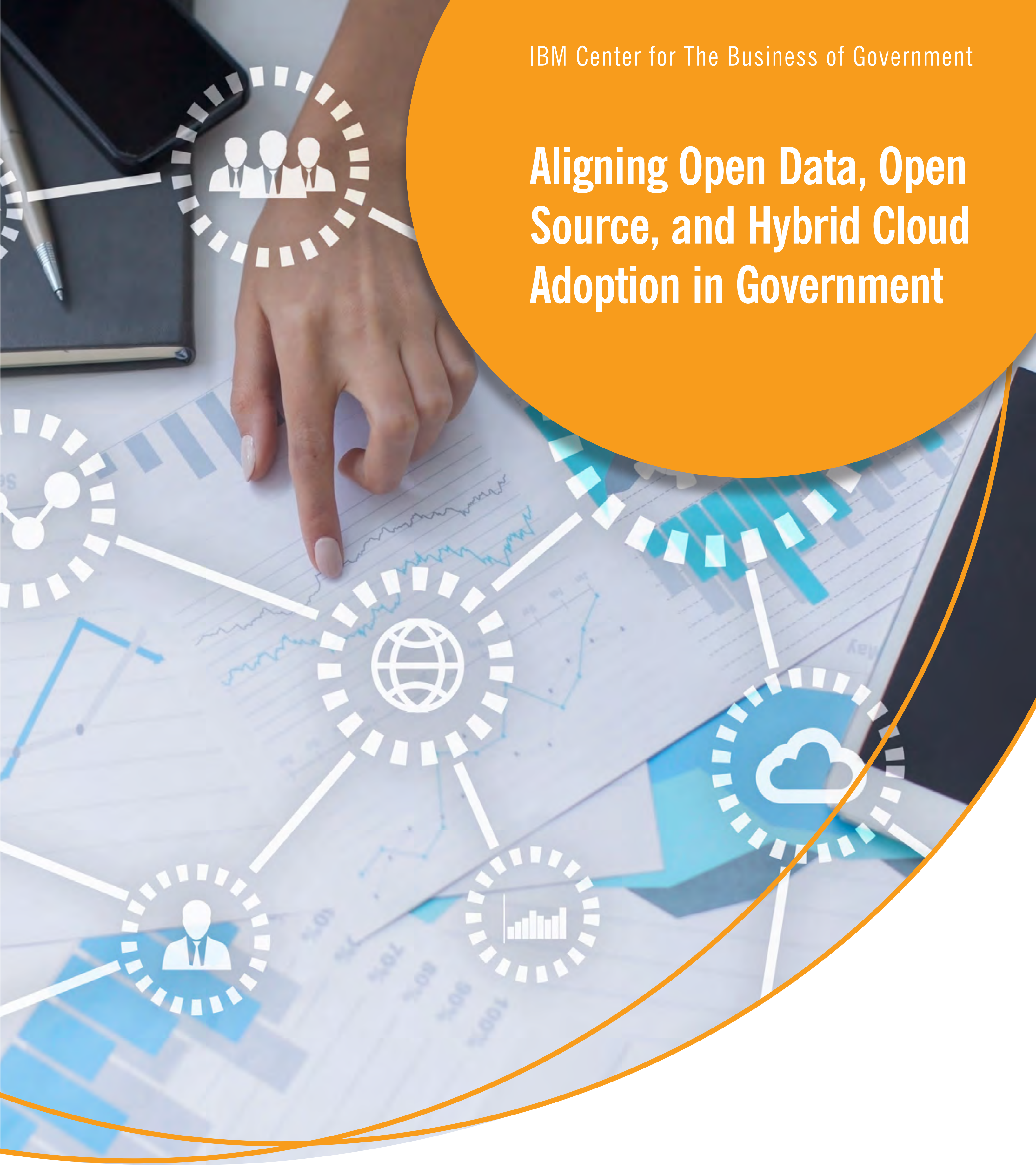


IBM Center for The Business of Government

Aligning Open Data, Open Source, and Hybrid Cloud Adoption in Government

Matt Rumsey and Joel Gurin
Center for Open Data Enterprise

IBM Center for
The Business of Government

2021

Aligning Open Data, Open Source, and Hybrid Cloud Adoption in Government

Matt Rumsey and Joel Gurin

Center for Open Data Enterprise

IBM Center for
The Business
of Government

2021

TABLE OF CONTENTS

FOREWORD

On behalf of the IBM Center for The Business of Government, we are pleased to release a new report, *Aligning Open Data, Open Source, and Hybrid Cloud Adoption in Government,* by Matt Rumsey and Joel Gurin of the Center for Open Data Enterprise (CODE).

Federal agencies are increasingly adopting cloud infrastructure, sharing their data across agencies and with the public, and relying on open source software (OSS) as they seek to enhance their capabilities, improve efficiency through IT modernization, and leverage their data for policymaking and administration. These developments have been acknowledged and formalized via government-wide strategic initiatives including the Federal Cloud Computing Strategy, the Federal Data Strategy and implementation of the Evidence-Based Policymaking Act, and the Federal Source Code Policy.

DANIEL J. CHENOK

To date, these domains have been addressed and implemented separately in agencies, with little focus on the manner in which they complement one another. Reviewing how best to implement them together will help make government information more transparent to the American public, promote efficiency in software development and interoperability across cloud domains, and increase the effectiveness and lower the costs of government operations.

DAVID EGTS

In this report, authors Rumsey and Gurin examine how these trends in analytics and technology intersect and can mutually reinforce one another. The authors draw on insights from an expert roundtable that brought together leaders in government use of data, software, and cloud approaches to discuss how these domains can they best be integrated under current federal policies—exploring how government leaders and stakeholders can leverage the intersection of open data, open source, and hybrid cloud models to drive improved performance and productivity.

The report assesses government progress in each domain, and then importantly reviews how the domains intersect. Based on this analysis, the authors present a series of recommendations for how government can best leverage the synergies across cloud, open source and open data. These recommendations include high-level considerations like improving workforce skills and sharing success stories, as well as specific proposals to update policy approaches in ways that promote an integrative data and technology strategy.

This report builds on prior Center work around agency IT and data strategy, including: *Making Agencies Evidence-Based: The Key Role of Learning Agendas* by Nick Hart, Kathryn Newcomer, and Karol Olejniczak; *Innovation and Emerging Technologies in Government: Keys to Success* by Alan Shark; *Data-Driven Government: The Role of Chief Data Officers* by Jane Wiseman; and *A Roadmap for IT Modernization in Government* by Gregory Dawson.

We hope that the analysis and recommendations from this report, drawing on the insights of senior leaders in the expert roundtable discussion, will help government agencies and stakeholders in understanding and updating strategies that take advantage of the evolving capabilities presented by open data, open source software, and hybrid cloud implementation.

Daniel J. Chenok
Executive Director
IBM Center for The Business of Government
chenokd@us.ibm.com

David Egts
Chief Technologist
North America Public Sector, Red Hat
degts@redhat.com

EXECUTIVE SUMMARY

Federal agencies have increasingly looked to enhance their technical capabilities, improve efficiency through IT modernization, and leverage data for policymaking and administration.

Governmentwide policies around open data, open source software, and cloud computing have been developed to help achieve these often-interconnected goals. Understanding and acting on the synergies across these three domains will help make government information more available to the American public, ensure that government decisions are driven by evidence, and increase the efficiency and lower the costs of government operations.

This paper presents findings from research and an expert roundtable discussion that explored how existing policies and strategies governing those three areas are working, and how they can be better integrated to reinforce one another and advance the federal government's IT modernization and data use goals. The report analyzes relationships between implementation of four highly relevant policies:

- Federal Cloud Computing Strategy
- Federal Data Strategy
- Rules that guide agency action around the Foundations for Evidence-Based Policymaking Act
- Federal Source Code Policy

These policies set the stage for addressing key issues of privacy, security, workforce, procurement, decision making, interagency collaboration, standards, and public value.

This paper describes areas where the agency practices under these policies overlap and support each other, and areas needing greater alignment. The report also describes how these "open" approaches intersect and can complement each other, addressing interfaces between cloud and open data, cloud and open source, and open source and open data—all of which can be developed to increase the effectiveness of each policy area.

Federal agencies can improve and expand their use of the cloud, open data, and open source in a number of ways. This paper provides high-level recommendations for strategic operating principles, and recommendations for action to consider in the near term. Additionally, the paper describes three themes that touch on many of the recommendations: the need for resources, leadership, and policy.

High-Level Recommendations

MAINTAIN FLEXIBILITY

Agencies should avoid one-size-fits-all approaches, given their different needs, operating scales, and timelines.

SHARE SUCCESS STORIES

Agencies have progressed to different degrees in adopting cloud technologies, embracing OSS, and opening data. Highlighting success stories can help inspire other agencies as they work towards change.

BUILD THE WORKFORCE

The current federal workforce is not well skilled to meet the demands of a government driven by cloud, OSS, and open data. New training, updated career paths, and dedicated recruitment strategies can help change this.

APPLY USER-DRIVEN APPROACHES

Agencies have recently begun to take a more customer- or user-centered approach to their operations and technology. This user orientation should continue and be made even more central to agency technology programs.

LEARN FROM THE PANDEMIC

Government agencies' adaptation to the COVID-19 pandemic—particularly the acceleration of remote work and increased sharing of scientific information—provide models for implementing more open methods.

Recommendations for Action

1. Fund Evidence Act implementation
2. Leverage the TALENT Act to build a data- and cloud-literate workforce
3. Provide vouchers for researchers, nonprofits, and others to use cloud resources for research, analysis, and policymaking
4. Update the Source Code Policy to encourage government use of existing, robust OSS products and further engagement with OSS projects and communities
5. Update the key federal data policies to better align them and fill gaps in those policies
6. Develop a plain language toolkit to explain how these policies intersect, what they mean for agencies, and how to implement them in ways that build strength across the three areas
7. Leverage agency data inventories to understand agency data systems

Three Key Themes

RESOURCES

Funding and other resources are critical to modernize systems and fully leverage data for decision making.

LEADERSHIP

The culture that can resist change at federal agencies won't improve without strong, consistent leadership that helps support change over time.

POLICY ALIGNMENT

Implementation of federal policies governing open data, open source, and cloud adoption has largely proceeded on separate tracks. More should be done to align all three.

INTRODUCTION

Federal agencies are increasingly adopting cloud infrastructure, sharing their data across agencies and with the public, and relying on open source software (OSS) as they seek to enhance their capabilities, improve efficiency through IT modernization, and leverage data for policymaking and administration.

