

Meeting 16 Notes and Actions
March 18, 2022

Next Meeting: ACDEB Meeting 17: May, 20, 2022

Meeting Agenda:

1. Meeting Introduction, Agenda Review, and Reminders
2. Thoughts from the Chair
3. Subcommittee Report: Other Services/Capacity-Building Opportunities
4. Subcommittee Report: Legislation and Regulations
5. Subcommittee Report: Technical Infrastructure
6. Subcommittee Report: Governance, Transparency, and Accountability
7. Subcommittee Report: Government Data for Evidence Building
8. Next Steps and Action Items

I. Meeting Introduction, Agenda Review, and Reminders – Emilda Rivers, ACDEB Chair and Alyssa Holdren, ACDEB Designated Federal Officer (DFO)

- a. Agenda Review:
 - i. Today is the first of three meetings to discuss the ongoing work of the five ACDEB subcommittees. Building off Year 1 Report, subcommittees have laid out agendas for next several months. Today they will discuss progress, plans, and priorities that build toward ACDEB's Year 2 Report due in October.
- b. Alyssa Holdren, ACDEB DFO, reviewed ground rules for member and public comments and questions.

II. Thoughts from the Chair – Emilda Rivers, ACDEB Chair

- a. Promise: The Evolving Data Ecosystem
 - i. There are many projects underway across the evolving data ecosystem, so what does this mean for the Committee's work?
 - ii. Different projects tackle core issues of data discovery, privacy protection, access, and use from different angles, so it is important for ACDEB to be concrete about plans and priorities for a National Secure Data Service (NSDS).
- b. Plans and Priorities: Year 2 Mandate
 - i. Acknowledging that ACDEB is part of an evolving ecosystem, ACDEB's Year 2 mandate is to leverage combined expertise and shape the path of evolution for the Federal statistical system by 1) guiding near-term actions and (2) fleshing out longer-term target areas it should be evolving towards.
- c. Plans and Priorities: Year 2 Roadmap

- i. Timeline is ambitious but doable with work already underway in subcommittees: recommendations approved by July, report drafted by September, and report delivered by October.
 - ii. To meet this timeline, ACDEB is having iterative conversations with OMB/ICSP, working on use cases, and engaging in virtual site visits with guest speakers.
- d. Plans and Priorities: ACDEB's Process
 - i. As a reminder and for transparency, a key step in Committee's process is "solicit initial feedback":
 - 1. Subcommittees draft recommendations that float through the Coordinating Committee and subcommittees before they reach the full Committee at a public meeting. This step in the process:
 - a. gives all members adequate time to review and provide feedback on recommendations before they are presented to the full Committee, and
 - b. creates a mechanism for identifying recommendations that can be approved through iterative discussions with the Coordinating committee, subcommittees, and those that require facilitated discussion at a public meeting.
 - ii. To meet schedule, a significant portion of deliberations occur at the subcommittee level and are then facilitated by the Coordinating Committee.
 - 1. All findings and recommendations will be reported out at public meetings with space left at public meetings to address issues that require input of full Committee in same place/time.
- e. Plans and Priorities: ICSP Conversations
 - i. At the January ACDEB meeting, OMB/ICSP kicked off an iterative conversation with the Committee that will continue over the next several months.
 - 1. Purpose is to provide ACDEB the opportunity to give feedback on existing initiatives in federal statistical system and use existing efforts to inform ACDEB's work.
 - a. These two pieces support the Committee's charge to advise OMB on CIPSEA 2018/Title III of the Evidence Act, both on here-and-now and target vision for an NSDS.
 - ii. In January, Bill Beach (ICSP) emphasized federal statistical system does not see implementing CIPSEA 2018 and building an NSDS as separate activities; they heavily overlap. By engaging ICSP, ACDEB provides timely, actionable, and

relevant recommendations for today that shape the evolution of the federal data ecosystem to support ACDEB's vision for an NSDS.

f. Plans and Priorities: Use Cases

i. Topics:

1. Topics for use cases originated in subcommittees and were reviewed/approved by full Committee.
2. Each use case is headed by a champion responsible for driving and tracking progress.
3. Proposals for use cases include implementation plans that outline timelines and key milestones in alignment with Year 2 Roadmap.

ii. Cross-Cutting Items:

1. Focus: current and evolving approaches to accessing, linking, and analyzing data across the federal, state, and local levels with consideration of how decision-making could be enhanced and facilitated; investigating improvements for current evidence-building ecosystem and weighing possibilities of an NSDS.
2. Rationale: Evidence Act with emphasis on Title III/CIPSEA 2018, Evidence Commission recommendations, ACDEB Year 1 Report including full Committee and subcommittee recommendations, ICSP workstreams.
3. Key Points:
 - a. Value of data access, linking, and analysis for evidence-building for data providers and users.
 - b. A wealth of potential data sources, including federal, state, local, private, official statistical products, administrative data.
 - c. Addressing barriers, challenges, and gaps such as legal and regulatory barriers, cultural resistance, differences in metadata, data quality, systems, interoperability, resource and capacity issues.
 - d. Gathering lessons learned and possible solutions; many opportunities for coordination within and across levels of government and with private sector.
 - i. Ex: data standards, consistency, interoperability; role of academic community, communities of practice, training,

resource sharing; technologies, tools, and advanced analytical methods.

