

July 8, 2019

Via Email

Re: Data Coalition Comment on the Federal Data Strategy Draft Action Plan (USBC-2019-0001)

Dear Mr. Droegemeier, Ms. Kelley, Ms. Kent, Ms. Potok, and Ms. Roat:

Thank you for the opportunity to provide comments on the Draft 2019-2020 Federal Data Strategy (FDS) Action Plan. The Data Coalition is America's premier voice on data policy and advocates for government information to be high quality, accessible, and useful. In most cases, this means transforming data to be standardized, open, and machine-readable information. Based in Washington D.C., the Data Coalition members include a broad cross-section of the data industry, including technology and data analysis companies, public sector consulting firms, and non-profit organizations.

As we indicated in past comments and public statements, the Executive Branch's efforts to develop a transparent and coordinated approach for developing the Federal Data Strategy are to be applauded.¹ The Data Coalition appreciates the incorporation of some of our comments and those of other stakeholders throughout the development of the strategy this past year.

While there is much promise in the current strategy, the draft action plan misses important opportunities to promote effective implementation over the next year of the Foundations for Evidence-Based Policymaking Act, the OPEN Government Data Act, and the Confidential Information Protection and Statistical Efficiency Act.² The absence of intentional actions that focus on key provisions of the Evidence Act is a major shortcoming of the draft strategy. The Data Coalition strongly recommends that the administration incorporate additional action items to reflect priorities for the data community in new legal requirements established with bipartisan support by Congress and signed by the president in January 2019. Our attached

1 Nick Hart. The Federal Data Strategy is a Trump Plan We Should All Support. NextGov, July 2, 2019. Available at: <https://www.nextgov.com/ideas/2019/07/federal-data-strategy-trump-plan-we-should-all-support/158153/>

2 Foundations for Evidence-Based Policymaking Act of 2018. Public Law 115-435.

comments include targeted suggestions to ensure the Action Plan aligns with the Evidence Act, among other recommendations.

The Data Coalition and its members look forward to continuing to support the development and implementation of this important strategy. Should you have additional questions about our comments, please contact the Data Coalition's Senior Director for Policy, Christian Hoehner (christian.hoehner@datacoalition.org).

Thank you again for the opportunity to provide feedback.

Regards,

A handwritten signature in blue ink, appearing to read "Nick Hart", with a long horizontal flourish extending to the right.

Nick Hart, Ph.D.
CEO
Data Coalition

DATA COALITION COMMENTS ON THE FEDERAL DATA STRATEGY YEAR 1 ACTION PLAN

1 - RECOMMENDATIONS FOR NEW COMMUNITY ACTIONS

ESTABLISH A GOVERNMENT-WIDE ORGANIZATIONAL IDENTIFIER WORKING GROUP

The federal government currently uses more than 50 organizational identifiers for companies, nonprofits, and other non-federal entities.³ Because the same entity, or organization, is identified differently by multiple federal agencies, entity matching can be excessively time consuming and resource intensive for developing regulatory analysis, implementing regulations, conducting investigations, implementing contracts, or collecting information for oversight and accountability purposes. The U.S. government previously conducted a working group to study the issue, though little progress has been made to adopt common identifiers over the past decade, notwithstanding global efforts -- in conjunction with some U.S. federal agencies -- to develop, adopt, and implement non-proprietary identifiers. For example, the Legal Entity Identifier (LEI), is a global standard that could be further considered for application across federal agencies alongside existing non-interoperable identification codes.⁴

In 2016 OMB convened the Employer Data Matching Workgroup which included fourteen representative federal agencies. The working group was charged with documenting the issues in “matching and uniquely identifying establishments and firms within and between data sets and over time,” identifying potential solutions to these challenges, and making recommendations to “improve the Federal government’s ability to identify and match unique firms and establishments (and the relationship between the two) within and across Federal data sets.”⁵ The working group also produced an inventory of various agency data resources

³ Duncan, J. A Jungle of Entity Identifiers and What We’re Going to Do About It. Washington, D.C.: Data Coalition, 2014. Available at: <https://www.datacoalition.org/a-jungle-of-entity-identifiers-and-what-were-going-to-do-about-it/>

⁴ Berner, Richard. Breaking Through Barriers Impeding Financial Data Standards. Washington, D.C.: Office of Financial Research, 2017. Available at: <https://www.financialresearch.gov/from-the-director/2017/02/02/breaking-through-barriers-impeding-financial-data-standards/>

⁵ Employer Data Matching Workgroup White Paper. Washington, D.C., 2017. Available at: https://www.cep.gov/comments/Employer-Data-Matching-Workgroup_White-Paper.pdf

used for identifying and monitoring U.S. employers.