These developments have been acknowledged and formalized via governmentwide policies including the Federal Cloud Computing Strategy[1] (the Cloud Strategy), the Federal Data Strategy[2] (the FDS), the Foundations for Evidence-Based Policymaking Act[3] (the Evidence Act), and the Federal Source Code Policy[4] (Source Code Policy).

These trends in analytics and technology intersect in a number of ways and can mutually reinforce one another. But how can they best be integrated, and how well are federal policies working? To answer those questions, this report from the IBM Center for The Business of Government in collaboration with the Center for Open Data Enterprise (CODE) explores how government leaders and stakeholders can leverage the intersection of open data, open source, and hybrid cloud models to drive improved performance and productivity.

This paper explores these topics and presents recommendations for action, drawing on an analysis of existing laws and policies, interviews with experts in open technology, and a virtual *Roundtable on Open Data, Open Source, and Cloud Adoption* held on October 22, 2020, by the IBM Center and CODE. It provides a high-level overview of the laws and policies governing open data, OSS, and cloud technologies across the federal government, analyzes how these various policies intersect across key areas including security, privacy, and governance, and presents insights and recommendations for ways that policymakers and practitioners can bring the three areas into greater alignment.

The Biden administration faces unique challenges related to technology modernization and data sharing brought on by the COVID-19 pandemic, as well as a changing federal workforce. Agencies will likely continue with digital transformation efforts—including open data and cloud adoption—started under President Obama and advanced through the Trump administration. The Biden administration is continuing these efforts and placing a renewed focus on the importance of science and technology at the highest levels.[5] President Biden has already appointed aides[6] who served in the Obama White House—including one with ties to the United States Digital Service—to key technology roles in his White House, and elevated the Office of Science and Technology Policy to a position[7] in his Cabinet. This paper and these recommendations are intended to help agencies and new leaders continue modernizing the federal government and achieve their technology policy and management goals.

[1] "Home | Federal Cloud Computing Strategy." n.d. Cloud.cio.gov. https://cloud.cio.gov/.

[2] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[3] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

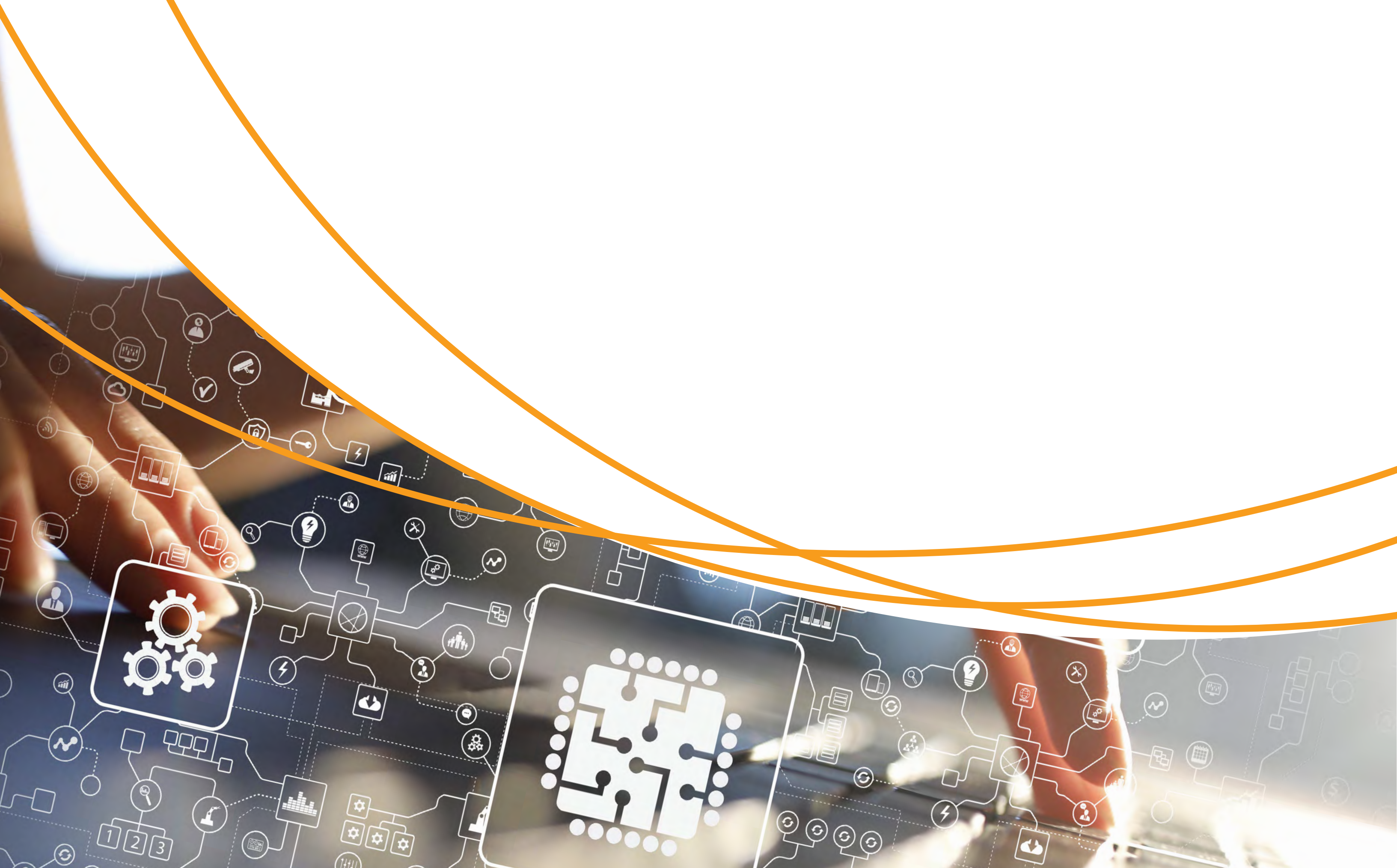
[4] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

[5] "Top Management Priorities for the Biden Administration." 2020. *Government Matters*. December 18, 2020. https://govmatters.tv/top-management-priorities-for-the-biden-administration/.

[6] Barnett, Jackson. 2021. "Biden Transition Team Names White House Tech Officials." *FedScoop*. January 5, 2021. https://www.fedscoop.com/biden-transition-white-house-technology-officials/.

[7] Kozlov, Max. 2021. "Biden Names Geneticist Eric Lander as Top Science Adviser." *The Scientist Magazine*®. Accessed February 18, 2021. https://www.the-scientist.com/news-opinion/biden-names-geneticist-eric-lander-as-top-science-adviser-68363.

Overview



Integrating strategies around open government data, OSS, and cloud computing will help make government information more available to the American public, ensure that government decisions are driven by evidence, and increase the efficiency and lower the costs of government operations. Several federal policies have been put in place to support these approaches and, in some cases, build on the connections between them. However, more can be done to integrate these approaches that enable government data to be shared and used for evidence-based decision making, while software and cloud capabilities can be leveraged across agencies in more cost-effective ways. The challenges include security and privacy, a federal workforce needing greater skills in 21st century technologies, under-resourced agencies, and questions about the return on investment for some technology improvements.

Overall, the need to open data for innovative public use, and more fully leverage data to guide government decision making, should serve as a driver across all three areas. Open and shared data is now mandated by law. This policy imperative can help create a virtuous cycle by driving adoption of cloud technologies and open source software, which can then support more robust, efficient, and flexible sharing of vital government data.

Open Data

The federal government has made its data more open and accessible to the American public for a number of years. Recently, the Evidence Act and the Federal Data Strategy (FDS) have used the force of law and federal policy to achieve this goal.

CODE's research identified the need to protect privacy and security as the major challenge with respect to the implementation of open data and data management programs across the federal enterprise. Both the Evidence Act and the FDS focus significant attention on privacy and security, with the Evidence Act specifically strengthening the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) to improve the privacy of statistical information.

The Foundations for Evidence-Based Policymaking Act (the Evidence Act)

The Evidence Act provisions include requiring agencies to use evidence for evaluation and policymaking, making open data the law of the land, giving legal structure to data management activities, and strengthening and reauthorizing existing data privacy protections.[8] The Act was introduced in 2017 and passed by the House that year. The Senate passed the bill at the end of 2018, and it became law in January 2019.[9] The Evidence Act encourages the use of evidence and data in policymaking, while opening data and protecting privacy and security. The law builds on existing efforts and, in many ways, lays a legislative foundation for other forward-looking work like the FDS.

Title II of the Evidence Act, the OPEN Government Data Act, puts a broad presumption of openness for government information into law and lays out a path to ensure that data is modern, machine readable, and of high quality. Title II includes governance initiatives like a requirement for data inventories that match strategies in the Source Code Policy (described below), and provides legislative guidance to the FDS in a number of ways.

[8] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[9] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

The Evidence Act aligns well with the goals of other efforts to modernize government information technology and data sharing systems. It makes it easier for qualified researchers to remotely access government data, lays the groundwork for modernizing federal data infrastructure, and strengthens a key privacy law (CIPSEA) while balancing it with appropriate data use.[10]

The Federal Data Strategy (FDS)

The Federal Data Strategy grew out of the 2018 President's Management Agenda, which introduced a new Cross-Agency Priority (CAP) goal to leverage data as a strategic asset.[11] That CAP goal expressed the need for a “robust, integrated approach to using data to deliver on mission, serve customers, and steward resources while respecting privacy and confidentiality,” and identified an “enterprisewide” Federal Data Strategy as a path to achieving that goal.[12] The FDS was developed by representatives from 23 federal agencies, who sought feedback from other federal employees and the general public.[13]

Overall, the FDS emphasizes aims to leverage data for the public good. The strategy takes a number of cues from the Evidence Act, while providing specific guidance for agencies. The FDS is grounded in principles and practices which represent aspirational and actionable goals, applied through a yearly Action Plan as well as specific agency plans.[14] The Action Plan “identif[ies] and prioritize[s] practice-related steps for a given year, along with targeted timeframes and responsible entities.” The FDS requires agencies to create open data plans developed with stakeholder engagement, inventory their data assets, and make qualitative and quantitative improvements to their data sets.

The FDS Principles, which serve as aspirational guidelines, are grounded in concepts of ethical governance, conscious design, and learning cultures.[15] Agencies are expected to practice effective data governance through sound data security practices and ensuring individual privacy and confidentiality, while allowing appropriate public use of federal data and developing skills and leadership within the federal enterprise. The FDS Practices lay out the broad steps that agencies can take to achieve these goals.[16]

The FDS Practices address regulatory, legal, and cultural barriers to data sharing and use across federal agencies. Agencies are urged to promote sharing for the purpose of better public use of data. The Evidence Act also creates a presumption of data asset accessibility for statistical agencies.

