixing problems with underperformed organizations, and improving the outcomes of those services. That's the whole deal behind automated risk checks. Where we want to go next with that is, once we deploy more machine learning, more predictive analytics within the risk tool to help further focus grantors attention onto the things that they consider most likely to impact the output of their grant funding.

We're continuing to engage with all federal departments on a regular basis to understand that they face. This is one of the most exciting things that we're seeing today, because people realize they can save time and do more with less. The next tool that I would like to just briefly highlight for you guys today, is

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

NCC Fast Track. This is again, another automation based tool. The idea behind it is, NCCS stands for non-competing continuations. Those are a type of grant application where you're issued a grant project.

You propose the government a grant project that spans five years and five million dollars, right? But due to federal funding cycles, the agency who wants to approve this project could only give you one year of funding upfront. You have to reapply each year for the next budget of funding. This is a very... Applying for a grant is a very time consuming process. If you guys have done it, there's a lot of documentation to fill out. Then there's a lot of reviews that need to be done by the federal government on the backend.

We had a request from one of our partners, the Administration for Children and Family. They were dealing with staff shortages, and reduced resources to do these types of checks. Their grantors that are in the [inaudible 02:06:51] were feeling extremely overwhelmed at the end of every fiscal year, trying to complete all their works before funding rolls over. They asked us... It opened a question. You guys have all of our grant data, what can you do to help us with our year-end processing? That way the lives of our staff... Our federal business experience is easier.

Then once again, they can put more of their effort into the important part of their work, which is, education, healthcare working with grantees to improve lives in the public. We set out with an open-end question. We investigated the worksheets that they did, and our initial analysis yielded that an NCC on average takes about three months from start to finish [inaudible 02:07:34] process. But of those NCCS, it's remarkable that over 99% of a NCC is ultimately awarded to the recipient. That makes sense, right?

If you were awarded the funds in year one, assuming something major isn't happening... You're doing the work, people get services. You can always argue about, are they doing the best job they can? It still makes sense to deliver those services, because at the end of the day, having some services in underserved areas, is better than no services, right? 99% of these things are ultimately awarded. We put natural language processing to review applications, to review budget information, and look at the themes that the agency uses to decide, are we going to move forward or not?

The end result is, we have now a process in place where the beginning steps of the basic application checks are automated through these workflows after being assessed by our machine learning algorithms. The net result is a reduction in NCC processing times by almost 25%. What used to take three months is now down to a little more than two months. Of those applications, if 99% get through, it allows grantors to focus on the 1% that are having some sort of challenge with delivering the grant project.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Now they can spend those two months worth of time that were spent reading routine, mundane project information, they can now spend that two months of time working with those problematic grant projects, and finding ways to overcome those challenges, that way people get the help that they need. This has been in pilot with the Administration of Children and Family. As we further refine the algorithms that we use there, we hope we can use that as a model to help other agencies adopt a similar type of technology to improve their efficiency, and once again, improve the outcomes that are delivered to the American public.

The real thing that has allowed us to be successful with ACF... I think that will allow us to be successful in even further adoption of these AI tools, is that the types of machine... The way we've set up machine learning in the grants business process, is not to replace the decision makers, right? The government agency leads the grant's management officers. They are still responsible for delivering on the promise of their grant funding. The way we've deployed assistive AI, doesn't take away anyone's decision making power.

It really just helps them focus on the things that are most important to them, which is not checking off all the stuff that is good, but trying to resolve issues that are preventing them from reaching their agency's goals. That's really the value proposition that we offer to the federal government in helping them be more successful, and tell a better story of where our taxpayer dollars go every single day. Those are the two initial AI projects that we're doing at the Department of Health Grant Solutions. In listening to you guys speak today, we are excited for the possibility of widespread adoption for AI in the future, because we believe it's going to be the key to success to unlocking further improved lives of the American public. Thank you very much for your time today.

Speaker 4: All right, Julius, thank you so much. We want to now turn to our final panel of the morning. This is panel four, all in person. I would like to welcome up to our stage, Nicole Dunn from Results for America, Sarah Di Troia, from Project Evident, David Corliss from Peace Work, Albert Lee from Summit Consulting, and David Dreisigmeyer from Trewon Technologies. We will have you speak in that exact order. Where you sit is up to you. Nicole, can I invite you up to the podium first? You'll each have six minutes. The clock will count you down from five. All right, go ahead, Nicole.

Nicole Dunn: Thank you. Thank you, Nick. Hello everyone. Thanks for having me here today, I'm Nicole Dunn. I'm the vice president of federal policy for Results for America. I'm going to talk a little bit about our federal standard of excellence, and how we have used this tool to help federal agencies incorporate equity into their data and evidence work. The Invest in What Works Standard, is a north star or a roadmap to help federal agencies improve their data and evidence practices. It's an annual snapshot of their progress, and hopefully a roadmap for them to help determine how to do better.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

The standard highlights significant progress. The nine federal agencies that participate in the standard have made to operate effectively and efficiently, including their early efforts to implement the requirements of the foundations for Evidence-Based Policy Making Act, which took effect in 2019. Results for America in 2013, began with urging federal agencies to prioritize evidence of effectiveness in their grant programs through annual Invest in What Works Standard of excellence. Americor was among the first agencies to respond, and it has increased its evidence-based grant making every year since then.

Last year, 51% of Americor state and national grants went to programs backed by strong evidence. The impact of this has been significant. For example, the shift helped Minnesota Reading Core, which improves reading skills for K through third grade students. According to multiple rigorous studies, they received 7 million more from Americor over five years, and the program was able to add 400 more tutors, and serve 6,000 more students. We have seen this similar impacts in other agencies, and are excited to continue to see this progress.

I'm just going to touch a little bit on how we have worked to incorporate equity. Building on the foundations for evidence-based policy making, the Biden-Harris administration has taken other important actions, including memos and executive orders on evidence-based policy and equity. They indicate the data driven evidence-based policy with a focus on equity should be the new normal for the federal government. To build on this, results for America took a stakeholder engagement approach to develop a series of equity focused questions, and included them in the federal standard of excellence.

With this addition, we helped agencies use these questions to serve as a guide to integrate a focus on equity into their work, to implement evidence based practices. Building on sophisticated and nuanced data practices can advance understanding of racial disparities through disaggregated data, and lead to improved program specific and agency-wide outcomes by focusing on reducing disparities and inequalities. RFA consulted with government leaders and experts, equity organizations, and others to build these core set of questions, and work with agencies to reflect on them, and implement them, and integrate them into their work.