- e. Privacy and confidentiality: privacy/quality tradeoffs, privacy protections, must comply with legal and ethical requirements.

- 4. Perspectives: have heard from specific subcommittees and expert groups.

g. Perspectives: The Lenses

i. ACDEB Subcommittees; for example:

- 1. Legislation and Regulations: looking at legislative barriers as well as challenges for breaking through them.
- 2. Governance, Transparency, and Accountability: governance frameworks, including mechanisms for ensuring transparency and accountability.
- 3. Technical Infrastructure: privacy enhancing technologies and technical infrastructure needed for systems to access and hold government data.

ii. Expert Groups:

- 1. Committee combines expertise that spans not just federal, but private sector, state and local parties, academia, privacy experts; the large scope of ACDEB's charge requires input from all of these perspectives.
- 2. Subcommittees mirror structure of full Committee with a cross-section of expertise in each of the five focus areas, so viewing use cases from subcommittee and expert lenses are compatible, not separate, activities.
- 3. Example perspective: ICSP Members (Statistical Agency Heads)
 - a. Highlighting this group because we are called out in the Evidence Act and given specific mandates and duties under Title III/CIPSEA 2018 and the charter of this Committee; beyond statistical official role outlined in Title I of the Act.
- 4. Other perspectives: ACDEB Experts
 - a. Encourage each ACDEB member to raise questions from their expert perspective when thinking about use cases with goal of harnessing combined expertise to make recommendations that benefit all.

h. Progress: Standard Application Process (SAP)

- i. In January meeting, learned about establishment of SAP as a requirement of Evidence Act, vision of its place in evolving federal data ecosystem, how it will serve as a unified front door for user access.
- ii. In virtual site visit, learned about implementation taking place, policies developed to support implementation, expected value that transparency will bring to federal agencies, how SAP engages stakeholders (like state/local and role ACDEB can play in facilitating engagement), and specific request for ACDEB's advice on SAP.
- iii. 3 ICSP Workstreams around SAP happening in the subcommittees: governance, technical development, and technical assistance.
- iv. OMB/ICSP wants the Committee's advice, so iterative conversation serves as a way to provide input on the SAP and flesh out vision for next steps that ultimately support the functions and services of an NSDS.
 - 1. For example, Governance group might want to engage how SAP should align with NSDS; ACDEB state and local representatives might want to know more about implications of how their data is used and advise how to inform policy development for SAP going forward.
- v. OMB/ICSP engagements can contribute to findings and recommendations for Year 2 Report as part of how ACDEB/NSDS fit into evolving data ecosystem.
- i. Legislative Underpinnings for an NSDS:
 - i. Recent legislation pieces are building framework for target NSDS; we expect this piece to evolve, too. SAP and NSDS legislation are just two examples of evolving data ecosystem.
 - ii. Important to keep moving and harness energy around these efforts in order to shape path toward evolution. Encourage all members to continue to engage in ICSP conversations, contribute to use cases, use wealth of information to develop findings, draft recommendations, provide input/feedback—all in order to yield a timely, actionable, and relevant Year 2 Report.
- j. **Committee Feedback/Discussion:**
 - i. **Question:** important to include all stakeholders in discussions of an NSDS and concerned that some conversations are already under development without participation of state/local representatives. How has ICSP, in developing the SAP, included state/local perspectives as policies are being developed?

Answer: ICSP Stakeholder Engagement group is part of SAP process; at stage in process where focus has been on getting started so people have something to provide feedback on; comment is being sought from the public (FRN, email

blasts, and other requests for engagement); also provides an opportunity for ACDEB to point out how this state/local engagement could look different.

Answer: work being done now is a foundation for what SAP could be; FRN is specifically asking about structural mechanisms in governance, for example; have scheduled conversations with Committee members for diverse feedback across stakeholder groups (ex: conversation with Other Services about improving engagement strategy); many opportunities for feedback to improve SAP as it evolves, especially if there is additional funding.