The FDS should include an Action Item which leverages and extends the initial steps to improve common entity identifiers. This could include an additional study on the disparate identification codes currently in use and a feasibility study on a full transition to a common nonproprietary federal entity identifier such as the LEI.

2 - RECOMMENDATIONS FOR EVIDENCE ACT IMPLEMENTATION

The draft Action Plan should take steps to address core requirements and provisions of the Evidence Act, including a number of provisions that appear to be omitted altogether from the draft Action Plan.

PROVIDE GUIDANCE FOR AGENCY-BY-AGENCY CDO DESIGNATIONS

The deadline in the Evidence Act to identify new Chief Data Officers is July 13, 2019. To date, OMB has not yet issued implementation guidance for how agencies should approach appointments, staffing, or infrastructure related to what will be a new role in more than 100 federal agencies. The Data Coalition strongly recommends that OMB prioritize the issuance of guidance to implement this important role, along with the corresponding enhancements the designated leaders will provide for managing privacy risks and improving data quality, access, and use. In particular, the Coalition recommends that the guidance clearly delineate and reaffirm the Evidence Act's intent that, generally speaking, CDOs serve as officials that are distinct and separate from Chief Information Officers and Chief Financial Officers. Guidance would also be useful for agencies seeking insights about how to appropriately resource the new function, including:

- Staffing and training an appropriate workforce;
- Identifying and sharing models and best practices for CDOs to acquire analytic platforms and tools
- Organizing the CDO to best fit within appropriate organization structures, including where FITARA funding may be appropriate for supporting implementation
- Further delineation of the eligibility requirements for CDOs, as explicated in the OPEN Government Data Act.

ESTABLISH AND EMPOWER THE CDO COUNCIL

In addition to guidance on the CDO positions and roles, we strongly recommend OMB include an Action Item addressing the establishment and empowerment of the statutory CDO Council.

In particular, the CDO Council's roles should be appropriately connected to the proposed OMB Council and the agency Data Governance Committees.

ISSUE GUIDANCE ON THE ESTABLISHMENT OF EVALUATION OFFICERS AND STATISTICAL OFFICIALS

Title 1 of the Evidence Act includes provisions that aim to bolster government's capacity to use data and to engage in evidence-based policymaking. The Draft Action Plan should include actions that recognize the need to designate evaluation officers and statistical officials, as well as to create respective policies and practices for evaluation as outlined by the Evidence Act. The administration has notably included other references to the importance of the evaluation function and key statistical activities in government-wide guidance and planning documents, but the data strategy must be clearly connected to these critical users of government data.⁶

ESTABLISH THE ADVISORY COMMITTEE ON DATA FOR EVIDENCE-BUILDING

When the U.S. Commission on Evidence-Based Policymaking issued its unanimous recommendations to Congress and the President in 2017, it suggested the establishment of the National Secure Data Service as a modern shared service for data linkage activities.⁷ The Evidence Act includes the creation of an advisory committee to further explore and make recommendations on how to go about achieving the Evidence Commission's vision. The President's FY 2020 Budget also includes references to the establishment of a data service. Initial steps can be undertaken, including the creation of the advisory committee, to ensure effective implementation of such a new entity. The Data Coalition recommends the action plan includes a specific action item that makes progress on developing the National Secure Data Service as a targeted strategy for improving access to data while strengthening existing privacy protections.

ISSUE A NOTICE OF PROPOSED RULEMAKING ON THE CODIFICATION OF STATISTICAL POLICY DIRECTIVE NO. 1

Title III of the Evidence Act includes a recommendation from the Commission on Evidence-Based Policymaking to codify OMB Statistical Policy Directive No. 1. The Data Coalition recommends that OMB immediately initiate the regulatory process to ensure SPD1 is reflected in administrative law as soon as possible. As new approaches for data management, curation,

⁶ See OMB Circular A-11 (2019) and White House Office of Management and Budget. *Delivering Government Solutions in the 21st Century: Reform Plan and Reorganization Recommendations*. Washington, D.C.: The White House, 2018.

⁷ U.S. Commission on Evidence-Based Policymaking (CEP). *The Promise of Evidence-Based Policymaking*. Washington, D.C.: Government Printing Office, 2017.

access, and use emerge, government agencies must take steps to ensure the American public can trust government to safely and securely steward this information. The codification of SPD1 represents an important set of guidance for federal agencies about protecting data in a manner that enhances public trust.