We saw some great outcomes. The administration for children and families helped build the capacity of their grantees to support culturally responsive evaluation through centers and toolkits. The administration for community living create produced research on accessibility challenges the deaf and hard of hearing encounter due to masks, and other COVID related issues. The U.S. Department of Housing and Urban Development offered technical assistance and resources to grantees to address racial disparities, especially those amplified by COVID-19.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

The Substance Abuse and Mental Health Administration engaged in a renewed focus on racial equity, diversity and inclusion. Through this office, grantees submit disparity impact statements to ensure SAMHSA programs are inclusive of underserved racial and ethnic minority populations in their services. We look forward to continuing to work with the agencies to incorporate an equity focus in their evidence and data work, and share more examples. Thank you.

Speaker 4: I don't see Sarah up here, so David, let's turn to you.

David Corliss: Hi, I'm David Corliss. I'm very, very happy to have you a chance to talk to the group today, and maybe continue the conversation this afternoon. A little bit of a background, my organization is Peace Work. You probably haven't heard of us. We are an NGO in Data for Good. We are an all volunteer organization, so nobody's getting paid for anything. People are donating their time to work on data science projects in for the public good.

Very often in issue-driven advocacy, poverty, education, success, and failure. That's frankly the sunshine and rainbows version. If you want to go and look at our website, that's peace-work.org. Look at our studies page. I like to call it the ugliest page on the internet. It seems like we have a lot of studies on people who do very bad things, but with the team working together on this, I wanted to give a shout out to some of the folks that have been working on this.

I'm very active in the American Statistical Association. I just had a chance to chair their conference on statistical practice. A lot of what we're going to hear now is going to be from the ethics panel we had there, which I host every year. Steve Pearson, some of you may know. American Statistical Association policy and advocacy director, Donald [Helone 02:18:31], also from the ASA. Also, Candace [Endahlia 02:18:34], who is an attorney here in DC, but is also a Peace Work volunteer, and is giving us a legal aspect on some of these.

As we've gone through the morning, we've heard the concerns. One of the things I haven't been hearing, is a whole lot of really, really major technological barriers. If we were having this conversation five years ago, two years ago, we might have been saying, "Oh my goodness, how are we ever going to get the computing platform to work?" I didn't hear anybody say that today. We've seen some really, really amazing products from government and from industry, that are making a huge impact. Technology is not the big area to accelerating AI.

We know what the needs are. We know what the needs are. We've heard about privacy. Brandon talked about equity. Sean talked about the problem with unethical applications. There have been a number of concerns. The things that are holding AI back, it's not the technology, it's not that we don't have a government support, it's not that we don't have amazing corporations developing this technology. The problems are ethical problems, and what this calls for, the way we're going to be able to move this forward.... What it calls

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

for, is for the creation of an ethical ethics best-Practices board within the data coalition.

That means inter-agency, not connected to one particular group or part of a specific organization. It can work across all agencies, the same. An ethical best-practices review board within the data coalition. You're going to be able to address all these different issues that we've heard over and over throughout the day. It's going to be able to advise federal agencies on developing best standards for AI, best practice, and also frankly, consistent practices. I'll give you an example. One of the things that we see a lot and with concerns about data privacy, the usual approach to data privacy is, here's a list of PII. Oh, and there's some sensitive PII, and we're going to make another list of things that are especially radioactive.

The first thing you'll find is, it's a list. It tries to solve an ethics problem with nouns. I want to talk about parts of speech for just a minute, and please bear with me. It talks about the things we have. If we have conversations and they're noun driven, that's going to be data and technologies. Those are things we've got. That's not the problem. If we talk about things that are verb driven, it's the actions we're taking. Those seem to be going along fairly well. That's not what's holding AI back. What's holding AI back is, ethical questions and ethics are described and written up in adverbs.

Equitably, evenly with diversity, and with an attitude towards being respectful of human rights. It's the manner in which we do the things we do. It's the adverbs that describe the ethics. This is the sort of thing that an independent ethics best panel group inside of the data coalition can focus on. It can work on consistent data, privacy standards, educating. It can also be public facing to produce information for the general public, but also build trust, because in the end, that's the critical issue. We've heard it mentioned briefly a couple of times.

The one thing that pulls together all the different talk we've heard today, that what's holding AI back if one thing... It's not the technology, that's way out in front. It's not the suppliers and people who are developing it. It's not the government partners and bringing together stakeholders. That's all working well, but AI can't advance unless people will put their confidence in it to trust that their data are being treated equitably and fairly. To trust that algorithms aren't going to treat people unfairly. These are all issues of trust.

The key point I want to deliver here, is that AI moves at the speed of trust. If we're talking about accelerating AI in the public sector, what that means is accelerating trust in the public sector, trust in AI that we use. That's the only way we're going to be able to accelerate this. How do we deliver on that? We're getting towards the end of the talk. How do we deliver on that? We're going to do that with an independent ethics best panel board that can create resources and material.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Can advise, can provide enforcement in the way that we have with the other safety review boards. Imagine on AI safety board, like we have a transportation safety board that invests accidents and the FDA, and so on. An AI safety board inter-agency as part of the data coalition. When things come up like the compass algorithm that keep people in jail because they're black, and things like this. It's a place where people can take direct action when something goes wrong with AI. Accelerating AI in the public sector means accelerating trust in AI, because AI moves at the speed of trust. Thank you.

Speaker 4: David, I hope you aren't suggesting that I'm running the AI safety board. I think I might have misheard you there at the very end, so let's come back to that later. MC programming note, I apologize. Sarah will be joining us virtually. We're going to modify this and put her at the end of the panel. I want to turn next to Albert and then David. So Albert, if you'll join me at the podium.

Albert Lee: I guess still good morning, everyone. I'm Albert Lee, happy to be here in person. I'm a partner of Summit Consulting. Listening to all the previous speakers, I realized that I probably was part of the problem. When I'm actually reflecting on this talk, and I reflect on my career as well, I want to actually couch the whole AI discussion in this long train of what I refer to as, algorithmic revolution, that has been happening in every aspect of our government agency. By algorithmic revolutions, I'm talking about agencies increasingly relying on data-driven algorithm, or to make certain type of decisions.

In my opinion, AI differs in degrees, not [inaudible 02:25:24]. Agencies that I had the privilege to work with, had begun to make decision of this type 20 years ago. The very, very first algorithm that I worked on and actually tested, was mortgage underwriting. That was the first piece. The second piece that I had a privilege to work on more recently, is about automated valuation. Think Zillow. Why do I actually single out that two pieces of algorithm? Because those decisions had not only a temporal impact on the particular borrower and asset holder, it has intergenerational impact from one generation to another generation.