Summary of chat comments: comments on state/local involvement in functionality and development, emphasizing request that state/local groups not just be involved in feedback but also be partners in the process of development; integrating state/local as partners and getting state/local input early in process prevents impression that state/local is an afterthought; importance of partnership since state/local are often originators of data to federal government.

Response: agree with importance of state/local governments and would like more specificity in terms of who to reach out to for advice from state/local perspective; how to help these federal initiatives in the data ecosystem reach right people in state/local governments? Some of this work is being done within ICSP and other agencies; had volunteer from Committee who wants to talk with ICSP on who should be involved from state/local.

- ii. **Comment:** Agree that considering resources is crucial; thinking about overlap between SAP, NSDS, and broader data ecosystem—how even the most advanced version of SAP progresses the vision for NSDS and what are additional tools needed; always a challenge to getting fast, high-quality data, but flexibility can be imbedded within vision of an NSDS; important that not all solutions are filtered through SAP, even in its most mature version.
- iii. **Question:** Did current version of SAP factor in ACDEB feedback from the Committee's first year? Need to reiterate feedback from first year or provide feedback beyond what was already shared?

Answer: Important to get feedback from ACDEB as things evolve, including with the SAP; important to stay current with what options are.

Answer (from chat): Because of the contract requirements for getting SAP 2.0 off the ground, the technical specifications were pretty far along before the ACDEB was stood up; its design was informed by feedback from those who used the version that the FSRDC produced.

Answer: this first version of the SAP is bare bones, so looking forward to getting targeted advice on rollout and governance of current 2.0 version and also to help develop specifications for where it goes next and what would be helpful for meeting needs of state/local/tribal governments; thinking about what resources are necessary for building capacity on user side and on agency side to make sure vision can be implemented.

- iv. **Comment:** General point about encouraging federal agencies to see that there is a White House statement from last year encouraging Indigenous and tribal knowledge in policy formulation and processes.
- v. **Comment:** SAP was mentioned as an example and is a good one; recently celebrated release of discoverable metadata; hope is that ACDEB uses that site as part of making comments on how to move forward.

Response: Excited that SAP has some metadata in it, but would like to stress it is dataset metadata and agencies need resources to get element level metadata and more information about provenance to help people figure out what data are available for evidence-building; have to have resources in order to have inputs ready for an NSDS.

Response: yes, and dovetails with fact that this is a work in progress and with resources will be another version.

- vi. **Comment:** Is there a particular subcommittee tasked with generating feedback for SAP? Or are they submitting feedback through traditional public submission?

Response: Several avenues and mechanisms (FRN, ICSP Stakeholder Engagement Group).

Response: partitioning SAP into segments: one in Governance, for example, with idea that once conversation is initiated in subcommittee, others from ACDEB will be pulled into conversation.

Response: 3 ICSP workstreams around SAP for subcommittee engagement: governance, technical development, and technical assistance.

k. **Discussion Wrap-Up:**

- i. Excited about what SAP is doing in terms of discovery role with data, access, and streamlining processes there. Excited for what we have accomplished with discovery feature (metadata).
- ii. Many federal agencies have longstanding relationships with state/local governments (Bureau of Labor Statistics, for example) to encourage conversation; this is an ecosystem where we will identify gaps in who we are

speaking with and then will be able to look comprehensively for who to reach out to more.

III. Subcommittee Report: Other Services/Capacity-Building Opportunities – Kim Murnieks, Ohio Office of Budget and Management, and Matthew Soldner, Department of Education

- a. Two use cases:
 - i. Communications-focused use case, necessary to convey value proposition of an NSDS to ensure service is useful and used.
 - ii. Technical-assistance use case that builds on data concierge notion from Year 1 Report and other services needed to help folks take advantage of an NSDS.
- b. Strategies:
 - i. Fact-finding; using first part of Year 2 to run down constituencies and resources needed to build recommendations.
 - ii. Developing personas that represents archetypal service users that will help pressure-test the range of technical assistance services to recommend making available through an NSDS.
 - 1. May ask other ACDEB members to engage in personas exercise in order to more fully brainstorm range of folks who may bring needs, questions, concerns to an NSDS.
 - iii. Summary: focused on ensuring NSDS recommendations are well-known, well-understood (communications focus) but also well-used (technical assistance); there are organizations out there doing a good job communicating about data and providing technical assistance, so we want to learn as much as possible from ACDEB members and the public about what exists.
- c. Findings:
 - i. Started with communications use case and benefitted from virtual site visits from Data Quality Campaign and Results for America.
 - ii. Results for America produced concrete recommendations for a communications strategy for an NSDS.
 - iii. Data Quality Campaign provided ways to think about specific artifacts to produce thoughtful communications materials for federal/state/local.
 - iv. Looking forward to synthesizing these materials and hope to bring communications recommendations to full Committee in next meeting.
- d. Next Steps:

- i. In April, will turn attention to data concierge and technical assistance, including SAP conversations with ICSP and fact-finding meetings with outside organizations; if Committee members have questions they would like included in these discussions, send to co-chairs and support team, and feel free to join these sessions if available.
- e. Discussion Questions:
 - i. Are there individuals or organizations missing that the full Committee would recommend for consultation?
 - ii. Are there things others are learning about communications or technical assistance in subcommittees that Other Services should be aware of?
 - iii. Wrap-up: have talked about communications, technical assistance, and engagement with state/local leaders to ensure value proposition is there for state/local (submitting data and using data for policymaking); key part of Other Services conversation is ensuring connection and usefulness of NSDS.
- f. **Committee Feedback/Discussion:**
 - i. **Comment:** when talking about concierge service and technical assistance, it would be great if this extended not just to data that is protected but to publicly available data; NSDS should help create a more data-driven society so there is help for all levels of sophistication (ex: NOAA makes curriculum available, including how to use NOAA data and what data are available for analysis).

Response: great point; subcommittee has discussed broader general public as stakeholder for both communications and technical assistance.
 - ii. **Question (from chat):** many federal agencies have Offices of Congressional and Intergovernmental Affairs (OCIA); do you think a similar group would work well for an NSDS?

Answer: question becomes for what remit? Short answer is yes, but if NSDS actively engages folks at all levels of government, will need to think about ways to go out and be effective for engaging those partners; please send specific ideas to Other Services.
 - iii. **Question (from chat):** could we use state/local experiences with such OCIA personnel to gauge how effective they are in engaging those partners?

Answer: absolutely.
 - iv. **Comment:** part of concierge service, if properly resourced, would also get at core of comment about element level metadata or understanding. While

concierge might not have expertise at dataset level, would have connector to someone who would—NSDS provides “connective tissue.”

- v. **Comment:** while listening to data concierge conversation, thought about South Carolina virtual site visit and how they talked about building knowledge so there was feedback or “two-way street” between Integrated Data System and other agencies; useful to consider as example.

Response: assume every agency benefits from such a feedback loop, opportunities to improve quality, think about use cases, new research directions.

- vi. **Comment:** would like to suggest dedicated concierge is useful but limited; could think more about how to develop community curation of data. When dealing with lots of administrative records, one person can’t act as “telephone operator.” There are models of community curation in computer science (ex: Stack Overflow, Kaggle competition). Suggest thinking about communities of practice and right incentive structure for providing feedback.

Response: mentioned support and development of communities of practice as role of data concierge in Year 1 Report; can build more on that and think about incentives.

- vii. **Question:** is there a way to see how can we improve those engagements that federal statistical system currently has in place with state/local partners?

Response: take offline as part of larger discussion.

g. Discussion Wrap-Up:

- i. Thank you to Other Services; highlights ecosystem—many pieces, including state/local conversation to continue offline in order to expand engagement so all voices are heard.
- ii. Communications conversation connects to SAP; good to see discussion on the components of how this connects to other parts of data ecosystem.
- iii. One option is phased approach; how to develop a system where you can build on things particularly as we all look for resources to get this incredible amount of work done.
 - 1. Response: crowdsourcing is less costly than expense of data concierge.
 - 2. Response: the more we field-trip the larger the potential technical assistance/user experience remit grows; a resource issue at some point.

IV. Subcommittee Report: Legislation and Regulations – Christine Heflin, Department of Commerce, and Nicholas Hart, Data Foundation

- a. Use case proposal: use case comes from points of view of Performance Improvement Officer and Evaluation Officer; using administrative data to track project impact, evaluate program impact, and track recovery and economic progress (American Rescue Plan, Infrastructure Investment and Jobs Act).
 - i. Typically use projections and then over the course of the life of the project there are reports; reporting burden can be significant and can vary; so, question becomes can we use administrative/statistical data for a few different purposes? Department of Commerce is exploring this question now.
 - ii. Can we use administrative and statistical data to look at a particular project's impact? Can it be used to evaluate the impact of a project and, if so, how? And lastly, issue with tracking economic progress in specific areas—challenge is lack of frequent and granular data, so can we track and eventually evaluate projects and, if so, how? What are constraints?
- b. In general, for all use cases, request that Committee members point out legislative and regulatory barriers to achieving a given use case and send to this subcommittee.
- c. Preliminary findings:
 - i. Had conversation with folks who have used administrative data to evaluate program impact; Department of Labor on Payroll Protection Program; suggested admin data could be useful for program evaluation.
 - ii. Could have a county or state recovering well, but Census tracts help identify pockets that get left behind; could create a national dashboard of Census tracts to show pace of recovery.
 - iii. Subcommittee wants to explore what it would mean to look at programs with states with data lakes; form partnerships and track impacts of programs together.
 - iv. Lastly, would it be possible to get a waiver from a business to use data that wouldn't typically be available given the regulations involved (ex: Small Business Administration)?
- d. Next Steps:
 - i. OMB guidance: how those guidelines might facilitate this kind of process for general use of administrative/statistical data for tracking and evaluation.
- e. Discussion Questions:
 - i. Wish there was a lawyer on subcommittee; folks at Census Bureau advised there is a white paper on legal issues associated with sharing data; question for

group: if anyone has white papers, legal briefs to share, please do; trying to widen spectrum of views on that.