Additionally, several Practices provide guidance on the intersection of data access and privacy protection, and explicitly prioritize data governance to achieve these goals. For example, the FDS recommends that agencies consider tiered data access to minimize privacy risk.

This focus on privacy and security does not come at the expense of data sharing and use. While FDS Practice 25 encourages coordination and sharing of data assets between federal agencies,[17] Practice 26 encourages sharing between federal agencies and state, local, and tribal governments. More broadly, the Practices lay the groundwork for data sharing, pushing

[10] Foundations for Evidence-Based Policymaking Act of 2018 Fact Sheet. Data Coalition. http://www.datacoalition.org/wp-content/uploads/2019/06/Evidence-Act-Web-version-2019.pdf.

[11] “President's Management Agenda.” n.d. Performance.gov. https://trumpadministration.archives.performance.gov/PMA/PMA.html.

[12] “Leveraging Data as a Strategic Asset.” n.d. Performance.gov. https://trumpadministration.archives.performance.gov/CAP/leveragingdata/.

[13] “Background - Federal Data Strategy.” n.d. Strategy.data.gov. Accessed February 18, 2021. https://strategy.data.gov/background/.

[14] “Action Plan - Federal Data Strategy.” n.d. Strategy.data.gov. Accessed February 18, 2021. https://strategy.data.gov/action-plan/.

[15] “Principles - Federal Data Strategy.” n.d. Strategy.data.gov. Accessed February 18, 2021. https://strategy.data.gov/principles/.

[16] “Practices - Federal Data Strategy.” n.d. Strategy.data.gov. Accessed February 18, 2021. https://strategy.data.gov/practices/.

[17] “Practices - Federal Data Strategy.” n.d. Strategy.data.gov. Accessed February 18, 2021. https://strategy.data.gov/practices/.

agencies to "proactively address the procedural, regulatory, legal, and cultural barriers to sharing data within and across federal agencies, as well as with external partners."

The FDS Practices also integrate data into broader capital and technical planning efforts. For example, Practice 18 emphasizes including data investments in annual capital planning processes to ensure that use of funds leverages data as a strategic asset. Meanwhile, agencies have taken advantage of the Evidence Act's statutory requirement for agency chief data officers (CDOs) to integrate data governance and use with IT planning, which may facilitate leveraging IT for data and evidence activities. This integration is often difficult to budget for cloud analytics and data sharing projects;[18] CDOs have a strong understanding of the potential audience for their open data or the amount of computing power and time data users may need, making planning and budgeting easier.

2020 Action Plan

The 2020 Action Plan is designed to build a foundation to help agencies implement the FDS over the next decade. The Plan also builds cross-agency capacity by developing and empowering communities of practice and shared services.

The 2020 Action Plan requires agencies to develop stakeholder-driven open data plans to ensure availability of open government data, while making quantitative and qualitative improvements to existing data sets. The Plan lays the groundwork for robust agency data governance by requiring all agencies to create[19] a "Data Governance Body (DGB)" that includes representation from across business units, and also prioritizes government-wide data governance by calling for the launch of a Federal Chief Data Officer Council.[20] The 2020 Action Plan also requires agencies to develop or evolve their data inventories—a requirement of the Evidence Act—with updated metadata. This will enable Data.gov, the government's central portal for open data, to catalog data from agencies, and will facilitate search engine optimization (SEO).

Next Steps

While the Biden administration has not yet committed to a deadline for a 2021 FDS Action Plan, the administration's early actions show a commitment to use data for evidence building across a wide range of policy areas. In his first month in office, President Biden signed executive orders that leverage the use of data to tackle challenges ranging from the environment and climate change to the coronavirus, racial equity, and economic recovery.[21] Biden's Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking also cemented the new administration's focus on rebuilding public trust using data and evidence.[22]

[18] Sharwood, Simon. 2020. "NASA to Launch 247 Petabytes of Data into AWS – but Forgot about Eye-Watering Cloudy Egress Costs before Lift-Off." The Register. March 19, 2020. https://www.theregister.com/2020/03/19/nasa_cloud_data_migration_mess/.

[19] "Action Plan - Federal Data Strategy." n.d. Strategy.data.gov. Accessed February 18, 2021. https://strategy.data.gov/action-plan/#action-2-constitute-a-diverse-data-governance-body.

[20] "Action Plan - Federal Data Strategy." n.d. Strategy.data.gov. Accessed February 18, 2021. https://strategy.data.gov/action-plan/#action-7-launch-a-federal-chief-data-officer-council.

[21] "2021 Joe Biden Executive Orders." 2021. Federal Register. https://www.federalregister.gov/presidential-documents/executive-orders/joe-biden/2021.

[22] Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking (2021). https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking/.

Challenges

In addition to ongoing concerns around privacy and security which are addressed in the Evidence Act and the FDS, one major challenge emerges regarding implementation of these policies: the lack of resources. The Evidence Act was passed without an associated appropriation. Agencies must fund this work from existing budgets, which significantly limits their ability to invest for real change. Chief data officers have small or nonexistent staffs and have to compete for limited resources, instead of collaborating to make change within their agencies. Funding could help with staffing and fulfilling requirements in the law, such as creating data inventories and evaluation plans. Ideally, funding for Evidence Act and FDS implementation could also be leveraged to make progress on related cloud and open source priorities.

Open Source

The federal government launched[23] its Source Code Policy[24] in 2016, officially encouraging agencies to share custom-developed source code across government. The Source Code Policy also marked the start of a pilot effort that required agencies to release at least 20 percent of new custom-developed code as OSS for three years.

Government agencies had previously engaged in OSS projects and released source code on an ad hoc basis, including some projects at the highest levels, with WhiteHouse.gov releasing open source code as early as 2010.[25] But the establishment of the Source Code Policy was the first time all agencies were directed to embrace openness with respect to custom-developed code. When agencies choose to procure custom-developed code, the policy requires them to obtain the rights to share that code with other government agencies and—if deemed appropriate—to release the code as OSS.

The Source Code Policy has several objectives. Specifically, it gives guidance to agencies as they consider whether or not to acquire custom-developed code, ensures that agencies retain appropriate control over custom code that they choose to procure, pushes agencies to consider OSS, and makes it easier for agencies to release code to the public. The Policy provides agencies with a three-step decision making approach for software procurement or development; it also directs agencies to consider hybrid solutions, modular architecture, cloud computing, open standards, and other targeted considerations.

The Source Code Policy shares governance approaches with governmentwide open data policies, including the requirement to inventory agency custom-developed code and the creation and maintenance of a central repository for information access—Code.gov[26] in the case of the Source Code Policy, Data.gov[27] in the case of open data efforts. Early and ongoing government open data efforts have also relied on OSS for their infrastructure needs. For example, Data.gov was built using the Comprehensive Knowledge Archive Network (CKAN), a prominent OSS project that provides a platform for open data and data sharing.[28]

[23] Scott, Tony. 2016. “The People’s Code. Code.gov (blog).” August 11, 2016. https://www.cio.gov/2016/08/11/peoples-code.html.

[24] Tony Scott and Anne E. Rung, “Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software” (official memorandum, Washington, D.C.: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

[25] “WhiteHouse.gov Releases Open Source Code.” 2010. April 21, 2010. https://obamawhitehouse.archives.gov/blog/2010/04/21/whitehousegov-releases-open-source-code.

[26] Code.gov website: https://code.gov/.

[27] Data.gov website: https://www.data.gov/.

[28] CKAN website: https://ckan.org/.

Challenges

While OSS makes sense for Data.gov as a governmentwide project, individual agencies often need more incentive to share their own code as open source or reuse code developed by other agencies.[29] CODE's interviews indicated difficulty in demonstrating ROI on open source. At the same time, however, closed-source software and existing procurement practices often give government agencies less flexibility than OSS solutions may offer.

To be successful, open source products require independent viable communities which contribute to the code and its governance.[30] In the context of government cloud adoption, open source technology must be based on open standards or formats moving towards standardization, must not restrict the use of intellectual property, and must feature application programming interfaces (APIs) that are usable across industry and not controlled by a specific vendor. CODE's interviews indicate that government OSS projects often struggle to develop engaged communities. However, agencies often hesitate to engage with existing OSS projects and robust communities that could serve their software needs, for fear of becoming dependent on communities that may or may not prioritize or meet government needs over time.

Openness can add significant flexibility, allowing software to be fixed on the fly or data to be shared and analyzed without unnecessary bureaucracy. This flexibility can be vital in high pressure or mission critical situations, reducing the burden of negotiating with a software vendor or tracking down the original owner of a dataset. One roundtable participant described the benefits of flexibility from the perspective of the Department of Defense, where "if you can't hack it, don't pack it."

The lack of flexibility and control associated with purpose built, closed source technology can also cost organizations extra money and time in the long run. As an example, in 2014 the contractor that built several important government systems, including FederalSpending.gov and the Federal Procurement Data System-Next Generation, went bankrupt.[31] The contracts gave control over these systems and their data to the contractor, not the government, forcing agencies to spend upwards of $30 million to buy the systems and data and keep them running while they found new contractors to manage them.

Cloud

Federal agencies have shifted their data storage and access needs from agency-owned, in-house data centers to cloud-based services since at least 2009. In a 2019 report, the Government Accountability Office (GAO) found that the federal agencies studied had all made advances in implementing cloud services, with some making more progress than others.[32]

The cloud, as defined by the National Institute of Standards and Technology (NIST), is "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[33] Models for cloud deployment include private, public, community, and hybrid.

[29] Data.gov website: https://www.data.gov/.

[30] Cloud Computing Team. 2012. "The Open Cloud: Red Hat's Perspective." Red Hat (blog). February 15, 2012. https://www.redhat.com/en/blog/The-Open-Cloud-Red-Hats-Perspective.

[31] Rumsey, Matt. 2014. "Government Data Should Be Public—Not Owned by Contractors." November 3, 2014. https://sunlightfoundation.com/2014/11/03/government-data-should-be-public-not-owned-by-contractors/.

[32] "Cloud Computing: Agencies Have Increased Usage and Realized Benefits, but Cost and Savings Data Need to Be Better Tracked." 2019. The U.S. Government Accountability Office (GAO). https://www.gao.gov/assets/700/698236.pdf.

[33] Mell, Peter, and Tim Grance. 2011. " The NIST Definition of Cloud Computing." COMPUTER SECURITY RESOURCE CENTER (CSRC). https://csrc.nist.gov/publications/detail/sp/800-145/final.

The most relevant models for federal government application are hybrid cloud models, which allow customers to use both public cloud infrastructure and private cloud systems that operate on private networks. This hybrid approach provides a balance between security and scalability and is popular among federal agencies. Hybrid models make it easy to share data while protecting privacy and ensuring robust security.

