

UNDERSTANDING THE FEDERAL DATA STRATEGY

Embracing the future of federal data management

INTRODUCTION

The Federal Data Strategy (FDS) describes a 10-year vision for how the Federal Government will accelerate the use of data to deliver on mission, serve the public, and steward resources while protecting security, privacy, and confidentiality.¹ In order to better understand the FDS, how it is being implemented, the challenges to its implementation, and how it will benefit government agencies and the public, Government Business Council (GBC) conducted a qualitative research campaign bringing together perspectives from experts within the federal government.

RESEARCH METHODOLOGY

GBC in partnership with Snowflake launched a qualitative research campaign in September, 2020. From December 2020 to January 2021, GBC conducted interviews with experts from the federal government who provided insight on the Federal Data Strategy, along with its progress, benefits, and challenges. The list of featured interviewees is as follows:



David Spirk
Chief Data Officer,
U.S. Department of Defense



Allen Hill
Deputy Assistant Commissioner for
Category Management, Office of
Information Technology Category,
*General Services Administration's Federal
Acquisition Service*



Dr. Matthew Graviss
Chief Data Officer,
U.S. Department of State

FEDERAL DATA STRATEGY

BACKGROUND: 20 ACTION ITEMS

In March 2018, the President's Management Agenda laid out a new Cross-Agency Priority Goal: Leveraging data as a strategic asset to develop and implement a comprehensive Federal Data Strategy. Based on this initiative the first government-wide data strategy was developed, along with a 2020 Action Plan for implementation.²

MISSION STATEMENT:

The mission of the Federal Data Strategy is to leverage the full value of Federal data for mission, service, and the public good by guiding the Federal Government in practicing ethical governance, conscious design and a learning culture.³

- | | | | | |
|---|---|---|---|---|
| one Identify data needs to answer priority agency questions | two Constitute a diverse data governance body | three Assess data and related infrastructure maturity | four Identify opportunities to increase staff data skills | five Identify priority datasets for agency open data plans |
| six Identify priority datasets for agency open data plans | seven Launch a federal Chief Data Officer Council | eight Improve data and model resources for AI research and development | nine Improve financial management data standards | ten Integrate geospatial data practices into the federal data enterprise |
| eleven Develop a repository of federal enterprise data resources | twelve Create an OMB Federal Data Policy Committee | thirteen Develop a curated data skills catalog | fourteen Develop a data ethics framework | fifteen Develop a data protection toolkit |
| sixteen Pilot a one-stop standard research application | seventeen Pilot an automated tool for information collection reviews that supports data inventory creation and updates | eighteen Pilot enhanced data management tool for federal agencies | nineteen Develop data quality measuring and reporting guidance | twenty Develop a data standards repository |

THE INTERVIEWS

Could you please provide a visual for our readers of why data management/analytics needed to be modernized in the federal government?

DS DAVID SPIRK

I think that we've all come to recognize that in the past we probably didn't place the right amount of value on our data. I think we saw it a bit as an afterthought, the exhaust, when we were putting systems in place. But at this point, we have our senior-most leaders recognizing that data is probably the most valuable asset that we have in our inventory. If used properly, and if managed properly, the decisions we can unlock, and the speed with which we can do that will allow us to retain that competitive advantage in the security space ... I think that really connects not just to us, but also to our partners and allies.

AH ALLEN HILL

We have too much duplication of data ... I think the intent of the federal data strategy is to make sure you know what data is out there and how it can be accessed so we're not duplicating it and we can make sure that it's secure.

MG DR. MATTHEW GRAVISS

Even in the pandemic, you're seeing why data is so important ... Our global workforce [at State]

is operating in more than 200 countries, and our mission is really giving them accurate information that helps them achieve their mission. So we need to put insights in the hands of all of our operators across the board. I would also say the global landscape is changing rapidly. The need for making decisions on the diplomatic front at a fast pace has never been more prevalent than it is today. To me, it's really important that we use data and analytics to make decisions faster ... Equipping the field with data and training is going to be really important as we move forward to modernize our data capabilities at the department.

“Data is probably the most valuable asset that we have in our inventory.”

- David Spirk

To what extent has the COVID-19 situation accelerated (if at all) the progression of agencies with the Federal Data Strategy?