3 - STRENGTHENING PROPOSED DRAFT ACTIONS

ACTION 1 - CREATE AN OMB DATA COUNCIL

The Data Coalition applauds OMB's recognition of the need to better coordinate data policy across government. Importantly, Action 1 builds on a suggestion from the Commission on Evidence-Based Policymaking that OMB may need to reorganize to better manage data policies moving forward.⁸ In implementing the OMB Data Council, the Data Coalition recommends leveraging the new CDO Council, the White House Data Fellows, and coordinate as appropriate with federal agencies tasked with government-wide data functions.

ACTION 2 - DEVELOP A CURATED DATA SCIENCE TRAINING AND CREDENTIALING CATALOG

Ensuring a strong workforce for data science and analysis is an important aspect of a successful long-term Federal Data Strategy. The effort to develop a catalog of training and credentialing programs should focus on the task and operational function of employees, link to the CDO Council infrastructure, and be updated routinely (e.g., every six months) to maintain relevance.

ACTION 3 - DEVELOP A DATA ETHICS FRAMEWORK

The Data Coalition applauds the inclusion of the ethics framework as part of the plan. Ensuring that data are ethically and responsibly used is essential to the success of evidence-based policymaking and informed decision-making in our society. To that end, the Data Coalition encourages that the ethics framework reflects decision points and directives that are mapped to specific examples to improve the salience for practitioners and the American public. This could be done, for example, along each of the lifecycle stages of data management articulated in the Federal Data Strategy. The ethics framework should involve a transparent and participatory development process, and include periodic updates based on public input.

ACTION 4 - DEVELOP A DATA PROTECTION TOOLKIT

⁸ See Rec. 5-3 in CEP, 2017.

The Data Coalition appreciates the development and application of a data protection toolkit, including efforts to update Statistical Working Paper 22 with corresponding guidance to federal agencies. The Evidence Act requires federal agencies to better monitor and assess the risk of re-identification of confidential records; the data protection toolkit will be an essential resource as agencies are implementing these new requirements. The Data Coalition suggests that the toolkit be closely developed in concert with the ethics framework (Action 3).

ACTION 5 - DEVELOP A REPOSITORY OF FEDERAL DATA STRATEGY RESOURCES AND TOOLS

The Coalition is pleased to see the emphasis placed on the establishment of a repository as required by the OPEN Government Data Act's (P.L. 115-435, Title II) Section 3511(c)(2) requirement for a "Repository". As the Act requires, the repository should also include, or point to, established schema standards to facilitate greater harmonization across federal data assets (as the law states, the Repository should include "an online repository of tools, best practices, and schema standards to facilitate the adoption of open data practices across the Federal Government"). Since data standards are essentially agreed upon processes and rules for describing and recording the format and meaning of data so that it can be shared, exchanged, and ultimately understood, it would be appropriate to consider standards for inclusion in the Repository as a means to support the goals of open data, data sharing, and harmonization.⁹ Additionally, the Repository's inclusion of data standards and relevant schemas will support such specific FDS Practices as "5. Prepare to Share", "20. Leverage Data Standards", "25. Coordinate Federal Data Assets", and "29. Design Data for Use and Re-Use."¹⁰

ACTION 6 - PILOT A ONE-STOP STANDARD RESEARCH APPLICATION

The proposed researcher application portal reflects an important privacy enhancement and burden reduction provision in the Evidence Act. The pilot should take efforts to ensure the benefits of proposed projects are clearly delineated and available for review by the American public. In addition, the pilot could consider approaches for qualified researchers and users to more efficiently apply to add qualified research assistants or appropriate support personnel for projects.

ACTION 7- PILOT AN AUTOMATED INVENTORY TOOL FOR DATA.GOV AND ACTION 8 - PILOT STANDARD DATA CATALOGS FOR DATA.GOV

⁹ <https://www.usgs.gov/products/data-and-tools/data-management/data-standards>

¹⁰ See OMB Memorandum M-19-18.

Ensuring data inventories with appropriate metadata exist is a critical aspect of protecting privacy and confidentiality of microdata in the long-term. The Data Coalition appreciates the inclusion of a pilot activity for rapidly advancing the development of the data inventory requirement in the Evidence Act and OPEN Government Data Act.