Federal agencies that want to leverage cloud technologies can look to the Cloud Strategy,[34] published by the Office of Management and Budget (OMB) in 2019 and superseded the Cloud First Strategy.[35] The Cloud Strategy (renamed from Cloud First to Cloud Smart) focuses on security, procurement, and workforce as three key areas that drive agency adoption of cloud solutions, and encourages agencies to leverage the distributed nature of the cloud and hybrid and multi-cloud solutions.

The Cloud Strategy recognizes that the software industry is moving to new capabilities offered at different system layers, enabling components to be managed by external vendors, government agencies, or a combination. The strategy encourages agencies to leverage the distributed nature of cloud and hybrid and multi-cloud solutions, and emphasizes a risk-based and integrated approach to securing cloud environments. As agency data flows through networks and resides within systems on premise or through cloud environments, agencies need the ability to detect malicious activity, integrate privacy programs, and ensure continuous data protection. FedRAMP (Federal Risk and Authorization Management Program) serves as the standard governmentwide approach to security assessment, authorization, and continuous monitoring of cloud services for federal agencies. To date, updates to existing cyber policies, particularly TIC 3.0 and the NIST Risk Management Framework, have cleared the pathway for a more accelerated adoption of commercial cloud capabilities.[36]

The Cloud Strategy pushes agencies to "leverage the strength of the government's bulk purchasing power, the shared knowledge of good acquisition principles . . . [and] relevant risk management practices" to overcome existing challenges associated with purchasing cloud services and technology.[37] The Strategy notes that procurement decisions should prioritize security, privacy, and continuity of service while avoiding vendor lock-in.

Finally, the Cloud Strategy acknowledges that moving to the cloud will have a significant impact on the Federal workforce.[38] To address this issue, agencies need to identify skills gaps, train existing staff to fill those gaps, and—where necessary—bring in new employees. The Cloud Strategy encourages agencies to pursue a range of approaches to workforce transformation.

Agencies now use the cloud for internal operations and to enable data and information sharing, including with qualified researchers, other nongovernmental partners, and in some cases the general public. For example the Big Data Program (BDP) of the National Oceanic and Atmospheric Administration (NOAA) leverages nongovernment cloud providers to more effectively share NOAA's massive open data holdings with the general public.[39]

[34] "Strategy | Federal Cloud Computing Strategy." 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[35] Kundra, Vivek. 2011. "Federal Cloud Computing Strategy." U.S. Chief Information Officer. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/federal-cloud-computing-strategy.pdf.

[36] See https://www.cisa.gov/trusted-internet-connections for more on TIC 3.0 and https://www.nist.gov/cyberframework/risk-management-framework For more on the NIST Risk Management Framework.

[37] "Strategy | Federal Cloud Computing Strategy," n.d. Cloud.cio.gov. Accessed February 18, 2021. https://cloud.cio.gov/strategy/#procurement.

[38] "Strategy | Federal Cloud Computing Strategy." n.d. Cloud.cio.gov. Accessed February 18, 2021. https://cloud.cio.gov/strategy/#workforce.

[39] "Big Data Program | National Oceanic and Atmospheric Administration." n.d. Www.noaa.gov. Accessed February 18, 2021. https://www.noaa.gov/organization/information-technology/big-data-program.

Challenges

While many agencies have begun to move on from their legacy systems in favor of cloud solutions, several challenges have impeded progress. These include inconsistent adoption, poor compliance, lack of interagency collaboration, and a mismatched workforce.

Cloud adoption has been inconsistent across federal agencies, with a recent Government Accountability Office report finding that despite a decade of progress, only a fraction of federal IT systems run in the cloud.[40] Several reasons exist for this slow transition, including complex and expensive security and privacy requirements, varying levels of technical capacity, a range of computing and data needs at agencies, and a lack of interagency collaboration.

FedRAMP often does not allow agencies and offices to easily embrace the cloud. For example, a subagency or office may need additional Authority to Operate agreements from its parent agency. Cloud.gov is a government managed platform intended to help agencies "buy, build, and authorize" cloud services and help with FedRAMP and similar compliance issues.[41] However, our interviews made clear that this value is not necessarily known to stakeholders.

Further, agencies have limited visibility into each other's cloud adoption efforts. Even though collaboration forums exist, they are not widely known or leveraged well. CODE's interviews highlighted the importance of getting CDOs involved in decisions around cloud implementation, given channels for cross-agency collaboration among CDOs through the CDO Council, CDOs' statutory role, and the cloud's utility for data sharing and analytics. The Chief Information Officer's Council (CIO Council) could also serve as a channel for collaboration in this area.

Issues do not just exist across agencies. Different components tend to function independently within agencies as well, resulting in expensive cloud systems and missed opportunities. One interviewee noted that different programs make cloud decisions independently. As a result, most agencies have contracts with all of the major cloud service providers, but lack an understanding of their cloud services at an enterprise level. Some agencies have addressed this issue by providing cloud adoption standards and frameworks that align the activities through the enterprise governance process, while allowing for flexibility in procuring and implementing within operating components.

Finally, the federal workforce is restricted by challenges that include insufficiently trained personnel, a lack of standardized roles and responsibilities across agencies, high attrition due to higher salaries in the private sector, and fluctuating contractor workforce.

Despite these challenges, cloud adoption is expected to continue and possibly accelerate in response to the COVID-19 pandemic.[42] The need to accommodate growing numbers of remote workers and increased demand for offsite data access highlights the cloud's value and helps to overcome institutional cultural barriers that previously stood in the way of innovation.

[40] "Cloud Computing: Agencies Have Increased Usage and Realized Benefits, but Cost and Savings Data Need to Be Better Tracked." 2019. The U.S. Government Accountability Office (GAO). https://www.gao.gov/assets/700/698236.pdf.

[41] Cloud.gov website: https://cloud.gov/.

[42] Davis, John. 2020. "Federal Cloud Adoption: Remote Working Accelerates the Urgency." *Federal News Networks*. May 25, 2020. https://federalnewsnetwork.com/commentary/2020/05/federal-cloud-adoption-remote-working-accelerates-the-urgency/.

Cross-Cutting Concerns

The Cloud Strategy, FDS, Evidence Act, and Source Code Policy approach their goals using different methods. However, they all face similar challenges, including security, privacy, governance, and more. To understand how these initiatives currently align and where they could be further harmonized based on implementation experience, it is important to see how they each approach these challenges. The tables below present high-level summaries comparing requirements across each policy area.

Comparing Common Issues Across Federal Policies

SECURITY	Security cuts across all of the policies, with a particular focus on data security and integrity to manage risk while realizing the value that comes with data linkage.
PRIVACY	All of the policies focus on maintaining individual privacy, which is closely linked to data security.
GOVERNANCE	Each policy approaches governance in unique ways, with some overlap including the use of inventories to track assets, a focus on privacy and security, and the need for public engagement.
WORKFORCE	The Cloud Strategy, FDS, and Evidence Act recognize the need for a data- and technology-literate federal workforce. Common approaches include capacity assessments and skill building throughout the enterprise.
PROCUREMENT	The Cloud Strategy, FDS, and Source Code Policy all pay at least some attention to procurement, with the Cloud Strategy and Source Code Policy encouraging rigorous analysis prior to purchase decisions.
USE IN DECISION MAKING	The FDS and Evidence Act are specifically focused on using data to guide decision making. The Cloud Strategy acknowledges that cloud computing can help agencies with decision making through analytics and easier data access.
INTERAGENCY COLLABORATION	Each policy acknowledges the need for increased interagency collaboration, with common themes around purchasing power, skill sharing, reduction of red tape, and progress towards shared goals.
USE OF STANDARDS	The FDS, Evidence Act, and Source Code Policy all encourage the use of standards to improve data quality, increase data access, and ensure that software can be easily used, adapted, and reused.
PUBLIC VALUE	Each policy acknowledges that agencies answer to the American people, and promotes principles of public value and access.

Security

Security concerns cut across all four policy areas, as they do across all aspects of the federal enterprise. There are similarities in how these policies address security, particularly between the Cloud Strategy and the FDS as well as between the FDS and the Evidence Act.

The Cloud Strategy encourages security at the data layer as well as at the network and physical infrastructure layers, which have been traditionally addressed. This provides additional security as data transitions to cloud environments, and matches nicely with the FDS' focus on protecting data integrity. The FDS also places an emphasis on harnessing safe data linkage to ensure security and privacy protection. The Evidence Act addresses the risks of the mosaic effect, which can cause security or privacy issues thanks to the unintended consequences of combining various data sets.[43] This focus on data linkage should add a strong security foundation.

The Federal Cloud Computing Strategy[44]	Requires adding security and privacy controls to the data layer in addition to the network and physical infrastructure layers
The Federal Data Strategy[45]	Instructs agencies to: • Exercise responsibility by practicing effective data stewardship, governance, and security • Protect data integrity through security best practices • Deploy data linkage and analysis tools that use secure and privacy-protecting technology
The Foundations For Evidence-Based Policymaking Act[46]	Requires agencies to consider security and privacy, including the mosaic effect, in decisions to open data
The Federal Source Code Policy[47]	Requires agencies to consider security and privacy while selecting software solutions

Privacy

Data privacy is closely linked to security and also a key issue for federal policymakers. All of the policies analyzed here focus on maintaining personal privacy. There are specific synergies between the Cloud Strategy, FDS, and Evidence Act.

The Cloud Strategy, FDS, and Evidence Act all call for risk assessment ahead of data release or technology adoption. They encourage or explicitly require safeguards that include improved governance and specific actions to protect privacy. The Cloud Strategy encourages collaboration and coordination across various functional areas, and encourages input from Senior Agency Officials for Privacy (SAOPs) on technology decisions that may impact privacy. Meanwhile, the FDS and Evidence Act encourage governance practices that consider privacy protection, while

[43] Breeden, John. 2014. "Worried about Security? Beware the Mosaic Effect." GCN (blog). May 14, 2014. https://gcn.com/articles/2014/05/14/fose-mosaic-effect.aspx.

[44] "Strategy | Federal Cloud Computing Strategy." 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[45] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[46] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[47] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

explicitly calling for risk assessments to be conducted before agencies open their data. The Evidence Act also includes language that strengthens existing confidentiality and privacy laws.