DS DAVID SPIRK

COVID has accelerated implementation of the Data Strategy in several ways. First, it prompted agencies to invest in modern, cloud-based collaboration tools to support a dislocated workforce. The DoD was already moving in this direction as part of its comprehensive Digital Modernization Program. It also provided an urgent operational need for the federal government to implement data sharing in order to coordinate our response to the crisis. Everyone has been very motivated to knock down obstacles to effective data sharing. Thirdly, it made clear the important role that data access and modern IT plays in enabling everyone to accomplish their work. CIOs (Chief Information Officers) and CDOs (Chief Data Officers) were seen as key players in providing mission resilience. These are hard won lessons and the results will be enduring.

AH ALLEN HILL

[COVID-19] emphasized the importance of modernization of the agency's infrastructure .. How your network infrastructure is structured is extremely important ... The way that we are structured now, a lot of agencies still have a legacy network architecture ... COVID-19 brought to light how we could [enhance] our infrastructure support in such a situation.

MG DR. MATTHEW GRAVISS

If anything, [COVID-19 has] been an accelerant that has underlined the need for the Federal Data Strategy. The pandemic kind of ripped the band aid off and forced people to accept change because they had to, rather than take a longer approach to cultural change. It really pushed us to be more flexible and innovative and on the IT side using more mobility and technology applications from that front. It's shown us that we're more motivated, certainly in this case, by an external force. It teaches us that we can in fact make changes quickly if we have the impetus for it.

Do you foresee this momentum carrying forward even after the COVID-19 situation as agencies begin to move back to the office?

DS DAVID SPIRK

Absolutely. As noted earlier, these are hard won lessons and the results will be enduring. DoD leaders, including the [Acting Defense]Secretary Norquist and [former] DoD CIO Dana Deasy, have made clear that modern data driven technologies are key enablers of enterprise resilience going forward. The DoD is applying insights from COVID challenges as we develop an information enterprise that is modern, secure, scalable, and enduring.

“Industry plays an important role in supporting the federal government. There’s a partnership there that’s necessary.”

- Allen Hill

What role (if any) does the commercial sector play in helping agencies achieve their data strategies?

DS DAVID SPIRK

Industry is a critical partner in our data journey, providing the hardware, software, and expertise needed to implement our data strategy. The private sector also helps us drive change by showing the art of the possible for an organization that fully utilizes data to create new advantages and opportunities. Few data challenges are really unique to government, so it is a priority for me to have an ongoing dialogue with industry leaders about overcoming data challenges.

AH ALLEN HILL

Industry plays an important role in supporting the federal government. There’s a partnership there that’s necessary. The federal government should focus on what its core mission is and leverage industry to help where it fits to support those things to where we’re not experts in.

MG DR. MATTHEW GRAVISS

It’s a partnership. When you look at the response to the current pandemic, you’re seeing companies rise to the occasion and create data sets. They’re making it available and not worrying about how they’re going to benefit from it. They’re working on making it available to as many people as possible. We’re leveraging that data, combining it with our own to make decisions on behalf of a major department of the government. So I do think it’s a really important partnership that we have and it’s part of the reason we’ve been successful.

How is your agency managing data silos? How might cloud be a vital asset to leveraging data and satisfying the requirements of the Strategy?

DS DAVID SPIRK

Ultimately, I think we're moving past data silos. We really have spent a significant amount of time over the last six months building the community of interest around data. At this point in the journey, with the senior leaders in the department really emphasizing the importance, and the junior-most officers and enlisted demanding the capability, I think everybody kind of got the point. That we are in a moment of culture change and we need to think data sharing first. ... With the community of interest that we've built, and the big tent we've put up for the CDM Council, we now really know who all the data leaders are. If we're having a challenge with the data silos, or we're having a challenge with a legacy system that maybe doesn't comport to that open data standard architecture, ... we know exactly who to pick up the phone and call to be our leader to help remediate what some of those hiccups could be. I'm really not finding any data manager or data owner unwilling. There might be some technical hurdles or some legacy capability that we've got to work through or around, until we can just continue to get those modernized systems in place that do embrace the cloud. Enterprise cloud capability is just a powerful way to continue not having additional silos.

What are some challenges you foresee agencies facing in the future with the FDS? How can agencies mitigate or prevent these challenges?

DS DAVID SPIRK

We need to ensure major organizations have dedicated and empowered CDOs [Chief Data Officers]. They need to have a seat at the table as Department requirements are set and resources are prioritized. In the national security community, it's also vital that we appropriately prioritize data security with data transparency. It's clear our adversaries are looking for opportunities to harvest any sources lacking appropriate access controls. Data aggregation is a powerful tool that requires data leaders to be proactive in considering potential ethical, privacy, or security implications that could arise from future unintended use. Trust [in] managing data must be hard earned.