Owning a house is the best way to pass wealth from one generation to another generation. If the algorithm tells me or tells you that you cannot actually have a mortgage, that not only impacts you, but it impacts your children. Asset prices. The house that you have, imagine that Zillow for whatever reason, actually gave a higher price in a white neighborhood, and actually a lower price in an African American neighborhood. Even if you own a home, the extent to which that you actually pass on the wealth to your next generation is reduced.

Where do these things come from? Well, data. Where do we actually get the bias from? In a lot of algorithm, be it in a typical model, or be it in AI, is usually trained with certain type of data. The data that agency has, is administrative data. They have not necessarily collected based on some type of randomness. It is based on the fact that agencies in the past had actually done a lot of work.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

These works were selected. They were based on expert judgment, and if the expert judgment was not executed appropriately, it embodied the bias.

That primarily is the sources of where we actually get the bias from. Part of the benefit of actually couching the whole AI discussion in this algorithmic revolution is because, while AI is new, the problem that we are confronting is not completely new. Office of inspector general for the past 20 plus years, had audited financial model day in and day out. It is actually completely part of their routine. If the IG begin to actually start looking into not just pure statistical and economic advocacy of some of this decision, but injected in their work what David said about certain type of values, transparency, accountability, equity, and fairness.

When we actually start to audit and understand that these type of models, not only they need to be efficacious, but it also needs to be fair. It also needs to be transparent. It also needs to actually create the right counter of accountability. Then I think that we will probably have the beginning of a discussion. Algorithmic decision making is not going away. AI is one step away to completely accelerate it. You know, this particular train. Sooner that we actually get a hold of this problem, the better off we will be. Thank you very much.

Speaker 4: All right, Albert, don't go too far for questions. All right, David, you're up.

David Dreisigme...: Hey everyone. I'm going to be talking a little bit differently, but everybody I think has touched on this, or at least a lot of people, is the workforce in the federal government, and using AI. I'm going to come from this at a perspective as a FDS fellow, Federal Data Strategy. I was also debt CDO at commerce. When I was working on putting together the first draft for the playbooks, for the workforce data skills part, for the federal data strategy, and then they also had a split off for the...

PART 5 OF 6 ENDS [02:30:04]

David: ... then they also had a split off for the data skills training catalog, also. When I was going around the federal organizations, the one thing that I noticed was a lot of them were putting together profiles for their workers. It's like a data scientist needs these skills, and they need a certain level of proficiency in these skills. Someone who's interpreting the data would need these skills, and certain levels of proficiency. It's not always the same skill set that you need, you don't always need to be an expert in a lot of these skills.

When we're thinking about AI, we're not just thinking about the data scientist. We would also be thinking about the consumers internal to an organization, which would be the management. If you want to have evidence based decisions being made, and you're using more analytic techniques to build the evidence and the story behind this, you need to also have people that are going to be able to interpret what you are developing.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

When we were writing the first draft of the playbook, what we used was the National Institute of Standards and Technology, their SP 1500, which was the big data framework. They had roles within that we tried to use to provide this kind of profile structure to the playbook. Then the final version that came out, there was a modified profiles and skills section that came out for the playbook that the FDS initially released. If we fast forward a little bit, when we were at commerce, and we were trying to do our data skills assessment, and we were using the SP 1500 to come up with profiles. We needed to relate to the individuals within each of the agencies, exactly how they fit into this whole data stream. Just because you are not programming in Python does not mean that you're not using data. The SP 1500 kind of fell flat, a lot of people didn't really like it, wasn't helpful to them, whatever.

If we look at Nest again, they have a cybersecurity framework. The national initiative for cybersecurity education, has a framework that talks about the tasks that you need, to do cybersecurity, and the skills and knowledge that you need to do these tasks. It would be nice to have a very similar framework to use for data skills, and AI within the federal government. If we begin to look at data itself, there's a life cycle to data. Part of that is analysis, which is where a lot of the AI and data scientists would fit in. There's also the business case of, what questions am I trying to answer, what data do I need to answer those questions? In that case you would need a different skill set, you have different tasks in there, you need different knowledge.

Then there's also the consumers of the data downstream, the people that would be making decisions, who are also going to have different tasks to perform, different skill sets, they need different knowledge. Then with that, if we have this framework for understanding the tasks during a data life cycle, if we know the skills, and the knowledge that needs to go into performing these tasks. You begin to create these profiles either generic ones, or for your agency specific where you can begin to relate to your workforce exactly where they fit into this whole data pipeline. If you're an archivist, for instance, you are very important because you're maintaining the data for the stakeholders. Whether that's internal stakeholders or external stakeholders.

Sorry, Nick, I thought this was going to be quicker.

That helps then by doing the staff skills assessment so you can find out where your organization is at. You can say, well, we need these skills at this certain level of proficiency for our organization, and our staff are at this level, we have a gap. How can we go about addressing this gap?

One way is to have training programs, obviously that try to fill this gap. You know, what your skills, and knowledge that you need to train to. There was an organization that had something very similar. They had on tracks and off tracks to gain competency, to gain new skills, new knowledge that you needed to do your job, advance your career, to go into another job. It also helps then to

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

create job descriptions if you're trying to hire somebody, because you can have a very specific list of skills, and knowledge that you need in order to do the tasks within your organization.

Another thing, it's common language that would be used by multiple organizations. If you're looking for instance, God forbid loan out your expert to another agency. There's a common language. One agency needs this skills, you have a person that meets that requirement, you can let them go on detail for six months to go and work on that project. Then also, if we look at the America Competes Act, and with the nice framework for cybersecurity, there is an initiative in there to begin looking at collecting statistics on the American workforce, as far as cybersecurity goes. If we have a similar framework for our data skills, that can serve a useful purpose. Just like the nice framework for beginning to assess where we are both as federal government, but then also if we begin to look at surveys for the general public too, we can use that framework.

Thanks.

Nick Hart: All right, thank you, David.

I will next turn us to our virtual speaker for this panel, Sarah, from Project Evidence.

Sarah, over to you.

Sarah: Wonderful, thank you for much.

Good afternoon, everyone.

I'm going to take a slightly expanded version of the definition of public sector. We have mostly focused on, actually, I think we've entirely focused on government over the course of the speakers this morning. I want to talk about the nonprofit sector, the practitioners. The reality is when you think about the social service safety net, and you think about many aspects of education, health, environmental policy, et cetera, et cetera, it's all delivered on the ground through a broad and diverse network of nonprofits. Project Evident is interested in harnessing the power of evidence, to drive outcomes, working with nonprofits, really putting them in the driver's seat with their own data. Too often their data is, frankly at behest of funders evaluators, and it actually isn't being used to drive their own knowledge, and learning and R and D. we are very interested in how AI is going to reshape the nonprofit sector, and their ability to harness, and mobilize data, and we're interested in understanding what's happening.