- ii. Any potential partners interested in exploring federal and state data to assess impact?
- iii. Exploring dashboard further; should public investments be tracked by Census tracts?
- iv. Problem of resources generally; Census considers resources dedicated to data sharing a significant barrier to timeliness; how can resources be made available?
- v. Wrap-up: subcommittee commentary about training activities, what does it mean to have conversations with lawyers around data access under existing law, how do we think about partnerships with nonprofit entities that governments can engage with, what are those kinds of levers we can pull? Welcome thoughts on training and education (a new area we've started exploring).
- vi. Updates coming soon:
 - 1. Subcommittee is meeting Monday with OMB/ICSP on Trust Regulation.
 - 2. NSDS Act passed House in bill called "America Competes" and expect movement in coming weeks.

f. Committee Feedback/Discussion:

- i. **Comment:** Paycheck Protection Program used ADP data, not administrative data

g. Discussion Wrap-Up:

- i. Thank you to Legislation and Regulations; particularly excited about how ACDEB can contribute to legislation and regulations discussions, how we can partner with people across the legal community; example of how evolving ecosystem is progressing (legislation moving forward).

V. Subcommittee Report: Technical Infrastructure—Amy O'Hara, Georgetown University, and David Park, National League of Cities

a. Framing:

- i. See NSDS as facilitating connection to data held where they are.
- ii. Clarify "infrastructure" as hardware and software in administrative and control systems that would be needed for entire user lifecycle when interacting with NSDS.
- iii. Need for NSDS to do things for users as well as to provide a critical research and development role for the ecosystem in order to bridge divide between what

CDOs and Statistical Officials are doing, including roles of fact-finding, myth-busting.

- iv. Hope is research and development component works on new methods, improving existing methods, and advancing ability to make informed decisions.

- 1. Ties into capacity development, tools to aid that, and hope that's part of technical infrastructure of an NSDS.

b. Use Cases (four broad categories):

- i. Looking into infrastructure that would support capacity to have federated data analysis; includes federal, state, local data streams as well as proprietary data.
- ii. Focusing on input and output privacy solutions; looking into privacy-preserving technologies (ex: synthetic data with validation servers, technical builds for formal privacy solutions).
- iii. The need to evaluate threat models, risk evaluations; what different data or different data projects need to consider when implementing privacy solutions, could bridge into how to monitor virtual access (tools out there to support such things).
- iv. Metadata as critical to a user's awareness of what data are available for evidence building, then to understand what they could request, and ultimately to understand quality and coverage characteristics of data; ways to automate metadata development or data inventory management, including interoperability between different organizations that have their own inventories.

- 1. Some of that will happen within the SAP but some will not; looking at state data sharing initiatives to see what's working and what is not.

c. Next Steps:

- i. Speaking to experts and continuing to review the literature to have information to shape recommendations; have scheduled meetings with experts through May.
- ii. Then coming back to Coordinating Committee with draft recommendations so they can be folded into Year 2 Report or discussed more broadly at next public meeting.

d. Discussion Questions:

- i. Will NSDS hold data (beyond pass-through for linkage)? Opens up another set of questions; not looking for a solution, but wanting to make sure we're describing the right capabilities to inform folks who take next steps on NSDS.

- ii. Will NSDS host projects? Place where people log in and access data? May be accessing data where they live at home, or with first bullet accessing data that are stored or managed locally.
- iii. Will users be able to contribute metadata? Just brought up in previous session. Have data sources that are constantly evolving, so as datasets are changing, can't expect one person or set of data concierges to be on top of that; state and local data providers are the experts, so how to take advantage of where the expertise is?
- iv. Will NSDS facilitate code-collaboration, code-sharing? Related to point above—if you have someone acknowledge anomaly and come up with way to address it, how would you be able to make sure that code would be available for others? Or have that community of practice for people to understand different ways of handling different shortcomings in the data?
- v. Comment:
 - 1. From local perspective, just coming off National League of Cities spring conference and heard a lot about staffing; this can't be everything to everyone, people you might want to talk to aren't there because you can't hire them.
 - 2. Second piece is driving decision-making and policymaking; talked about longer-term projects (months, years) but lots of folks on the ground need timely data and are looking for a solution in a matter of days.

e. Committee Feedback/Discussion:

- i. **Question:** holding data as fundamental question—can data be in their own cloud server, can I do a project where every time I want to do a run on data I have to re-link everything? Cloud isn't always fast enough to process; important question where the answer is a function of the technology.