AH ALLEN HILL

Security is going to be the hardest challenge. It's going to require modernization of the infrastructure. Some agencies are more mature than others in their modernization efforts. It's going to take time. It's not going to be something that you just do in one year. ... So we get out of that life cycle. A lot of large agencies are in this lifecycle situation where they're trying to replace old components and they just can't stay up-to-date ... It comes down to a funding issue too. You have to invest the dollars. That's going to

“I think we have to keep our eye on the prize, which is really solving business problems across the board and making sure that governance and data management are enablers for solving problems.”

- Dr. Matthew Graviss

take some strategic initiatives of how you use your dollars and get there too ... I think that they have made drastic progress over the past couple of years, and they'll continue to make more.

MG DR. MATTHEW GRAVISS

I'd say that the top priority or the top challenge that we continue to face is balance. When I say balance, what I think a bout is the balance between mission-oriented initiatives and management-oriented initiatives, and to find more operational or organizational efficiencies, versus a more diplomatic mission. It's also a balance between supporting executive decision makers and doing projects that enable all analysts across the department to enhance their data acumen and solve more problems. I think we have to keep our eye on the prize, which is really solving business problems across the board and making sure that governance and data management are enablers for solving problems. At the end of the day, that's what we need to be focused on. Solving mission-oriented problems and management problems using data. That's really a cultural change that we need to work on.

Of the 20 action items on the Federal Data Strategy, which is the most pressing to accomplish for your agency? Which is the most difficult to accomplish?

MG DR. MATTHEW GRAVISS

I'd say it has to do with data skills... It's really about multiple communities within the State Department. One is the analyst community, people who have signed up to do analytics and how we get them the tools and the training to do their jobs. Another group

“Protecting data integrity ... knowing the public can trust that what we have of public data is truly protected. To me, protecting is the most important aspect of responsibilities that fall in my preview.”

- Allen Hill

is folks that aren't necessarily hired as the analyst, but analysis is becoming part of everybody's job ... The other group to really focus on is the executive leadership. We want them to get to the point where they're saying, "what does the data say," before they make a decision.

How important is inter- and intra-agency data sharing for the work of your organization? What are some steps your agency has taken to achieve improved data collaboration?

MG DR. MATTHEW GRAVISS

Data sharing is everything. The department's Center for Analytics is meant to be an Enterprise Analytics Hub, and there's a couple of ways we do that. One is working to enable those capabilities across the department like training. Another is to do these through more complex analytics, which are through partnerships with other bureaus ... The Chief Data Officers Council is a really phenomenal working group that they set up for federal CDOs to talk about common issues, and data access is certainly one of them. There's a lot of work going on within the council to do that. There are also a lot of common challenges that we're all going through as Chief Data Officers of various agencies, so the ability to collaborate with that group has been really great.

How do you imagine most agencies are progressing with the Federal Data Strategy (now versus six months ago?)

DS DAVID SPIRK

From everything I have seen, there is broad support for the Federal Data Strategy and its implementation. Different agencies are at different places in their journey, and have different resource

constraints, but that's the benefit of a venue like the Federal CDO Council: it provides a place to collaborate on problems, share insights, and drive accountability. That leadership has played a key role in moving the Strategy forward as have the partnerships that Ted Kauok (CDO - USDA) and Kshemendra Paul (Executive Director, Data Governance and Analytics, Chief Data Officer) have played in their leadership roles.

Could you name a few technologies/practices that have helped assist your agency in making progress with organizing and using data?

DS DAVID SPIRK

We move the fastest when we leverage current commercial practices, including a data architecture based on open data standards. The data analytic platforms we use depend on a modern, secure IT infrastructure that can push information out to the tactical edge. That is why data is embedded in DoD's Digital Modernization program that includes investments in cloud, DevSecOps, Artificial intelligence, Cybersecurity, and modernized C3. These investments are vital to enabling us to collect, understand, and seamlessly share the right data at the right time. It also needs to be understood that data management in an organization with the breadth of tasks such as the Department of Defense

is a team sport. Spending time cultivating the relationships between CDOs, connecting the Community of Interest to each other, and finding opportunities to benefit from lessons learned and borrow or scale successes is how we are moving fast and why we expect to accelerate in 2021.

What can agencies do to benefit from the Strategy in the long run?