What I want to give you is like actually a view from the ground. We interviewed about 30 practitioners, largely in the education, and social services space.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Vendors who were creating SAS, AI tools and applications, as well as other thought leaders in the sector, just to understand what was happening. The good news is that there is a tremendous amount of opportunity that nonprofits see in the advance of AI. Not surprisingly the bad news is there's some extraordinary barriers to nonprofits actually being able to access, and get the full power of using AI with their data to drive outcomes. There's no shortage of evaluation work that's happened, discovery, R and D, to try and shift the outcomes for black, brown, and low income populations in the US.

Unfortunately, most of the evidence that's been generated, and what works is locked in reports. It's all in unstructured data, it's very hard to access, and it may not be relevant to specific populations, and it's really limiting the usefulness of that information. When we talked to nonprofits and asked them about what they saw as the opportunity for AI in the future, right at the top of the list was the ability to generate, use, and mobilize evidence for outcomes. I think this is a really interesting contrast with the for-profit sector. If you think about the journey of AI and the for-profit sector, which was initially really seen as a way of driving efficiencies. Now is really seen as a way of strategic imperative, and competitive advantage. I think the nonprofit sector is already seeing competitive advantage. In their words, that's around driving better outcomes for constituents.

They also believe that AI is going to help them identify disproportionality. Facilitate R and D solutions, and to help people on the ground make better decisions to address systemic racism, and enact better policies and practices. They really see this as a way of shifting the way the program is implemented, and having higher, and better, and more equitable outcomes. It also increases their impact from moving data to real time insights. Right now, data frequently is used in the sector in an opportunistic manner. As a way of proof, of proving what we're doing, and then collecting administrative data. This would really shift the way nonprofits can have access to data on a real time basis to drive, and learn, and increase impact immediately. It would broaden access to information to new audiences. They also see it a way as facilitating scale of their program models. Finally, obviously there's efficiency for them as well, but that is really last on the list. Really first on the list is around impact, and equity.

Now, the challenge of course, is there's some extraordinary barriers for nonprofits to be able to access AI. We've all touched on talent, across many elements of the market, whether it's in government, in other places, in stem education. The real challenge right now though, is that nonprofits have access to very few people who have talent, and knowledge around AI, and how to enable scale solutions for data collections, synthesis, and mobilization. Unfortunately, what we heard is that even when they find those folks or they begin training up their own staff, and investing, and up skilling them, they frequently jump from the nonprofit sector, and they move to the for-profit sector where it's so much more lucrative.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

There's a real talent challenge, a real talent gap, in the nonprofit sector for being able to access AI. The second challenge is just the cost. AI is considered more overhead for nonprofits, and that's both the investment of the technology, the up skilling of staff, the hiring for new capabilities, and also the change management. I love how some of our earlier speakers were talking about how technology is really about people, this is about moving to a place of constant change. The beginning of that is really implementing change management, and new learning routines inside an organization. Very few nonprofits, if any, have full cost accounting. That means every single time you serve a beneficiary, you create a gap of funding that has to be filled from another service, another place, often that's from private philanthropy. In the absence of full cost contracting, or even R and D money, it is very hard to see how nonprofits are going to be able to incorporate AI into their overall tech stack.

We've talked a lot about equity, and justice, and bias over the course of this morning. There is a pretty significant equity gap between practitioners equity commitments and what's happening currently in the technology products. Often the way technology comes into the nonprofit sector from the for profit sector, is it moves from custom build into SAS applications, let's say. Which are much more lower cost and more attainable on a financial basis for nonprofits. The challenges when we interviewed SAS application providers, the for profit clients at the moment are not asking them about equity, and design justice. Those products are not being developed with equity, and design justice in mind so those products are going to be frankly, not compatible with what the customer needs are for the nonprofit sector. That leaves them in a place where they have to do custom builds on algorithms, which is more expensive, which puts them further away from being able to access the technology.

Finally, there's a real challenge on the fragmentation of just offerings in the sector, which really puts the nonprofit in a general manager, they have to build their own taxonomies, they have to talk to somebody else about data cleaning, they have to talk to somebody else about an algorithm development. The nonprofits themselves, as I've said, they're at a talent desert around AI knowledge, and so they're in a very bad place to be general contracting across a whole suite of vendors, and consultants to support them. That again, limits their ability to be able to access AI technology. When I think about the nonprofit sector, there is huge opportunity to think about equity, and scale, the most important outcomes, and increasing, and bettering those outcomes for people in America. The real challenges are again, not technology, unless we talk about the equity inside of technology, but it is around talent, and it is around dollars.

When I think about the call to action for the nonprofit sector, as a partner to the public sector, a partner to government, and actually delivering upon its commitment, I think about risk capital. Government is one of the least risk tolerant investors in nonprofits. They want to know it's going to work a hundred percent before they want to make an investment. Really thinking about what

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

are the pools of risk capital that are available for early adopters to come forward, and begin experimenting with AI. That also means, and this has been mentioned previously, that folks in government need to be educated about AI, and how to make those investments, and how to evaluate those investments, even while embracing more risk.

Finally I want to underscore the algorithm of transparency. I think one of the collateral benefits of the last several years, it's hard to see any benefits, but the nonprofit sector really moved firmly into being platformed. Whether it's Microsoft, or Alphabet, or Salesforce, many more non profits have strong technology platforms, but we know that all those platforms are going to begin releasing algorithms, and AI as part of their standard package. We're really concerned if nonprofits haven't had a chance to experiment, and learn about what works for them, that they're going to be stuck with black box algorithms that can't be tested for equity, and for justice. Algorithm transparency is sort of the last call to action I put out there for the nonprofit sector.

Thank you.

Nick Hart:

First of all, Sarah, thank you so much. I appreciate the recognition that the public sector, is actually a much broader term than we've given it credit to, largely in our own framing of this event. In many ways, our shuffle of the agenda to put you at the very end was nicely done. I'm sorry that we shuffled you around, but I'm actually grateful for the way that that worked, since I did not know exactly what you were going to say.

I want to start by actually asking you a question, and then I'm going to open it up to the rest of the panelists here, but I'd be curious to get your perspective on how government itself could specifically help in this nonprofit context. There's a lot that we've talked about in government, if you have a suggestion for what government might do.

Before we turn to you for the rest of the speakers, I'm going to ask you the same question that I've asked many of the other panelists today, before we turn to our closing speaker for the morning session. If you had just a few seconds with a senior policymaker, and in fact there is a congressional hearing coming up next week, I think it's next week, at the end of this month, which will be on AI, and regulatory technologies, which is the topic of our afternoon session. If you were meeting with the congressional members, what would you tell them as a recommendation for how to advance this work in the near term?