Answer: in subcommittee meetings, continually discuss the lack of capacity where many datasets are generated to stand ready every time someone wants their data; they don't have a separate research cloud waiting for NSDS or academics; would be a burden on state/local agencies that are generating data streams and it's unclear where resources and technical capacity would come from.

- ii. **Comment:** need better data to drive day-to-day operations—very important point from the chat; one challenge is delay in processing and using data; timeliness is importance piece of quality and value; suggest using summary tabulations (Midwest Collaborative), portfolio of options (synthetic data, secure remote access, broad group of users who have access to tabulations that are at

a lower suppression level and they sign an agreement that says not to re-identify)

Response: do you think NSDS holds data?

Answer: think it has to because you can't do data linkage without it; quality of data linkage can be very "noisy," especially for marginalized people (ex: individuals with double last names, Latinx community, people coming out of Corrections, when trying to match based on zip codes for mobile population (lower income, transient, younger))—so many ways in which failures to have data in same place means quality of linkages will be very suspect and need other sets of eyes on them; would argue same for confidentiality protections; would like to see analysis of impact of different data protection techniques take place before the data are released inside the NSDS in conjunction with the people whose lives are going to be impacted as a result of those different types of techniques applied, and then once impacted populations agree on tradeoff and are aware of it, then get release occurring.

Response: great aspirational model, but trying to reconcile it with need for quick data release?

Answer: referring to synthetic data and different types of technical approaches (not summary tabulations).

- iii. **Question:** thoughts on whether or not NSDS will host software that lets users search metadata (similar to automated concierge service)?

Answer: would be great; great to test it with SAP and determine what needs to be built as a wraparound service or bridge to data that are not from statistical agencies.

Comment (from chat): would be a great intersection to Title 2 of the Evidence Act, if OMB issued its guidance. T2 requires the comprehensive inventories to include metadata. So that is a more global question about government than just NSDS.

- iv. **Comment (from chat):** (1) state/local government officials are desperately aware and in need of better data to help drive day-to-day operations as well as policy-making. We are all looking for more ways to rapidly scale and apply solutions to meet citizens needs and expectations, and (2) the price point for such technology and process improvements are not easily funded and we often cannot generate revenue off of providing data to those who request it.

- v. **Question (from chat):** A related question is: will the NSDS permanently host data or just host it long enough to support a project and perhaps for a period of time for validation or replication?

Answer: further refinement of earlier question posed to the Committee; if NSDS has the right protective technologies in place, it could be a safer place to provide access than many of the agencies; thoughts from perspective of Commission?

Response: that was a goal of the Commission, if there was more unified structure, when access to data would be given and how they were released would improve overall security of data; acting as a whole rather than disparate parts; valid point to keep track of.

f. Discussion Wrap-Up

- i. Thank you to Technical Infrastructure; a good point for us to think about what advice to give to OMB about regulations (presumption of accessibility, access regulation); what kind of opportunity do we have to shape advice to OMB?
- ii. And, as we think about how to answer such a question, considering categories: how long to hold or store data? Validation and replication? Metadata and moving beyond a dataset? What agencies actually have?
- iii. Ecosystem perspective when thinking about answers to these questions: through presumption of accessibility, stat agencies have ability to acquire, hold, and link data. If agency can collect data, how to ensure component pieces speak to each other? NSDS as service, philosophy, and place—helps crystallize where we need to be specific about NSDS and what it will do.

VI. Subcommittee Report: Governance, Transparency, and Accountability – Charles Cutshall, Commodity Futures Training Commission, and Julia Lane, New York University

- a. To date in Year 2, Governance has been diving into specific use cases to further explore and refine recommendations and attributes included in Year 1 Report; use cases cover 3 areas: Education and Workforce, Health Statistics, and Labor Market Activity
- b. Education and Workforce, Gregory Fortelny
 - i. Ability to Benefit program allows students without a high school credential who are enrolled in a career pathway program to be eligible for Title IV federal student aid, provided their program enables them to obtain a high school credential.
 - ii. Accessing, linking, and analyzing educational workforce data in this space provides a number of federal, state, and local benefits, as well as benefits to learners themselves.