The Federal Cloud Computing Strategy[48]	• Requires agencies to coordinate between information security and privacy programs to ensure compliance with privacy requirements • Makes Senior Agency Officials for Privacy (SAOPs) responsible for managing risk
The Federal Data Strategy[49]	Requires agencies to: • Exercise responsibility by practicing effective data stewardship, governance, and privacy and security measures • Govern data in a way that provides appropriate access to confidential information while protecting privacy and public trust • Review data release for disclosure risk
The Foundations For Evidence-Based Policymaking Act[50]	• Requires privacy risk assessments, including the risk of re-identification for de-identified data, prior to the public release of data • Reauthorizes and strengthens existing privacy legislation (CIPSEA) for statistical agencies • Requires protection and use of confidential data for statistical purposes only
The Federal Source Code Policy[51]	Requires agencies to consider security and privacy when selecting software solutions required to later switch vendors, and availability of quality support

Governance

These policies require processes and procedures for implementation. Each outlines governance priorities with somewhat different approaches. The Cloud Strategy makes it clear that individual agencies must choose an appropriate governance model for their own cloud-hosted data, while the FDS and Evidence Act lay out policies and best practices that apply across government.

There are also similarities across the policies. For example, the Cloud Strategy and FDS both focus on agencies' responsibility to the public and try to ensure that actions will respond to taxpayer needs. The Source Code Policy and Evidence Act both include inventories to track custom code and data sets respectively. Finally, the Cloud Strategy and Evidence Act both place high importance on security and privacy.

[48] "Strategy | Federal Cloud Computing Strategy." 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[49] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[50] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[51] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

The Federal Cloud Computing Strategy[52]	• Requires agencies to develop agreements with providers regarding access and use of data describing changes to cloud-hosted information • Makes agencies responsible for their own governance model for cloud-hosted data • Makes agency heads responsible for managing enterprise risk, even in contractor managed systems • Requires agencies to articulate roles and responsibilities on a granular level, establish clear performance metrics, and implement remediation plans for noncompliance with their cloud service providers
The Federal Data Strategy[53]	Requires agencies to: • Ensure relevance by ensuring data quality and integrity • Plan for interoperability and future uses from the start • Gather and incorporate stakeholder input through a circular feedback process • Prioritize governance by ensuring that sufficient authorities, roles, organizational structures, policies, and resources are in place • Include data investments in annual capital planning processes to leverage data as a strategic asset
The Foundations For Evidence-Based Policymaking Act[54]	• Requires OMB to compile agency input into a unified report to cross-pollinate best practices • Reduces existing limits on interagency data sharing • Streamlines application processes for accessing restricted data • Requires agencies to develop and maintain open data plans
The Federal Source Code Policy[55]	• Requires agencies to maintain an inventory of all custom software and publish the code on code.gov • Encourages agency leadership to work with public affairs, open government, web, digital strategy, and other staff to identify and collaborate with OSS communities

Workforce

Ensuring that the federal workforce can operate and use the cloud and open data is as important as the new technologies and approaches themselves. The Cloud Strategy, FDS, and Evidence Act recognize the need for a data and technology literate federal workforce, and include approaches for skill development.

[52] "Strategy | Federal Cloud Computing Strategy." 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[53] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[54] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[55] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

All three look to capacity assessments to measure current skills and identify gaps where they exist. The Cloud Strategy and FDS go further by highlighting the need to develop leaders and skills at all levels of government, not just the top. The FDS specifically argues for investments in training, tools, communities, and more in order to build the necessary skills and support learning cultures at agencies. Meanwhile, the Cloud Strategy suggests agencies use a range of strategies—including public-private partnerships, interagency collaboration, and retraining where appropriate—to close skills gaps and build a modern federal workforce.

The Federal Cloud Computing Strategy[56]	• Encourages Agency CIOs, Chief Human Capital Officers (CHCOs), and SAOPs to collaboratively conduct a skills gap analysis that maps current IT workforce resources to future skill and position requirements. Analysis should focus on identifying technical and nontechnical skill and position gaps • Encourages agencies to take a multidisciplinary approach to hiring and training • Encourages agencies to embrace multiple models for workforce transformation, including development programs, apprenticeship programs, initiatives to convert non-IT personnel, or exchanges of personnel through public-private partnerships or interagency details
The Federal Data Strategy[57]	Requires agencies to: • Promote learning cultures through ongoing investment in infrastructure and human resources • Increase capacity by investing in training, tools, and communities • Develop data leadership throughout the workforce by investing in training and development
The Foundations For Evidence-Based Policymaking Act[58]	Requires agencies to: • Conduct periodic capacity assessments and report on their ability to perform statistical evaluation • Conduct and report research and analysis in their daily operations, identifying where key skills and competencies are falling short
The Federal Source Code Policy[59]	Encourages agencies to strengthen internal capacity to efficiently and securely deliver OSS

[56] "Strategy | Federal Cloud Computing Strategy." 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[57] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[58] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[59] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

Procurement

Procurement considerations play a key role in policy implementation, particularly for new software, systems, and services. The Cloud Strategy, FDS, and Source Code Policy all pay at least some attention to procurement. In some cases, guidance in one area may bolster the goals of others. For example, the FDS encourages the use of "collaborative computing platforms" like the cloud, in order to lower costs while improving performance.

The Cloud Strategy and Source Code Policy both encourage analysis ahead of purchasing decisions. The Cloud Strategy specifically focuses on managing security and privacy risks proactively during the procurement process. The Source Code Policy's assessment requirements require identifying the most appropriate method to acquire new software, placing preference on existing federal solutions over the purchase or development of new, potentially duplicative services.

The Federal Cloud Computing Strategy[60]	Requires agencies to: • Consider security and privacy risks and ensure that agencies maintain visibility of their assets in the cloud environments when making cloud procurement decisions • Leverage the government's "bulk purchasing power" and shared knowledge of best practices for procurement • Leverage "category management" to reduce "duplicative contracts" and improve procurement practices
The Federal Data Strategy[61]	Requires agencies to: • Leverage collaborative computing to minimize costs, improve performance, and increase use • Leverage buying power to promote efficiency and reduce costs related to private sector data assets, services, and infrastructure
The Foundations For Evidence-Based Policymaking Act[62]	Requires agency open data plans to have "requirements for meeting the goals of the agency open data plan" that include implementing procurement standards that allow for the acquisition of "innovative" public and private solutions
The Federal Source Code Policy[63]	Requires agencies to: • Conduct strategic analysis and analyze alternatives prior to initiating a technology procurement or custom code development • Consider existing federal and commercial solutions ahead of custom development. If custom code is chosen, agencies should consider OSS • Ensure government rights in contracts with software vendors including the ability to share and use code amongst agencies

[60] "Strategy | Federal Cloud Computing Strategy," 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[61] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[62] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[63] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

Use in Decision Making

While the FDS and Evidence Act specifically focus on using data to guide decision making, the Cloud Strategy also acknowledges that cloud computing can help agencies with decision making through analytics and easier data access.

The FDS focuses on harnessing existing data proactively while using data more generally to guide decision making. The Evidence Act goes even further, directing agencies to consider their evidence-building activities through multiyear learning agendas and evaluation plans.

The Federal Cloud Computing Strategy[64]	Instructs agencies that they cannot outsource decision making to commercial suppliers
The Federal Data Strategy[65]	Requires agencies to: • Prioritize existing data to inform research and policy, and to collect new data where necessary • Use data in policy, planning, and operations to guide decision making
The Foundations For Evidence-Based Policymaking Act[66]	• Prioritizes evidence-building at agencies through multiyear learning agendas and evidence plans • Aims to help each agency collect and analyze data to inform program design and policymaking
The Federal Source Code Policy[67]	Not currently covered in this policy

Interagency Collaboration

Too often agencies work in silos, missing opportunities to share resources and best practices to achieve their missions more efficiently and effectively. Each of the policies acknowledges the need for increased interagency collaboration and outline several ways to achieve results.

The Cloud Strategy and Source Code Policy both recognize that leveraging the power and size of the entire federal enterprise can help negotiate better deals for the American taxpayer. Meanwhile, the Cloud Strategy and FDS recognize the power of collaboration in building a robust, skilled workforce. Finally, the FDS and Evidence Act highlight ways that data sharing can help reduce red tape and make progress towards shared goals.

[64] "Strategy | Federal Cloud Computing Strategy." 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[65] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[66] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[67] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, D.C.: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

The Federal Cloud Computing Strategy[68]	Instructs agencies to: • Consider personnel exchanges or interagency detail opportunities to fill workforce gaps • Leverage the government's "bulk purchasing power" and shared knowledge of best practices for procurement • Leverage "category management" to reduce "duplicative contracts" and improve procurement practices • Promotes the development of common Authority to Operate (ATO) agreements. Reusing ATO agreements can drive "better and more automated control inheritance and monitoring" • Encourages standardizing cloud service level agreements (SLAs) to identify good candidates for SLA use across agencies and improve procurement outcomes for agencies
The Federal Data Strategy[69]	Requires agencies to: • Proactively address existing barriers to data sharing within and across agencies • Connect data functions across agencies through communities of practice • Coordinate and share data assets across agencies • Facilitate data sharing between state, local, tribal, and federal agencies
The Foundations For Evidence-Based Policymaking Act[70]	• Requires OMB to consolidate agency reports into a unified report to foster coordination and sharing of best practices • Reduces existing limits on agency data sharing
The Federal Source Code Policy[71]	Pushes agencies to leverage existing solutions when considering new software needs. Encourages sharing across agencies

Use of Standards

Standards help agencies derive the most value from data, and from software or systems that need integration across users. The FDS, Evidence Act, and Source Code Policy all encourage the use of open standards to increase interoperability and ensure that users are on the same page. Data standards can maximize data quality while facilitating access and use. Meanwhile, open standards for software ensure that software can be used—or adapted and reused—without restrictions and interface with other systems as needed.

[68] "Strategy | Federal Cloud Computing Strategy." 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[69] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[70] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[71] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, D.C.: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

The Federal Cloud Computing Strategy[72]	Not currently covered in this strategy
The Federal Data Strategy[73]	Requires agencies to adopt, adapt, or create data standards to maximize data quality and facilitate use, sharing, and interoperability
The Foundations For Evidence-Based Policymaking Act[74]	• Encourages the consistent application of rules by directing the CDO Council to promote standardization of data and rules throughout government • Defines "open government data" in part as data that is "based on an underlying open standard that is maintained by a standards organization" • Requires development of a repository for tools, best practices, and schema standards to facilitate open data across the government
The Federal Source Code Policy[75]	Urges all government software procurements and development projects to consider utilizing open standards to increase interoperability

Public Value

At the end of the day, federal agencies are responsible to the American public. Each policy analyzed in this paper acknowledges that fact and takes steps to promote principles of public value and access.

The Cloud Strategy focuses on the cloud's ability to help agencies fulfill their missions and deliver services to the public more effectively. Similarly, the Source Code Policy ensures that the government maintains control over code that it develops in order to ease sharing between agencies and save money for the public.