That's going to be the question for everybody in the room, but Sarah, let me turn to you first.

What can government do to help the nonprofit sector?

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

Sarah: A couple things, one is that the nonprofit sector's been deeply appreciative about way government has made data more available. It moves us out of a situation where we have to collect our own data to do randomized control trials, because we can have access to far more data. The challenge is that folks don't often think about making that data available in a way that nonprofits can access it. They think about local governments accessing it, they think about large institutions, and academic institutions accessing it. If government could think about how could data be made available so that smaller nonprofits, which is the backbone of social services, and education services in our country, how they can actually make best use of that data to drive outcomes, and learning.

Nick Hart: Go down the line here, we'll start with Albert.

Albert: If I have a few minutes with a senior officer, I will probably impress up on him or her about the distillation of the values, because I don't think that we can actually make effective policies if we do not get the value. Transparency, accountability, fairness, and due process. I think that if we are actually focusing on these values that actually promote the democratic society that we are living in right now. I think it is absolutely key because otherwise AI would just become a tyrant, and it would exist outside of the democratic norms, and I think that it would actually do more harm than good to the society.

Nick Hart: David?

David: I guess mine would be to fund the CDOs more than what they currently have. These are new officers, they're not going to be at the same level necessarily as the CIOs, as far as funding, and influence within their own organization. A lot of the CDOs would be the natural place to begin to integrate AI, and to work on a cultural change that you need order to really effectively leverage this in the government. If they're not fully staffed, if they're not funded to a sufficient level, it's very difficult to pull that off.

Nick Hart: All right, David.

Speaker 5: Create an AI safety board like we've got for transportation, and financial services, and food and drug. Go down the line, create an AI safety board with a twin mandate safety, just what we're talking about when things go wrong, when algorithms hurt people, and the second mandate to create resources that are going to commonize ethical human rights, driven best practices across the federal apparatus.

Nick Hart: Clarity, it's good.

Nicole final word,

Nicole: Certainly support these recommendations (silence) quickly.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

That's okay. I can speak loudly, I think, nope? That okay.

Sherry: [inaudible 02:50:53].

Nick Hart: It's flashing at us. We'll use the other one, here you go.

Nicole: I would support the recommendation on the supporting the work of CDOs, and setting aside funding to build capacity in the agencies to do this work, and ensure equity is integrated as others have said throughout the process.

Nick Hart: All right.

To all of our speakers for panel number four, we want to thank you for sharing your insights and your perspectives. This has been fantastic. Thank you so much.

We'll free you from the stage, and I would like to welcome now our final speaker of the morning. Sherry Bennett is the chair of the Data Coalitions Advisory Board, every year, the Data Coalition Advisory Board elects a new chair. We're very pleased to have her DLT Solutions is a partner level member of the data coalition. Sherry's been a great leader, and advocate champion for us. Sherry, I'd like to invite you to offer some closing remarks for our morning session.

Sherry?

Sherry: Thank you Nick, and good luck with you around your next event.

Thank you, thank you.

Thanks again, Nick, and actually the entire data coalition team. He's got a fantastic team. They're a pleasure to work with, and they're so busy. Providing this venue, and this forum to have these conversations that are so important is really a great thing.

Just a couple of brief comments, want to thank all of our panelists for this amazing session, and to our attendees online, and in person here for tuning in. We've heard an incredible set of perspectives, and recommendations, and I don't know about you, but I feel like I've learned a lot. I appreciate all those different perspectives, and these conversations need to continue. On behalf of the data coalitions members, thank you again to our speakers for sharing your expertise, your discussion about your projects, and guidance on best practice for use of AI in the public sector. As we've heard this morning, there's a lot of exciting things happening around AI, and the deployment of projects, and pilots. There's still a great notable exception in areas to think about when it comes to how to successfully develop, and deploy these solutions.

This transcript was exported on Mar 22, 2022 - view latest version [here](#).

It is critical for government to make progress, implementing the use of a AI more so than it has so far. It's time to move beyond the research, the mini pilots, proof of concepts, and really begin to put AI in practice at an enterprise level. We know this is really complicated, but we need to begin this hard work. The government obviously should not do this alone, I think we all agree on that, in fact, government can't do this work alone. Collaboration with industry leaders in the private sector, and experts from academia is absolutely crucial to ensure that the federal government has the best, most secure, and the most equitable AI technology capability. Government needs to maintain, and I think in most cases needs to really expand, collaboration with outside experts, and partners to make progress, and continue innovation in the years ahead.

Today's forum furthers the ongoing efforts, and the collaboration that we all seek. We've heard a number of very important topics today, equity and AI systems, and applying a ethics framework. Realistic steps to inject transparency, and accountability in algorithmic processes, and the need for trustworthiness in AI. Data protection, and privacy measures, certainly the ability to manage data, and platforms that are stable, cataloging the data and so forth.

We have heard a lot about ways to support government workforce capacity. The needs for particular skill skillsets, and machine learning, and other AI technologies and methods. Literacy around those issues, and using things like machine learning, and other AI technologies to carry out the Evidence Act requirements, and open up agency data. Our panelists discuss their own projects, their research, and their advice for using AI to improve our government. If well developed, taking into account all of the topics we heard about from our speakers this morning. AI really does, as we all know, have the potential to transform how government provides services to its citizens. By automating many of the manual processes that exist, the government can direct human capital to improve programs, and address the many needs of the American American people. Efficiency and transparency, and data sharing and access strengthens trust, and the capabilities of our government, and ultimately strengthens our very own democracy.

To ensure AI is ethical, equitable, and transparent, public / private exchanges like this very one must continue. We need to share best practices, identify key concerns, and take advantage of the many opportunities that we know, and see every single day that are out there. We look forward to continuing these conversations with you and the policy makers in the months and years ahead.

Thank you.

PART 6 OF 6 ENDS [02:56:18]



3138 10th Street North
Arlington, VA 22201-2149
703.522.4770 | 800.336.4644
f: 703.524.1082
nafcuhq@nafcuhq.org | nafcuhq.org

National Association of Federally-Insured Credit Unions

May 12, 2022

The Honorable Bill Foster
Chairman
Committee on Financial Services
Task Force on Artificial Intelligence
U.S. House of Representatives
Washington, DC 20515

The Honorable Anthony Gonzalez
Ranking Member
Committee on Financial Services
Task Force on Artificial Intelligence
U.S. House of Representatives
Washington, DC 20515

Re: Tomorrow's Hearing: "Keeping Up with the Codes – Using AI for Effective RegTech"

Dear Chairman Foster and Ranking Member Gonzalez:

On behalf of the National Association of Federally-Insured Credit Unions (NAFCU), I am writing to share NAFCU's views on granting the NCUA third-party vendor authority in conjunction with tomorrow's hearing. NAFCU advocates for all federally-insured not-for-profit credit unions that, in turn, serve over 130 million consumers with personal and small business financial service products.