ACTION 9 - IMPROVE DATA RESOURCES FOR AI RESEARCH AND DEVELOPMENT

The Data Coalition strongly supports Action 9 to improve data resources for the research and development of artificial intelligence.

ACTION 10 - IMPROVE FINANCIAL MANAGEMENT DATA STANDARDS

This Action should also address the continued implementation of the Digital Accountability and Transparency Act (DATA Act; P.L. 113-101). OMB should work with the U.S. Treasury Bureau of the Fiscal Service to create a consolidated set of recommendations addressing remaining data quality issues. Specifically, a plan should be developed to examine award systems maintained externally to agency financial systems which USAspending.gov relies on for spending information. Recommendations should be developed to assure these systems more closely integrate with legally established DATA Act management guidance, reporting processes, and data standards.

The FDS should also outline how the Administration can further leverage the U.S. Treasury maintained DATA Act Information Model Schema (DAIMS) to facilitate more automated integration with other federal performance and budget reporting. For instance, OMB's June 2018 memorandum on "Appendix A to OMB Circular No. A-123, Management of Reporting and Data Integrity Risk"¹¹ outlines a number of Financial and Non-Financial federal reporting areas which could integrate relevant data elements defined in the DAIMS data model to facilitate automated integration and informative connections across different areas of federal reporting.

Future federal reporting requirements should leverage modern standard-setting best practices to incorporate schema, taxonomy, and ontology into the data standard models. Federally maintained data standards should seek to specify the relationships and hierarchies between data elements (e.g., program activity, budget function, CFDA, Treasury Account, etc.), relevant technical specifications (e.g., text structure parameters, appropriate metadata such as time parameters), and data quality validation rules.

ACTION 11 - IMPROVE GEOSPATIAL DATA STANDARDS

¹¹ See OMB Memorandum M-18-16.

The U.S. Census Bureau has routinely demonstrated the value of applying geospatial data to improve the quality and efficiency of data collection activities. The Data Coalition supports inclusion of a targeted action for the geospatial data community to develop critical standards and support implementation of the Geospatial Data Act.

ACTION 12 - CONSTITUTE A DIVERSE DATA GOVERNANCE BODY

The Coalition is pleased to see the emphasis placed on the agency Data Governance Body. It would be appropriate for the OMB Data Council to provide recommended models for the agency CDOs to establish these governance bodies based on the agencies that have already established such committees. Agency data governance bodies should have an expectation to interface with the CDO Council to share similar mission related data challenges and best practices across the government (e.g., grant making and federal assistance, regulatory enforcement, information collections and statistical analysis, R&D, fund administration, etc.).

ACTION 13 - ASSESS DATA AND RELATED INFRASTRUCTURE MATURITY

This action should propose a model for how the CDO can accomplish this work while leveraging existing resources maintained by the agency CIO. A data maturity assessment model, such as the internationally recognized Data Management Capability Assessment Model (DCAM)¹², should be recommended to agencies to assure a consistent approach and manner for expressing the findings. The CDO Council should be encouraged to establish a Working Group to consider thematic results which might require coordinated, government-wide projects or resources. Additionally, the GSA team responsible for the Repository should be integrally involved in this process.

ACTION 14 - IDENTIFY OPPORTUNITIES TO INCREASE STAFF DATA SKILLS

The assessment of staff data skills is an important attribute of ensuring government's workforce can meet emerging data needs. To support agencies in conducting such assessments, consideration could also be given to including additional questions as a supplement to the annual Federal Employee Viewpoint Survey.

ACTION 15 - IDENTIFY DATA NEEDS TO ANSWER KEY AGENCY QUESTIONS

This action must be clearly connected to the learning agendas and evidence-building plans required of agencies by the Evidence Act and articulated in OMB Circular A-11. The action plan

¹² <https://edmcouncil.org/page/aboutdcamreview>

could also articulate that the August 2020 deadline is intended to precede the final submission of interim learning agendas, due to OMB in September 2020.¹³

The Evidence Commission's recommendation for establishment of a learning agenda indicated that a key benefit is the identification of knowledge gaps, then determinations of how to align existing data resources (including potential sharing of data) and new data collection activities to address key questions. Agency leaders must ensure the development of learning agendas¹⁴ and identification of data needs is not limited to currently collected data or data 'owned' by a single agency or organizational unit.

The development of learning agendas should also involve stakeholder feedback beyond the government agency. The Data Coalition and its members look forward to supporting meaningful progress on this action.