The Source Code Policy also aims to encourage community through engagement on open source code. It recognizes that sharing code publicly can help address shared challenges, build new communities, and ultimately improve the code used by government agencies. Similarly, the Evidence Act acknowledges that open data and evidence building should help improve public trust by requiring the government to maintain objectivity, independence, and confidentiality where necessary.

The FDS takes this issue of trust even further, explicitly asking government agencies to uphold ethics and improve public communication and transparency around data use while promoting wider public access to government information.

[72] "Strategy | Federal Cloud Computing Strategy." 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[73] "Welcome - Federal Data Strategy." 2019. Data.gov. 2019. https://strategy.data.gov/.

[74] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[75] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

The Federal Cloud Computing Strategy[76]	• Stresses that federal IT is integral to the delivery of services to the public. Notes that leveraging modern technologies and practices can help agencies harness new capabilities and expand existing abilities to enable their mission and deliver services to the public faster • Encourages agencies to make cloud adoption decisions based on their ability to meet mission goals and act as good stewards for taxpayer-provided funds
The Federal Data Strategy[77]	Requires agencies to: • Design checks and balances and monitor federal data practices to protect the public and ensure ethical data use • Document processes and products and articulate the reasons for data use to build trust • Promote access to machine-readable data through multiple channels including nonfederal paths to meet stakeholder needs • Communicate planned and potential uses of data
The Foundations For Evidence-Based Policymaking Act[78]	• Requires agencies to provide public access to collected data whenever possible and provide it in an open and machine-readable format • Pushes the government to maintain objectivity, independence, and confidentiality in evidence-building by requiring statistical information and the interpretation of data be “accurate,” “unbiased,” and “credible”
The Federal Source Code Policy[79]	• Requires agencies to secure the full scope of the government’s rights, including—but not limited to—sharing and using the code with other federal agencies • Encourages agencies to develop and release OSS code to foster communities, feedback from nongovernmental users, and federal employees and contractors to contribute to external OSS communities

[76] “Strategy | Federal Cloud Computing Strategy.” 2017. Cio.gov. 2017. https://cloud.cio.gov/strategy/.

[77] “Welcome - Federal Data Strategy.” 2019. Data.gov. 2019. https://strategy.data.gov/.

[78] Foundations for Evidence-Based Policymaking Act. H.R.4174. (2018). https://www.congress.gov/bill/115th-congress/house-bill/4174/text.

[79] Tony Scott and Anne E. Rung, “Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software” (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

Topical Intersections and Takeaways

CODE's research and interviews, and our roundtable, found many ways in which these "open" approaches intersect in their potential. There are intersections between work on the cloud and open data, between the cloud and open source, and between open source and open data.

Cloud and Open Data

The cloud is a natural pathway to increased data sharing and can support open data programs. However, there are challenges to realizing that potential, given the cost and varying requirements and needs among agencies. Making open data available on the cloud raises concerns for agencies that operate on fixed cost budgets and may not be able to handle the additional costs of cloud hosting if use of a dataset unexpectedly increases—a common model for cloud contracts.

Agencies differ on the path to data sharing and cloud adoption and have different computing needs. Some have strong technical infrastructure but poor data quality; others have good data, but trouble sharing it; and some have both data and technical challenges. Agencies with large amounts of data may already use supercomputers for analytics and be reluctant to switch to the cloud, while agencies with smaller data holdings may find it easier to host their data without leveraging the cloud at all.

Despite these challenges, the cloud has already proved its worth as a pathway for sharing open data and making it easier to use. For example, NOAA's Big Data Program has successfully improved access to open data for the public and other agencies while controlling costs to government and providing additional analytical capabilities.[80] For agencies with large open data portfolios who may already have invested in high performance computing (HPC) infrastructure, the cloud can help with pre-planning for HPC use, by allowing data scientists to run initial analysis or check results without using limited HPC time—and providing a convenient way to break down and share data publicly in more manageable pieces.

The COVID-19 pandemic has helped highlight the value of the cloud for data sharing and remote work. One roundtable participant shared the experience of providing secure cloud access to data from the Census Bureau for researchers who previously had to travel to physical locations. Although some leaders within the Census Bureau had already pushed for cloud access to these restricted data sets, the pandemic drove the Bureau to overcome inertia and cultural resistance to this change.

Other agencies have also begun to develop similar cloud-based programs. The Centers for Medicare and Medicaid Services (CMS) runs a Virtual Research Data Center (VRDC), which "provides timelier access to Medicare and Medicaid program data in a more efficient and cost effective manner" than other alternatives.[81] This approach could be implemented on a broader, cross-agency scale through a National Secure Data Service or similar body, as recommended by the U.S. Commission on Evidence-Based Policymaking.[82]

[80] "Big Data Program | National Oceanic and Atmospheric Administration." n.d. www.noaa.gov. Accessed February 18, 2021. https://www.noaa.gov/organization/information-technology/big-data-program.

[81] "CMS Virtual Research Data Center (VRDC) | ResDAC." n.d. Www.resdac.org. Accessed February 18, 2021. https://www.resdac.org/cms-virtual-research-data-center-vrdc.

[82] Hart, Nick, and Nancy Potok. 2020. "Modernizing U.S. Data Infrastructure: Design Considerations for Implementing a National Secure Data Service to Improve Statistics and Evidence Building." Data Foundation. https://www.datafoundation.org/modernizing-us-data-infrastructure-2020.

CODE's interviews and the roundtable confirmed that transitioning to cloud infrastructure represents a prime opportunity to inventory and clean an agency's data with an eye towards opening it up. However, user-driven strategies should guide cloud migration and open data sharing, with priority given to future data needs and the most important data sets. Once cloud infrastructure and processes exist, legacy and lower-priority data can be added to the cloud.

Cloud and Open Source

OSS solutions may provide value to agencies as they move to the cloud by providing increased flexibility and lowering barriers to leveraging the cloud's computing resources. Additionally, "moonshot" projects that benefit from cross-organizational collaboration also benefit from OSS solutions during development, deployment, and distribution. Furthermore, OSS has played a significant role in the development of cloud computing. For example. Kubernetes is an increasingly popular, open source "system for automating deployment, scaling, and management of containerized applications."[83]

OSS can give agencies flexibility as they choose between competing cloud providers, or manage between operating divisions that may contract with competing providers. Agencies may feel locked into a specific cloud vendor after making large, upfront investments. Conversely, different agency components may rely on different cloud providers, causing integration challenges. OSS can help agencies better understand their varying cloud environments, and integrate across or within these domains.

OSS can also enable agencies and outside stakeholders to use cloud computing resources. One roundtable attendee suggested using OSS when training employees in cloud technologies, since this is often cheaper and easier to take up than proprietary tools. For example, R, a popular statistical computing tool, is free, open source, and relatively easy to learn.[84]

Finally, OSS is a vital part of so called "moonshot" projects tackled via collaboration across large groups. For example, much of the most valuable, early work on COVID-19 data was done by diverse communities using open source tools. That open data was then distributed via open channels.

Roundtable attendees highlighted a number of important considerations for agencies to maximize the value of OSS as they move to the cloud. To align workforce and technology, roundtable participants suggested a hiring initiative to bring in more employees with technology skills that specifically align with open source goals and values. Agencies should also proactively identify the value in OSS solutions and build OSS foundations, before moving into the cloud or other new technologies. This will help them optimize software and data for the cloud, rather than carrying over old systems and ways of working to the new technology.

Open Source and Open Data

Combining OSS and open data provides advantages including flexibility, enterprisewide efficiencies, improved quality, and the ability to build trust in government activities. However, challenges have limited their robust use across the federal enterprise, including the perception that open source and open data conflict with privacy and security, which the roundtable indicated are manageable issues as long as they are considered early in the OSS or open data process.

[83] "Kubernetes." n.d. Kubernetes.io. Accessed April 2, 2021.

[84] "R: The R Project for Statistical Computing." 2019. R-Project.org. 2019. https://www.r-project.org/.

OSS and open data can save organizations money and lead to improved quality control for federal agencies. Active OSS projects and popular open data sets are reviewed by large groups, making it more likely to detect and fix bugs or data issues. OpenStreetMap,[85] for example, combines open source and open data principles to provide accurate map data for thousands of organizations and projects,[86] ranging from Amazon to the governments of Italy, Lithuania, and New York City. OpenStreetMap's approach enables more functionality than maps for everyday uses by providing accurate information to disabled communities,[87] as well as serving up-to-date information during disaster scenarios.[88] Ethical concerns, such as algorithmic bias in artificial intelligence (AI) applications, can also be addressed with more open and transparent approaches.

Concerns around privacy and security present major challenges for agencies that want to implement OSS and open data. However, roundtable participants argued that security and privacy should not block progress, as long as they are considered early in developing OSS or open data sets. As CODE has described[89] in earlier[90] work, a growing number of technical and data management approaches make it possible to share sensitive data with privacy protection.

Ultimately, openness allows agencies to launch and scale projects more rapidly, and allows for data and software reuse, expansion, and improvement by a variety of actors—including other agencies as well as the public. A network effect emerges when multiple agencies use the same OSS. For example, a number of agencies use DKAN, a Drupal version of CKAN, to support open data programs—ultimately, allowing them to both contribute to and benefit from the broader DKAN community.

Government data and software, at their core, belong to the public. Whether developed by or for government organizations, they are paid for with public money, and this impacts decisions about their openness. The public will often find novel uses and derive additional value from open software and open data. Opening data and sharing code can also help rebuild trust in government, by giving the public more control and improving the quality of government data and software services.

CODE's interviews, research, and the roundtable provided insights that can help the federal government procure and develop more OSS and release more of its data to more potential users. They also showed that open principles do not conflict with privacy and security. Open data and code are ultimately more useful and contribute more to mission success when shared.

[85] "OpenStreetMap." 2019. OpenStreetMap. 2019. https://www.openstreetmap.org/about.

[86] "Who Uses OpenStreetMap? | OpenStreetMap." n.d. Welcome.openstreetmap.org. Accessed February 18, 2021. https://welcome.openstreetmap.org/about-osm-community/consumers/.

[87] Mikel. 2020. "The Best World Map for Accessibility." OpenStreetMap Blog (blog). December 2, 2020. https://blog.openstreetmap.org/2020/12/02/the-best-world-map-for-accessibility/.

[88] "Humanitarian OpenStreetMap Team | Disaster Response." n.d. www.hotosm.org. Accessed February 19, 2021. https://www.hotosm.org/impact-areas/disaster-response/.

[89] "Balancing Privacy with Health Data Access." 2019. Center for Open Data Enterprise (CODE). https://healthdatasharing.org/wp-content/uploads/2020/07/RT2-Privacy-Summary-Report-FINAL-2020.07.28.pdf.