NAFCU and our member credit unions believe that cybersecurity, including the security of vendors that credit unions do business with, is an important issue. However, we are opposed to granting additional authority to the NCUA to examine third parties at this time. NAFCU believes in a strong NCUA, but we also believe that the NCUA should stay focused on where its expertise lies—regulating credit unions. It is important to note that credit unions fund the NCUA budget. Implementing such new authority for the NCUA would require significant expenditures by the agency. The history of the NCUA's budget growth has shown that these costs would ultimately be borne by credit unions and their 130 million members.

There are tools already in place for the agency to get access to information about vendors. We believe the agency's time and resources are better focused on reducing regulatory burden by coordinating efforts among the financial regulators. The NCUA sits on the Federal Financial Institutions Examination Council (FFIEC) with the Federal Deposit Insurance Corporation (FDIC), the Office of the Comptroller of the Currency (OCC), and the Federal Reserve. The FFIEC was created to coordinate examination findings and approaches in the name of consistency, and to avoid duplication.

In September 2020, the NCUA's Office of Inspector General (OIG) released a report titled "[Audit of the NCUA's Examination and Oversight Authority Over Credit Union Service Organizations](#)" (the Report).¹ The Report makes several observations regarding the extent to which the NCUA may participate during joint exams of technology service providers (TSPs) led by FFIEC agencies. According to the Report, the other federal banking agencies have adopted a "guiding principles document" for vendor reviews that

¹ NCUA OIG, Audit of the NCUA's Examination and Oversight Authority Over Credit Union Service Organizations (September 1, 2020), available at <https://www.ncua.gov/files/audit-reports/oig-audit-cusos-vendors-2020.pdf>.

The Honorable Bill Foster
The Honorable Anthony Gonzalez
May 12, 2022
Page 2 of 3

prevents the NCUA from participating in examinations of technology service providers.² It is unclear whether this is the primary barrier the NCUA faces, and NAFCU understands that there have been occasions when the NCUA has obtained access to a joint examination report for a TSP.

The restrictions in the guiding principles document likely derive from a legal opinion letter produced by one or more of the banking agencies sometime between 2013-2014.³ The Report summarizes the legal opinion as concluding that the NCUA lacks the legal authority to accompany federal banking examiners during vendor reviews.⁴ The NCUA has not publicly shared this legal opinion letter. Accordingly, it is difficult to assess whether the NCUA's apparent inability to join FFIEC vendor exams is a true statutory limitation or merely a legal interpretation proffered by another federal banking agency.

A reassessment of the federal banking agencies' position on whether the NCUA can participate during joint exams of TSPs offers a more straightforward and simpler solution than granting the NCUA a new authority with potentially unlimited scope and budgetary impact. As a member of the FFIEC, the NCUA should be able to request the results of an examination of a core processor from the other regulators and not have to send another exam team from the NCUA into their business and duplicate an examination that has already taken place. This would seem to be an unnecessary burden on these small businesses. Additionally, if the NCUA did its own examination, the likelihood of finding anything the other regulators did not would seem to be close to nil.

Recognizing alternatives to vendor authority, Congress should require the NCUA to measure the costs and benefits of developing a parallel vendor supervision program versus obtaining vendor examination reports from the FFIEC agencies. The NCUA should also supply a clear description of the stated objectives and scope of a third-party supervision program. For example, it is unlikely the NCUA will have the resources to supervise every vendor, so a risk-based prioritization framework would need to be developed. The NCUA has offered little indication of how it would tailor such a policy in order to appropriately manage its administrative resources.

More troubling is the NCUA's lack of transparency regarding the overall potential budgetary impact of a new vendor supervision program. In one recent whitepaper, the agency equivocates on the issue of cost: "[w]hile this may increase the NCUA's budget due to the addition of more examiners with specific expertise, the agency does not expect a dramatic increase."⁵ Yet the NCUA provided no specifics or estimates. Furthermore, the NCUA has suggested that the availability of other federal banking agency resources will minimize the cost of implementing a vendor supervision program—a fact that seems to acknowledge that existing resources, rather than new NCUA exams, can generate useful supervisory data at lower cost.⁶

² See *id.* at 14.

³ See *id.* at 20.

⁴ See *id.*

⁵ NCUA, Third-Party Vendor Authority, 10 (March 2022).

⁶ See *id.*

The Honorable Bill Foster
The Honorable Anthony Gonzalez
May 12, 2022
Page 3 of 3

Instead of granting the NCUA vendor examination authority, Congress should encourage the agency to use the FFIEC and gain access to the information on exam findings on companies that have already been examined by other regulators. If that option is not available for the NCUA due to the decisions of other regulators, Congress should consider compelling the other regulators to share the information. This would seem to be a much more preferable route than raising costs on credit unions and their 130 million members for the creation of a duplicative NCUA program. Supervisory reports for core providers will likely have significant cross-applicability; according to the NCUA, approximately 5 core processor vendors control approximately 85 percent of credit union data.⁷ Use of existing reports for other TSPs would also address the NCUA's concerns without creating additional costs to credit unions or increasing regulatory burdens on credit unions and small businesses. As such, we urge Congress to oppose granting the NCUA this new authority and urge you to oppose H.R. 7022, the *Strengthening Cybersecurity for the Financial Sector Act of 2022*, in its current form.

Thank you for the opportunity to share our thoughts on this issue of great importance to credit unions. Should you have any questions or require additional information, please do not hesitate to contact me or Chad Adams, NAFCU's Senior Director of Legislative Affairs, at (703) 842-2265 or cadams@nafcu.org.

Sincerely,



Brad Thaler
Vice President of Legislative Affairs

cc: Members of the House Financial Services Committee

⁷ NCUA OIG, Audit of the NCUA's [Examination and Oversight Authority Over Credit Union Service Organizations](#) at 3.



Office of the Comptroller of the Currency

Washington, DC 20219

July 13, 2022

The Honorable Bill Foster
Chairman
Task Force on Artificial Intelligence
Committee on Financial Services
United States House of Representatives
Washington, D.C. 20515

Dear Chairman Foster:

Enclosed please find my response to the question for the record submitted following the May 13, 2022, hearing entitled "Keeping Up with the Codes – Using AI for Effective RegTech."