- iii. Accessing, linking, and analyzing data requires a variety of sources: federal data from Department of Education and IRS, state data from longitudinal data systems, and a variety of non-government sources.
 - iv. With that variety of data sources comes a variety of challenges: data governance and privacy, mosaic of federal and state laws, privacy/quality tradeoff.
- c. Health Statistics, Brian Moyer
- i. Use case primarily focuses on CDC and National Center for Health Statistics (NCHS) efforts to modernize National Vital Statistic System; collaborative effort between federal government and all 57 state and jurisdictional vital records offices.
 - ii. Effort seeks to improve not only the quality and timeliness of vital data through enhanced data standards and data interoperability but also seeks to improve how the data are shared, both at federal and jurisdictional levels, for purposes of evidence building.
 - iii. In terms of learning agenda, proposing to engage with CDC and NCHS modernization teams working on this effort, also proposing to engage with state vital records offices and other state-level stakeholders.
 - iv. Focus for governance is 1) better data sharing and linking among states, 2) value of better data sharing and linking between federal government and states, and 3) enhanced communication, training, and technical assistance through robust community of practice that has already been stood up and has proven to be successful.
- d. Labor Markey Activity, Julia Lane
- i. States have the ability to capture very granular data on employment and unemployment activity using administrative records, so the potential for using information about unemployment wage records is vast.
 - ii. Advantage for national statistics and state/local information sharing, decision-making, especially in tracking flows of workers across states.
 - iii. Advantage of sharing information in a secure environment, improve information about employment outcomes at a granular level in a way that gets information to the people who are going to use it (workforce boards, governors' offices).
 - iv. Potential to go from claims to claimants to cohorts: unemployment insurance data has very detailed information on a range of categories, so can see differences in duration of unemployment and quality of re-employment at weekly frequency.

- v. Challenge being faced is potential to be blocked because of uncertainty about legislative restrictions, so NSDS can provide clarity on what is permitted and what is not; privacy issues, how privacy should be operationalized and institutionalized, and privacy/quality tradeoff.
 - vi. In general, these use cases are trying to highlight the value proposition of horizontal and vertical data linkages; can generalize from these use cases to broader use.
 - e. Overarching Findings (preliminary, not final recommendations)
 - i. Articulating value proposition for key stakeholders is foundational; important to extend beyond state/federal institutions to include people/organizations.
 - ii. NSDS must itself provide products and best-in-class services to facilitate use of data for evidence building; ties in well with input/output conversation.
 - iii. Challenge of navigating federal and state laws, regulations, and policies; NSDS should identify requirements and provide service that meets requirements.
 - iv. Efficiency and innovation as critical to success; NSDS should provide efficient avenue to access data without adding bureaucratic burden; NSDS must offer clearly defined standards, frameworks, process.
 - v. NSDS should offer diverse privacy protecting services while itself epitomizing an organization with utmost reverence for privacy of individuals.
 - vi. NSDS must be public-private partnership; federally owned but privately managed.
 - vii. Plan to present these preliminary findings and draft recommendations as soon as feasible to Coordinating Committee, then circulate to full Committee.
 - f. **Committee Feedback/Discussion:**
 - i. **Comment (in chat):** <https://coleridgeinitiative.org/national-convening-march-2022/>
 - ii. **Comment:** question on public-private partnership: good idea to enlist help from private sector, but concerned there are actors in private sector trying to monetize technologies; need to be aware that any technology that is developed is part of the public domain to be widely shared, widely used, and tested for reliability/safety by researchers.

Response (in chat): need to balance engagement from private sector given eagerness to gain access and monetize, which puts government at disadvantage when we cannot generate the type of revenue needed to really make a

difference with modernizing our technology and processes and sustaining innovation going forward.

Response (in chat): agreement and disagreement; need to consider both private and public sector concerns.

Response (in chat): agree that there are public sector concerns, but concerns are different. I don't think the public sector is likely to try to monetize new technology whereas the private sector has a strong incentive to do this. Also, acknowledging potential conflict of interest, there might also be a role for nonprofit research institutions (the semi-private sector?).

- iii. **Comment:** 1) need to be conscientious of not only linking current data but thinking about missing datasets; from agency perspective, we're recruiting industries, so how do you layer that information so you have real-time labor market information in addition to linking the data; 2) when thinking about the SAP, not all states ask same claimant information, so is there a case to be made for having states collect same information? Lack of mandatory standards makes sharing data a challenge.

Response (in chat): that is one way in which federal agencies can help with providing standards.

g. Discussion Wrap-Up

- i. Thank you to Governance presenters; provided a lot of information that puts ACDEB in good position to work through the process and to come up with relevant, timely, and actionable recommendations.
- ii. Hearing about use cases gives ACDEB a framework to build on, so recognizing through these use cases that there is a lot of existing information/approaches already; can use these use cases to think about gaps, how to meet needs, how to balance private sector interests in larger ecosystem.