[90] Gurin, Joel, Matt Rumsey, Audrey Ariss, and Katherine Garcia. 2017. "Protecting Privacy While Releasing Data: Strategies to Maximize Benefits and Mitigate Risks." In The Social Dynamics of Open Data, 183–200. OpenAIRE: African Minds. https://zenodo.org/record/1117782#.YC_UMWRKhmr.

Recommendations

CODE's research and interviews, and the roundtable, provided many ideas for federal agencies to improve and expand their use of the cloud, open data, and open source. This section provides both high-level recommendations for operating principles, and recommendations for action by the Biden administration. Three themes—resources, leadership, and policy—emerged across the project and touch on many of the recommendations. This section addresses these themes, which will be explored in greater detail in the conclusion.

High-Level Recommendations

Maintain flexibility

Agencies should avoid one-size fits-all approaches, because they have different needs and operate at different scales and on different timelines. Early stage efforts to mandate cloud adoption were often too restrictive; at the state level, most have since pivoted to more flexible policies that allow agencies to make their own cloud adoption decisions. Similarly, federal agencies have a range of needs and challenges in cloud adoption, data needs, and technical capacity to adopt OSS.

Share success stories

Agencies have made different levels of progress in adopting cloud technologies, embracing OSS, and opening data. Highlighting success stories can help inspire other agencies as they work towards change. Embracing open principles can also help agencies scale more quickly. For example, open source security policies that build compliance into code can be shared in machine readable formats, which can help other agencies with their own implementation and efforts to get FedRAMP approval or ATO.

Build the workforce

The current federal workforce is not well matched to the demands of a government driven by cloud, OSS, and open data. Training/reskilling efforts for current employees, the addition of a data science job series to the General Schedule (GS) classification and pay system, and dedicated efforts to recruit new talent with appropriate skills could all help improve workforce capacity. A centralized, cross-government approach to technical workforce development could enable trained personnel to move between agencies as needs evolve.

Apply user-driven approaches

Agencies have recently begun to take a much more customer- or user-centered approach to operations and technology. This user orientation should continue and become even more central to agency technology programs. For example, the General Services Administration (GSA) has listened to user complaints about the difficulties in becoming become a vendor for the federal government. In response, GSA has modernized systems for doing business with agencies to open up more opportunities for small or startup enterprises.

Learn from the pandemic

Government agencies' adaptation to the COVID-19 pandemic—particularly the acceleration of remote work and increased sharing of scientific information—provide models for implementing open methods. Cloud adoption is more necessary than ever during and after the pandemic, with most employees needing remote access to systems and data. Meanwhile, openly sharing scientific data and research results can help fight COVID, while also helping prepare for the next global public health crisis. Leaders should apply these lessons as they plan for the future, rather than waiting for the next crisis or inflection point.

Recommendations for Action

Fund Evidence Act implementation

Fully funding Evidence Act implementation could help advance goals across all three types of open initiatives described in this paper, specifically in the areas of workforce development, technology adoption, and interagency collaboration.

Effective data sharing requires more than just strong policies and governance approaches. It also "requires investments in hardware, software . . . network infrastructure," and a workforce that can manage it all in the service of data sharing within the federal enterprise and with the public.[91] As it stands, chief data officers and agency data programs are often underfunded and understaffed. Evidence Act funding could help agencies purchase necessary technology, develop their workforces, and create better connections with other agencies.

In 2020, the Data Coalition, a nonprofit organization that supports open data and related policies, called for $2 million for each CFO Act agency to "to specifically encourage development of leadership and expertise for these newly appointed officials under the Evidence Act," as well as an "interagency transfer fund to allocate additional resources to leading agencies of up to $50 million . . . to launch pilot projects or support governmentwide coordination of certain aspects of data governance, management, and analysis capabilities."[92] This request as well as additional resources may be needed to fund Evidence Act implementation.

Leverage and build on the TALENT Act to build a data- and cloud-literate workforce

The Tested Ability to Leverage Exceptional National Talent (TALENT) Act was the last bill signed into law by President Obama.[93] It codified and gave GSA responsibility for the Presidential Innovation Fellows (PIF) program, making it easier to recruit software engineers, designers, and other technically skilled candidates into the government for one-year engagements. TALENT Act implementation should continue with a full complement of PIF's chosen every year, and its ethos expanded to bring technical talent into government agencies.

The positive example of the TALENT Act and PIF program is already being expanded upon to bring larger numbers of technically skilled individuals into government service. For example, the United States Digital Service (USDS) recently released[94] a data scientist job posting[95] to hire for open positions at 10 different agencies. This sort of initiative does not represent a structural change, but can serve as an example and starting point to help build a more technology-literate workforce across the federal enterprise. Ultimately, the Office of Personnel Management should create new job series to represent data scientists and other increasingly important technical roles.

[91] Wiseman, Jane. 2020. "Silo Busting: The Challenges and Success Factors for Sharing Intergovernmental Data." IBM Center for The Business of Government. http://www.businessofgovernment.org/sites/default/files/Silo%20Busting.pdf.

[92] Hart, Nicholas. 2020. "SUBJECT: RECOMMENDATIONS FOR FISCAL YEAR 2021 APPROPRIATIONS TO SUPPORT IMPLEMENTATION OF FOUNDATIONS FOR EVIDENCE-BASED POLICYMAKING ACT OF 2018 AND OTHER DATA PRIORITIES IN GOVERNMENT." Data Coalition. March, 13, 2020. http://www.datacoalition.org/wp-content/uploads/2020/11/Copy-of-HOUSE-FY2021AppropsRequests.DataCoalition.pdf.

[93] TALENT Act. H.R.39. (2017). https://www.govtrack.us/congress/bills/115/hr39/summary.

[94] "Data Scientist Positions Open across the Federal Government." n.d. Subject Matter Expert Qualification Assessments. Accessed February 19, 2021. https://smeqa.usds.gov/info/data-govwide/.

[95] "Data Scientist." n.d. USAJOBS. Accessed February 19, 2021. https://www.usajobs.gov/GetJob/ViewDetails/588499000.

Provide vouchers for researchers, nonprofits, and others to use cloud resources for research, analysis, and policymaking

Individuals or small organizations often lack the incentive or resources to use cloud-hosted data for analysis and decision making. For example, scientists with access to high performance computing (HPC) resources on site are likely to use those resources and their own data, rather than explore open data hosted in the cloud. While HPC time is built into their costs, paying for cloud computing time adds to their budget. Meanwhile, citizen scientists lack resources to access HPC or cloud analytics.

To encourage more use of open data in the cloud for scientific analysis, some agencies have started to make it easier for their grantees to access and use the cloud. For example, the National Institutes of Health (NIH) has developed the Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES) initiative.[96] STRIDES "allows NIH to explore the use of cloud environments to streamline NIH data use by partnering with commercial providers" to provide scientists with discounts and credits to use cloud resources. These and similar programs could be considered for expansion or application in nonscientific domains, where use of cloud resources and open data could be better encouraged. One interviewee suggested a large-scale voucher program to allow researchers, nonprofits, and other stakeholders to access and use cloud resources for analysis. Research, data, and insights created through this sort of program should be considered government products and subject to existing requirements for openness.

Update the Source Code Policy to encourage government use of existing, robust OSS products and further engagement with OSS projects and communities

While the Source Code Policy clarifies OSS as an option for agencies as they assess their software needs, agencies could make greater use of OSS products and participation in existing open source projects. For example, the Source Code Policy currently requires agencies to use a three-step process when considering their software needs. This process ostensibly puts OSS on a "level playing field" with proprietary solutions, but in practice fails to give appropriate weight to existing open source products that may fit the government's needs.[97] Instead, the process highlights prioritizing existing government software (a cost efficient goal), exploring existing commercial services, and assessing custom software development, in that order. The policy could be strengthened by specifically prioritizing commercial off-the-shelf (COTS) products built using OSS over products built with proprietary software. As a follow-up, the Government Accountability Office could examine the effectiveness of the Source Code Policy's pilot project requiring that 20 percent of new custom code developed by the federal government be released as OSS. That study could analyze and recommend ways to encourage deeper federal engagement with OSS projects and communities.

More broadly, agencies should be encouraged to participate more closely with the private sector to build and maintain open source communities. The public sector's current lack of engagement gives the appearance of avoiding conflicts of interest, but in practice it limits the government's ability to leverage and build on existing private sector instances of successful open source products. In fact, participating in public, open source communities along with private entities could help the government avoid favoritism while ensuring that open source products continue to serve their needs over time.

[96] "STRIDES Initiative | Data Science at NIH." n.d. Datascience.nih.gov. Accessed February 19, 2021. https://datascience.nih.gov/strides.

[97] Tony Scott and Anne E. Rung, "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software" (official memorandum, Washington, DC: Executive Office of The President, 2016), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2016/m_16_21.pdf.

Update the key federal data policies to better align them and fill gaps

The FDS, Evidence Act, Cloud Strategy, and Source Code Policy already align across several thematic areas. However, CODE's research shows the need for more alignment across the documents. Additionally, specific gaps in the policies could be filled, where one policy fails to address important issues that are covered by the others. For example:

- The Source Code Policy only obliquely addresses the issue of workforce. It could be updated to include more specific language supporting training and hiring OSS literate talent, in order to boost engagement with OSS projects.
- The Cloud Policy does not currently call for the use of interoperable data or software standards. This omission could ultimately limit agencies' ability to easily transition between cloud providers or share data more widely. The policy could be updated to require the use of interoperable standards in cloud projects, drawing on language already included in the FDS, Evidence Act, and Source Code Policy.

Develop a plain language toolkit to explain how these policies intersect, what they mean for agencies, and how to implement them in ways that build strength across the three areas

A plain language toolkit could address what the policies mean for agencies, and how to implement them in ways that build strength across the three areas. While the Federal Data Strategy and Code.gov present some plain language information, documents could explain all three of these policies and how to implement them in simple, readable language. A plain language toolkit would explain each policy, show how they intersect, and lay out approaches to implementing them that build strength across the three areas. This could be led by a central government office like OMB or the GSA, or a nonprofit or academic organization with interest in these areas.

Leverage agency data inventories to understand agency data systems

CODE's research highlighted the inefficiency caused by the lack of agencywide cloud adoption strategies, which has resulted in nearly every major cloud service provider having a presence in nearly every major agency. Agencies may lack an internal understanding of their cloud contracts, and how the cloud could be used strategically to improve the value of their data assets. Agencies can leverage ongoing efforts to inventory their data to better understand the systems on which their data live, and to make more informed decisions about how they leverage cloud technologies.

CONCLUSION

The recommendations above could help promote and coordinate progress on open source, open data, and cloud adoption. Progress on these recommendations will require a unified approach that covers three key areas: providing adequate resources, establishing leadership that can change culture and bring the federal workforce along, and aligning across policy areas in new ways.