If you have questions or need additional information, please contact Carrie Moore, Director, Public Affairs and Congressional Relations, at 202-649-6737.

Sincerely,

A handwritten signature in black ink that reads "Kevin Greenfield".

Kevin Greenfield
Deputy Comptroller for Operational Risk Policy

Enclosure

Mr. Greenfield, I am concerned that the continued adoption of AI in regulatory and supervisory technologies may lead to unfair disadvantages to women and people of color, and older American consumers of certain financial services, if these concerns are not addressed sooner rather than later.

Mr. Greenfield, what are the steps that we, as Congress, can make today to ensure unbiased AI for the future?

Response: The OCC shares your concerns about the potential for unfair disadvantages to vulnerable groups due to the use of new and emerging technologies. I can assure you that OCC examiners have significant experience in supervising banks' use of sophisticated mathematical models and tools. Additionally, we expect banks to monitor for and identify practices that create risks of adverse outcomes and violations of law that involve or could result in unfair treatment of consumers. If we identify concerns, risks, or deficiencies during our examinations, the OCC will use the range of tools available to take supervisory or enforcement action, as appropriate.

We welcome Congress' continued oversight in this evolving area and would encourage ongoing dialogue with stakeholders, including industry representatives, advocacy groups, and the regulators, to promote clarity and understanding of this issue. We look forward to continuing to engage with Congress to ensure that all consumers are treated fairly and provided fair access to financial services.



 National Credit Union Administration

**United States House Financial Services Committee
Task Force on Artificial Intelligence
Keeping Up with the Codes – Using AI for Effective RegTech
May 13, 2022**

Response to the Question for the Record

Question for Ms. Kelly Lay, Director, Office of Examination and Insurance, National Credit Union Administration (NCUA), from Representative Sylvia Garcia:

- 1) How can AI transparency benefit financial institutions, such as credit unions and consumers specifically?

Response:

Transparency in artificial intelligence (AI) plays a vital role in developing understanding and trust among the financial institutions and consumers. For the most part, the methodologies and parameters for these systems are protected intellectual property and trade secrets that are not available to the public to review and evaluate. As a result, AI-based systems, how they are used, and their potential benefits are often misunderstood.

AI only works as intended when it is explainable and understood by its user(s). The NCUA continues to foster a greater level of certainty and understanding about how AI systems work, the outcomes they produce (that is, what the machine-learning or AI-based technology does and the parameters it uses to evaluate a scenario and to reach a decision), and how those results will be used. The NCUA, along with the Office of the Comptroller of the Currency, the Federal Reserve System, the Federal Deposit Insurance Corporation, and the Consumer Financial Protection Bureau, issued a March 2021 Request for Information (RFI) and Comment on Financial Institutions' Use of AI, including machine learning, to examine issues such as bias, fairness, and transparency. The NCUA is still evaluating potential next steps based on the few comments received with respect to that RFI.

Some of the benefits of AI transparency are universal, like the reduced potential for abuse, errors, and fraud. Greater levels of transparency help to identify unintentional biases, data inequality, and unfairness with AI models. Transparency also assures stakeholders of the level of objectivity in the system's decision-making process, helping to build greater trust.

The NCUA continues to research and explore using various AI tools to help monitor and regulate credit unions and improve efficiencies within the agency's examination and supervision program. If the NCUA were to implement an AI-based system in the future, the agency would disclose what data was used and how, how the results developed were derived, and how the NCUA intends to use the results.



May 6, 2022

**Statement for the Record by [SecurityScorecard](#)
Before the House Committee on Financial Services Hearing Titled,
*"Keeping Up with the Codes—Using AI for Effective RegTech"***

Chairwoman Waters, Ranking Member McHenry, thank you for the opportunity to present SecurityScorecard's testimony related to the use of emerging technologies such as Artificial Intelligence (AI) and Machine Learning (ML) by financial regulators to support their supervisory, regulatory, and enforcement efforts.¹

This hearing could not come at a more important time. Several federal agencies recently issued new or updated regulations and guidance documents to improve cybersecurity across the federal enterprise and the private sector. For example, the National Institute of Standards and Technology (NIST) has updated, or published drafts for comment on, three foundational cybersecurity frameworks; the Department of Defense continues to improve its Cybersecurity Maturity Model Certification (i.e. CMMC 2.0); and, most recently, the Securities and Exchange Commission (SEC) proposed two separate rules which would require cyber incident reporting, reporting on executive and board cybersecurity expertise, and other cybersecurity provisions.²

SecurityScorecard believes that AI and ML tools, like security ratings and continuous monitoring platforms, can drive cybersecurity insights, improve oversight by regulators, and provide incredible company value by reducing cybersecurity risk. As the committee considers the role of AI and ML in Regulatory Technology (RegTech) and Supervisory Technology (SupTech) to keep up with changes in cybersecurity threats and policies, SecurityScorecard would like to present a case study to consider.³ Recently,

¹ Memorandum, Committee on Financial Services, U.S. House of Representatives, Mar. 25, 2022, <https://financialservices.house.gov/uploadedfiles/hhrg-117-ba00-20220330-sd002.pdf>.

² See, for example, "SP 800-161 Rev. 1 (Draft)," National Institutes of Standards and Technology, accessed on Mar. 28, 2022, <https://csrc.nist.gov/publications/detail/sp/800-161/rev-1/draft>; "About CMMC," Department of Defense, accessed on Mar. 28, 2022, <https://www.acq.osd.mil/cmmc/about-us.html>; "SEC Proposed Rules," Securities and Exchange Commission, accessed on Mar. 28, 2022, <https://www.sec.gov/rules/proposed.shtml>.

³ "RegTech" refers to the use of emerging technologies such as AI and ML by financial institutions to ensure compliance with applicable laws and regulations. "SupTech" refers to the use of AI and ML by financial regulators to support their supervisory, rulemaking, and enforcement efforts; see, "FinTech and

SecurityScorecard partnered with the New York Department of Financial Services (DFS) to modernize their oversight of financial services institutions operating in New York State through adoption of our AI/ML-based continuous monitoring and security ratings platform. This effort serves as a model not just for improving oversight, but to enhance collaboration with the private sector as well.⁴

A Model from the States - New York Department of Financial Services

Five years ago, New York State adopted a cybersecurity regulation commonly known as "Part 500," which establishes cybersecurity requirements for DFS-regulated financial services companies, including insurance, and banking, as well as cryptocurrency and student loan services.

As part of its oversight responsibilities, DFS conducts examinations of regulated entities to determine compliance with Part 500. DFS seeks to modernize the enforcement and compliance tools DFS uses to meet the dynamic cyber threats companies face. To do so, they looked to industry for solutions, specifically the insurance industry, which led them to security ratings.