VII. Subcommittee Report: Government Data for Evidence Building—Kenneth Troske, University of Kentucky, and Anna Hui, Missouri Department of Labor and Industrial Relations

- a. Government Data is focusing on four use cases: South Carolina Data Sharing Program and Data Inventories, Environment-health (PFAS), and building onto Labor Market Activity and Health use cases already spoken to in Governance subcommittee.
- b. South Carolina data inventory and metadata use case, Elisabeth Kovacs:
 - i. The data that is produced can only be as good as what is captured.
 - ii. South Carolina has used templates and surveys to try to determine what data elements are captured and how much data is captured (for example, SC only

captures 5 elements while other states capture up to 24) as a way to think about standards.

- iii. The other model is JEDx who presented at a virtual site visit, which captures data through data inventory related to labor market activity, so this can serve as a guide for best practices.

c. Environment-health (PFAS) use case, Richard Allen:

- i. PFAS is a broad category of chemicals (over 12,000) that are persistent in environment, build up in human and animal tissue through drinking water and consumption.
- ii. Physiological effects are well-studied with substantial data on concentration levels in drinking water: currently monitor concentration levels of 6 chemicals and proposal to monitor additional 20+
- iii. Collect data from states who are important partners because they do their own monitoring at finer geographic resolution, and federal partners would be EPA and CDC, as well as other federal agencies with equities.
 - 1. Benefit of proposing this use case is to see how NSDS could reduce friction in sharing these human health and environmental conditions data.
 - 2. Better leverage national system of data collection and data sharing, and combine these data with state data and expertise from academic institutions, then can drive a more complete understanding of these chemicals.
 - 3. NSDS serves as a community builder and resource for collecting disparate data.

d. Findings:

- i. SC has been successful in developing a data inventory system, but less successful at sharing data.
- ii. University of Florida has used SWIS data to track graduates who have moved to other states; not sure if other Universities are doing the same.

e. Next Steps:

- i. Getting approval for SC and PFAS use cases.
- ii. Speaking to experts: SWIS virtual site visit, which sounds like NSDS with SSN matching; look at governance, MOUs).
- iii. Review JEDx presentation.

iv. Update on how to modernize UI system:

1. DOL put out grant opportunities to states, Midwest Collaborative looked at grant opportunities, thought about a number of potential demographic aspects to get more information on, and wrote same language shared among states to incorporate into grant requests they are making; an opportunity where state/federal engagement could prove fruitful.
2. Another example of states' modernization efforts and federal funds to help with those efforts is thinking about how to leverage consortium/collaboration; Missouri is leader of Innovate UI consortium based on concept of "cousin" states and sharing code among these states; there is a baseline level of code that could be jointly maintained in shared environment going forward and how to better utilize data aspects we engage for federal reporting for state purposes.

f. Discussion Questions:

- i. Main question focuses on challenge of many different, disparate efforts at federal/state levels, as well as in private sector, to combine efforts. Year 1 report asserted that coordination would be critical function of NSDS, so how can NSDS support coordinating efforts?

g. **Committee Feedback/Discussion:**

- i. **Comment:** PFAS use case also aligns with options presented by BPC (see #13 & #17) for EPA on some data integration activities that syncs well here:
[https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2020/02/Strengthening Transparency and Accountability at EPA.pdf](https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2020/02/Strengthening%20Transparency%20and%20Accountability%20at%20EPA.pdf)
- ii. **Comment:** Wondering about creating a couple of common fields with same syntax for state and federal; facilitating linkage by proposing common fields for data sets?

VIII. Next Steps and Action Items – Emilda Rivers, ACDEB Chair

- a. Thank you to Government Data; balance is necessary when thinking about public and private sectors as well as decentralized federal system.
- b. CIPSEA 2018, federal laws and regulations, examples like the SAP all show how we can pull together to come up with standardization that helps us work toward goals that helps us be seamless to users while meeting our individual missions.
- c. Federal statistical system is also moving in evolution of data ecosystem that is beyond disjointedness; always room for improvement to ensure progress.

- d. Transparency report is available on CNSTAT site that gets into metadata and data standards that can provide food for thought on standardizing in that space.
- e. Excited about what we heard today in terms of being in a good position for July meeting where goal is to have all recommendations submitted, reviewed, and approved by full Committee to fullest extent possible.
 - i. Practically speaking, this means no new recommendations at that meeting; at latest, draft recommendations should be submitted to Coordinating Committee by July 11 so we can process wealth of information and vet it to ensure inclusion in full report.
- f. Next Meeting: ACDEB Meeting 17: May, 20, 2022