Resources

Funding and other resources are critical to modernize systems and fully leverage data for decision making. Sufficient resources would support foundational work such as establishing data governance and building technical infrastructure, and implementation projects like technical upgrades and data sharing and migration.

Agencies have a range of needs based on their size and function. Small agencies may lack technical capacity, while large agencies have to manage large and complex systems. To that point, resources should be spread strategically so that all agencies can modernize technologies and embrace data sharing and evidence-based policymaking, in ways that make the most sense for their particular needs.

Leadership

Changing culture at federal agencies requires strong, consistent leadership over time. Given how often leadership turns over in government, mechanisms to help transition between leaders would be helpful—particularly with respect to offices that deal with technology transformation and long-term systems change. For example, Obama administration priorities around open data and evidence building were maintained and moved forward in the Trump administration, thanks to continuity of leadership in key management functions as well as ongoing and bipartisan leadership and oversight from Congress.

When it comes to implementation, leadership at the OMB, the CIO and CDO Councils, and agency CIOs and CDOs should focus on aligning efforts and finding synergies across the domains of open data, open source, and cloud. While programs and business units should have flexibility to make their own decisions with respect to cloud adoption, data sharing, and software acquisition, agencywide governance and planning will help integrate across the enterprise. CIOs and CDOs should play complementary and not competing roles in this effort. For example, CIOs can focus on systems and technology while CDOs work on data governance; the two roles can work together on integration challenges. Importantly, both roles should participate in decision making on technology policy and operations.

Policy alignment

Beyond efforts to align the source code policy with open data policies, federal policies governing open data, open source, and cloud adoption have proceeded on separate tracks. More can and should be done to align all three.

The FDS and the Evidence Act are policy imperatives—and statutory requirements, in the case of the Evidence Act—while cloud and open source have been treated as "nice to haves." Data policy imperatives could be leveraged to make advancements on cloud and open source. Currently, CIOs are very engaged in Evidence Act and FDS implementation—which provides useful leverage to increase data sharing at agencies—but CIOs should also consider how the two imperatives can help advance cloud and open source goals.

Government leaders can look to open data, open source software, and the cloud in advancing management priorities and modernizing federal IT and decision processes. A unified approach that considers open data, open source software, and cloud adoption as interrelated pieces of a larger puzzle will help achieve those priorities.

ACKNOWLEDGEMENTS

Interviews and Roundtable Attendance

CODE and the IBM Center would like to thank the following individuals for participating in interviews or the roundtable meeting for this project:

- **Madina S. Ali** (Roundtable only)
- **Phil Ashlock** (Interview only)
- **Mark Bolter** (Interview and roundtable)
- **Debbie Brodt-Giles** (Interview only)
- **Dwight Chamberlain** (Roundtable only)
- **David Egts** (Roundtable only)
- **Margie Graves** (Interview and roundtable)
- **Nick Hart** (Interview and roundtable)
- **Jed Herrmann** (Interview and roundtable)
- **Alex Howard** (Interview and roundtable)
- **Joe Klimavicz** (Roundtable only)
- **Sanjay Koyani** (Interview and roundtable)
- **Stephen LeNard** (Roundtable only)
- **Katie Malague** (Roundtable only)
- **Dave McClure** (Roundtable only)
- **Travis Methvin** (Roundtable only)
- **Rick Miller** (Roundtable only)
- **Jonathan O'Neil** (Roundtable only)
- **Nancy Potok** (Roundtable only)
- **Maria Roat** (Interview and roundtable)
- **Michael Tiemann** (Roundtable only)
- **Meredith Ward** (Roundtable only)
- **John Wilbanks** (Interview only)
- **Jane Wiseman** (Roundtable only)
- **Elias Yishak** (Roundtable only)

ABOUT THE AUTHORS

Matt Rumsey is the research and communications manager at CODE. He has experience in open government data as an advocate, policy analyst, and researcher. Prior to joining CODE, Matt worked as an independent researcher with organizations including the Data Foundation, CODE, and the Sunlight Foundation. He conducted research and wrote policy and briefing papers on topics including DATA Act implementation, grants data standardization, legal entity identification, the interaction between open data and privacy, and open research data. He got his start in open data at the Sunlight Foundation, where he worked on federal policy initiatives. He advocated for passage and effective implementation of the DATA Act, conducted research and advocacy around executive branch open data efforts, and helped to conceptualize and draft what would become the OPEN Government Data Act. Matt has a B.A. in history from American University in Washington, D.C. You can email him: Matthew@odenterprise.org.

MATT RUMSEY

Joel Gurin is the president and co-founder of the Center for Open Data Enterprise (CODE) and an internationally recognized expert on open data. His book *Open Data Now* (McGraw-Hill), written for a general audience, is considered a benchmark publication that helped define this emerging field. Before launching CODE in January 2015, he conceptualized and led the development team for the GovLab's Open Data 500 project, the first thorough study of the use of open government data by the private sector. Joel's background includes government, journalism, nonprofit leadership, and consumer issues. He served as chair of the White House Task Force on Smart Disclosure, which studied how open government data can improve consumer markets, and as chief of the Consumer and Governmental Affairs Bureau of the U.S. Federal Communications Commission. For more than a decade he was editorial director and then executive vice president of Consumer Reports, where he directed the launch and development of ConsumerReports.org, which was then the world's largest paid-subscription information-based website. He is a graduate of Harvard University with an A.B. in biochemical sciences, Magna Cum Laude, Phi Beta Kappa. You can follow him on Twitter at @joelgurin or email him: joel@odenterprise.org.

JOEL GURIN

Additional research for this project was conducted by Nidhisha Phillip, who was a research fellow at CODE from July 2019 through July 2020.

Special thanks to Dan Chenok, Ruth Gordon, Susan Murphy, Maureen Carlin, and Elias Yishak from IBM, and Temilola Afolabi, Sophie Hu and Paul Kuhne from CODE for their help planning and executing the roundtable. Additional thanks to David Egts at Red Hat and Margie Graves at IBM for their guidance throughout the project and helpful comments on the draft of this manuscript.

KEY CONTACT INFORMATION

To contact the authors:

Matt Rumsey
Center for Open Data Enterprise
matthew@odenterprise.org

Joel Gurin
Center for Open Data Enterprise
joel@odenterprise.org

REPORTS FROM THE IBM CENTER FOR THE BUSINESS OF GOVERNMENT

For a full listing of our publications, visit www.businessofgovernment.org

Recent reports available on the website include:

Agility:

The Road to AGILE GOVERNMENT: Driving Change to Achieve Success by G. Edward DeSeve
Transforming How Government Operates: Four Methods of Change by Andrew B. Whitford
Agile Problem Solving in Government: A Case Study of The Opportunity Project by Joel Gurin, Katarina Rebello
Applying Design Thinking To Public Service Delivery by Jeanne Liedtka, Randall Salzman

Digital:

Innovation and Emerging Technologies in Government: Keys to Success by Dr. Alan R. Shark
Risk Management in the AI Era: Navigating the Opportunities and Challenges of AI Tools in the Public Sector by Justin B. Bullock, Matthew M. Young
More Than Meets AI: Part II by the Partnership for Public Service, The IBM Center for The Business of Government
Financial Management for The Future: How Government Can Evolve to Meet the Demands of a Digital World by Angela Carrington, Ira Gebler
The Impact of Blockchain for Government: Insights on Identity, Payments, and Supply Chain by Thomas Hardjono
A Roadmap for IT Modernization in Government by Dr. Gregory S. Dawson

Effectiveness:

Federal Grants Management: Improving Outcomes by Shelley H. Metzenbaum
Government Reform: Lessons from the Past for Actions in the Future by Dan Chenok, John Kamensky
COVID-19 and its Impact: Seven Essays on Reframing Government Management and Operations by Richard C. Feiock, Gurdeep Gill, Laura Goddeeris, Zachary S. Huitink, Robert Handfield, Dr. Rodney Scott, Sherri Greenberg, Eleanor Merton, Maya McKenzie, Tad McGalliard
How Localities Continually Adapt Enterprise Strategies to Manage Natural Disasters by Katherine Willoughby, Komla D. Dzigbede, Sarah Beth Gehl
Measuring the Quality of Management in Federal Agencies by James R. Thompson, Alejandra Medina
Mobilizing Capital Investment to Modernize Government by Steve Redburn, Kenneth J. Buck, G. Edward DeSeve

Insight:

Making Federal Agencies Evidence-Based: The Key Role of Learning Agendas by Dr. Kathryn E. Newcomer, Karol Olejniczak, Nick Hart
Improving Outcomes in Government through Data and Intelligent Automation by The IBM Center for The Business of Government, Partnership for Public Service
Silo Busting: The Challenges and Successes of Intergovernmental Data Sharing by Jane Wiseman
Integrating Big Data and Thick Data to Transform Public Services Delivery by Yuen Yuen Ang
A Practitioner's Framework for Measuring Results: Using "C-Stat" at the Colorado Department of Human Services by Melissa Wavelet
Data-Driven Government: The Role of Chief Data Officers by Jane Wiseman

People:

Distance Work Arrangements: The Workplace of the Future Is Now by John Kamensky, Emily G. Craig, Michaela Drust, Dr. Sheri I. Fields, Lawrence Tobin
Preparing the Next Generation of Federal Leaders: Agency-Based Leadership Development Programs by Gordon Abner, Jenny Knowles Morrison, James Perry, Bill Valdez
Assessing the Past and Future of Public Administration: Reflections from the Minnowbrook at 50 Conference by Tina Nabatchi, Julia L. Carboni

Risk:

The Rise of the Sustainable Enterprise by Wayne S. Balta, Jacob Dencik, Daniel C. Esty, Scott Fulton
Managing Cybersecurity Risk in Government by Anupam Kumar, James Haddow, Rajni Goel

About the IBM Center for The Business of Government

Through research stipends and events, the IBM Center for The Business of Government stimulates research and facilitates discussion of new approaches to improving the effectiveness of government at the federal, state, local, and international levels.

About IBM Global Business Services

With consultants and professional staff in more than 160 countries globally, IBM Global Business Services is the world's largest consulting services organization. IBM Global Business Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues, and the ability to design, build, and run those solutions in a way that delivers bottom-line value. To learn more visit ibm.com.

For more information:
Daniel J. Chenok
Executive Director
IBM Center for The Business of Government

600 14th Street NW
Second Floor
Washington, DC 20005
202-551-9342

website: www.businessofgovernment.org
e-mail: businessofgovernment@us.ibm.com

Stay connected with the IBM Center on:

or, send us your name and e-mail to receive our newsletters.

IBM Center for
The Business of Government