Cyber insurers use security ratings, to (1) mitigate enterprise cyber risk, (2) manage third-party risk, and (3) maintain regulatory compliance. Ratings allow cyber insurers to detect vulnerabilities within the cloud and Internet of Things (IoT) environments and prioritize remediation activities for the most impact. They also allow insurers to manage third-party risk by centralizing, automating, and scaling third-party risk management (TPRM) activities and monitoring more suppliers without commensurate increases in security staffing and spending. Last, and importantly, insurers leverage AI and ML to automatically track compliance and minimize the risk of regulatory penalties as regulatory frameworks constantly evolve.

For regulators like DFS, AI and ML technologies like security ratings reframe the dynamic between them and their regulated entities into a more collaborative relationship. As described on DFS's website⁵ and discussed during its March 29

market structure in financial services: Market developments and potential financial stability implications," Financial Stability Board, Feb. 14, 2019, <https://www.fsb.org/2019/02/fintech-and-market-structure-in-financial-services-market-developments-and-potential-financial-stability-implications/>.

⁴ "Case Study: SecurityScorecard Helps New York State Department of Financial Services Ensure Financial Institution Cybersecurity Compliance," SecurityScorecard, Mar. 2022, <https://securityscorecard.highspot.com/items/6238bfcbf7532de6b47d839e?lfr=sfp.0>.

⁵ "Cybersecurity Resource Center," New York Department of Financial Services, accessed online on May 5, 2022 at, https://dfs.ny.gov/industry_guidance/cybersecurity#examinations.

Cybersecurity Symposium,⁶ DFS is using SecurityScorecard's cybersecurity ratings and analysis to assess the strength of the cybersecurity programs of DFS's nearly 3,000 regulated entities.

By providing continuous, outside-in visibility of every regulated entity, security ratings convert the regulatory "stick" (i.e., penalties and fines) into a conversation between DFS and regulated entities. For the regulated entities themselves, security ratings platforms - many of which, like SecurityScorecard's, are free to anyone for self-monitoring purposes - allow regulated entities to leverage the same groundbreaking AI and ML technology to "see what the regulator sees," and to communicate with the regulator using the same lexicon and objective data sets.⁷

Specifically, DFS identified several new tools to support their work. First, DFS uses third-party risk management and monitoring tools such as SecurityScorecard to provide DFS with a more informative starting point for their cyber examinations, as well as to create a more collaborative environment to work with covered entities to identify and mitigate cybersecurity risks. Security ratings are useful in settings like DFS' evaluations, Peterson explained, because they measure a large pool of data and enable an outside-in viewpoint to complement an inside-out viewpoint that DFS plans to collect via a questionnaire process.

DFS further identified that security ratings platforms help organizations remove blind spots, monitor and prioritize vulnerabilities, and better understand third-party supply chain risks. Indeed, NIST recently identified security ratings as a "foundational capability" for enhanced vendor risk management, in their recent update to Special Publication 800-161 r1, "Cybersecurity Supply Chain Risk Management Practices for Systems and Organizations."⁸

⁶ "The Unrelenting Cybersecurity Battle: Five Years of Evolving Threats and Controls," (Mar. 29, 2022), <https://meetny-events.webex.com/meetny-events/lr.php?RCID=646f83859cb3a77578aba1eefcb5d837>.

⁷ "Free Security Ratings," SecurityScorecard, <https://securityscorecard.com/free-security-ratings>.

⁸ "Improving the Nation's Cybersecurity: NIST's Responsibilities Under the May 2021 Executive Order," National Institutes of Standards and Technology (NIST), accessed online May 5, 2022, <https://www.nist.gov/itl/executive-order-14028-improving-nations-cybersecurity>.

“By combining the traditional exam data and incorporating the cyber risk tools such as the Cybersecurity Questionnaire (CIBRQ) and SecurityScorecard, we’ll have a more robust and current understanding of a regulated company’s cybersecurity posture,” Peterson said in his remarks. “This will enable us, as regulators, to make better informed supervisory and policy-making decisions.”

Continuous Monitoring with Security Ratings

DFS also leverages AI and ML risk management tools, including SecurityScorecard’s security ratings platform, to better track risk over time, and across the 3,000 entities that DFS regulates. The real-time risk monitoring that security ratings provide, both of covered entities and of their third-party risks, provides DFS with an outside-in, data-driven approach to complement and enhance DFS’s existing enforcement and compliance framework.

SecurityScorecard encourages the committee and Federal regulators to consider the DFS model using AI and ML technologies, like security ratings to support their supervisory rule-making, oversight and enforcement efforts.

Legislative Considerations

Since 2015, the Financial Stability Oversight Council (FSOC) has highlighted a need for congressional action to address the emergence of AI and ML technologies.⁹ SecurityScorecard believes the need to adopt AI and ML, and specifically security ratings, continuous monitoring, and third-party risk management, is critical to elevate the cybersecurity health of the financial services sector, and defend the nation against malicious cyber threats and malicious nation-state actors.

At the beginning of May, SecurityScorecard submitted comments to the Securities and Exchange Commission (SEC) in response to their proposed rule on cybersecurity risk management, strategy, governance, and incident disclosure for public companies.¹⁰ In our submission, we reviewed why AI and ML tools like third-party security ratings, continuous monitoring platforms, and third-party risk assessments are a cost-effective, comprehensive, and standardized way for organizations to assess and manage their cybersecurity risks, and why they are an increasingly necessary component of cyber risk management programs. We urged

⁹ Annual Reports, FSOC (accessed on Mar. 23, 2022); See also Oversight of Prudential Regulators: Ensuring the Safety, Soundness, Diversity, and Accountability of Depository Institutions, House Financial Services Committee (HFSC), 117th Cong. (May 19, 2021), and Cyber Threats, Consumer Data, and the Financial System, HFSC Subcommittee on Consumer Protection and Financial Institutions, 117th Cong. (Nov. 3, 2021).

¹⁰ RIN 3235-AM89. Cybersecurity Risk Management, Strategy, Governance, and Incident Disclosure.

the SEC to include security ratings as a cost-effective and efficient tool to support board oversight of cybersecurity, and to require covered organizations to employ continuous monitoring of vendors and supply chain companies as a cybersecurity best-practice.

We urge the committee to consider legislation requiring the use of continuous monitoring by publicly traded companies for third-party risk management and to disclose to their corporate boards and the Securities and Exchange Commission the company's cybersecurity risk ratings..

SecurityScorecard looks forward to working with the Committee as they consider ways to advance AI and ML adoption, and RegTech, across the financial services sector.

Thank you for the opportunity to provide this testimony.