ACTION 16 - IDENTIFY PRIORITY DATASETS FOR AGENCY OPEN DATA PLANS

Efforts to implement this action should consider the priority data identified by the Evidence Commission as a starting point for engagement with stakeholders.¹⁵ The Data Coalition and its members recognize that it would be infeasible for all eligible information to be made accessible as open data over the next year, and we strongly encourage agencies to prioritize implementation of this provision of the OPEN Government Data Act, in consultation with key stakeholders and data users.

In addition, the Data Coalition strongly recommends that the administration prioritize enabling improved access to income and earnings data, with supporting privacy and confidentiality protections.¹⁶ Devising strategies to improve open data access for some components of this information and ensuring these data are captured in enhanced data inventories will be a critical component to successfully using this information to determine whether a vast number of federal programs are achieving their intended outcomes.

5 - COMMENTS ON IMPLEMENTATION RESOURCES

Without question the actions in the Draft Action Plan will require resource reallocations across government to be fully successful. The Data Coalition recommends OMB issue an additional FY 2020 budget request to Congress to support the establishment, organization, and

¹³ See OMB Circular A-11 (2019).

¹⁴ CEP, 2017.

¹⁵ CEP, 2017; Appendix D.

¹⁶ See also discussion of Rec. 2-6 in CEP, 2017.

empowerment of new leadership positions, including the Chief Data Officers, Chief Evaluation Officers, and Statistical Officials, in addition to developing and apportioning any necessary reprogramming or reallocations in the current fiscal year. Such a request should be based on OMB’s assessment of the resource needs in the short-term.

As part of the Action Plan and upcoming budget guidance for the FY 2021 budget, OMB should also encourage agencies to assess resources needs for implementing the Evidence Act. The Data Coalition strongly recommends agencies share resource requests with OMB for the FY2021 Budget when submitting annual requests to OMB in September 2020; we further strongly recommend OMB develop a pool of resources or a new targeted interagency budget account that can be used for supporting targeted needs in the FY2021 budget process.

OMB should also direct the new OMB Data Council to explore potential updates to OMB Circular A-25 to ensure fee mechanisms align with the current guidance, consider development of Evidence Incentive Funds as recommended by the U.S. Commission on Evidence-Based Policymaking,¹⁷ and consider how government can apply flexibilities in procurement processes to achieve additional efficiencies.¹⁸

6 - OTHER GENERAL COMMENTS

STATE AND LOCAL GOVERNMENT PARTICIPATION

The action plan appears to skirt the issue of the role state and local governments play in supporting federally-funded but locally-implemented programs. Additional attention should be given to ensuring a collaborative and participatory process for co-implementers of key government data activities.

IDENTIFY “RESPONSIBLE INDIVIDUALS”

The draft action plan appropriately identifies agencies the actions will be assigned to. However, as we have learned through decades of implementing management activities in government, such as the Government Performance Results Act Modernization Act of 2010, a more effective strategy is to identify individual senior leaders who can be held responsible both inside government and publicly for achieving the intended goal.¹⁹

¹⁷ See discussion in Rec. 5-5 of CEP, 2017.

¹⁸ See discussion in Rec. 5-4 of CEP, 2017.

¹⁹ Nick Hart and Katherine Newcomer. Presidential Evidence Initiatives: Lessons Learned from the Bush and Obama Administrations. Washington, D.C.: Bipartisan Policy Center, 2018.

PERFORMANCE METRICS

We applaud the intent to include metrics through which the American public can hold the Executive Branch accountable for achieving the individual actions in the draft plan. However, the draft action plan includes numerous performance metrics for individual actions that fail to satisfy basic measurement criteria as articulated in OMB Circular A-11.²⁰ In several instances, the measures are largely misaligned with the stated action or offer mere binary determinations for whether an action was achieved. For example, in Action 1 for the creation of an OMB Data Council, the establishment of the Council itself is less important than whether the group achieves activities of value. With an increased focus in government on achieving meaningful outcomes, the measurements should include well-articulated and well-defined measures.

TIMELINES

To improve accountability for the final action plan, we recommend in lieu of relative timeframes, that the action plan include specific deadlines for completion or progress on individual actions.

LEGAL OBLIGATIONS

As noted throughout these comments, more attention can be given to expressly delineating how the actions align with new legal obligations in the Evidence Act and other statutory authorities. This could minimally be achieved by listing the statutory authority at the end of each action in addition to the “supported practice” statement.

²⁰ OMB Circular A-11, 2019.