DS DAVID SPIRK

Lead from the front — showing that data will be used for all important decisions and we aren't waiting for it to be perfect. Senior most-leader emphasis is also a critical component. In DoD, [Acting Defense] Secretary Norquist has set the tone for the entire Department by using enterprise data analytics in all his senior decision meetings. As with any bold change in organizational culture, it will take courage and perseverance, but we owe our best to the public and military members we serve.

INDUSTRY PERSPECTIVE

Nicholas Speece | Chief Federal Technologist at Snowflake



Do you think the federal government needed to modernize data management/analytics? If so, what value do you think the Federal Data Strategy brings for the government?

So first absolutely yes, but when we say that we have to modernize, but what does that mean? How do we get there? What services do we need to provide to our citizens? That's where the federal data strategy really starts to cut away some of the ice on top of this massive glacier of government systems. We've got to start figuring out what we're going to build and what we want to accomplish with whatever improvements or modernization we have. I think that's what the federal data strategy does. It gives us not the answer but an answer. It allows us to be iterative and start building improvements and focusing on technologies that can answer questions by giving us a framework to do so.

How important is inter- and intra-agency data sharing for the federal government?

It matters more than anything else you can do with data. When we started looking at COVID-19 and how the pandemic was going to affect the country, we didn't have an infrastructure to share data. We had SFTP, FTP, and Excel spreadsheets going through email. Some of it is still done that way, but what I think the pandemic has taught us, at least from the perspective of a systems engineer, is that data collaboration is much more important than almost any other aspect of data management.

How has COVID-19 changed the progression of the Federal Data Strategy? Is the FDS still on track?

Certain aspects have been massively accelerated. Data sharing and collaboration have come to the forefront. The data strategy for a long time was about looking inward and finding what data we had, building data catalogs, and identifying the gaps in our own infrastructure in each of the agencies. I think the focus became outward very quickly and in those veins I will say that the strategy and the CDOs upholding that strategy have responded extremely well to a very fluid situation.

What are some challenges you foresee agencies facing in the future with the Strategy? How can agencies mitigate/prevent these challenges?

I think a lot of agencies out there are trying to do data shared in collaboration, but they're doing it with legacy tools and that's really challenging. Snowflake is as successful as it is because we've brought the data cloud to the world. We've embraced data sharing, private marketplaces, and data exchanges. It's like an app store for your phone, but it's datasets. Getting somebody else's data set in a seamless way with no movement and no copying or FTP in a securely managed way, that adds more value to those kinds of challenges than anything else we've done in the legacy data world in the past. So adopting strong platforms like Snowflake is how we get there.

That's how we really move the needle. It's not the executives that will get things done in the end. It really comes down to everyday people in their work centers making this stuff happen.

- Nicholas Speece

What can agencies do to benefit from the Federal Data Strategy in the long run?

We have to embrace it from the top down. Every time I talk to an audience, I talk about culture because it's so evidently important to everything that we do. If the folks who were in these agencies aren't bought in, and if they don't know the value of what you're giving them, then they're never going to take advantage of it. They have to understand its value to their job, to their role, or to their lives by having access to this data. They'll have to embrace those changes to the data strategy and what it really means to their agencies. That's how we really move the needle. It's not the executives that will get things done in the end. It really comes down to everyday people in their work centers making this stuff happen.

Can you tell me about Snowflake's solutions and how they can provide agencies looking for technology modernization?

Snowflake is a brand new look at what a data platform could be. We've had databases for decades and they worked very well in the architectures that existed decades ago, but we're not there anymore. We're in a very modern architecture, a cloud forward architecture. I remember when big data became the favorite buzzword, we were talking about gigabytes of data, and now we're talking about petabytes or zetabytes of data being generated and handled and analyzed. It's really easy to say, "I put 10,000 spinning disks in my data center, I can store all my data," but the question is: can you run compute against it? Does your data do anything for you? How can it scale and be cost effective and be useful? Snowflake answers these questions when our cloud computing platform works with partners like AWS, Microsoft, Azure, GCP, and we don't break the bank. Data sharing, data collaboration, and artificial intelligence can all be done on Snowflake without breaking the bank and in a very easy to start way.

ENDNOTES:

1. <https://strategy.data.gov/assets/docs/2020-federal-data-strategy-framework.pdf>
2. <https://strategy.data.gov/background/>
3. <https://strategy.data.gov/overview/>

ABOUT

**Government
Business
Council**

As Government Executive Media Group's research division, Government Business Council (GBC) is dedicated to advancing the business of government through analysis, insight, and analytical independence. An extension of Government Executive's 50 years of exemplary editorial standards and commitment to the highest ethical values, GBC studies influential decision makers from across government to produce intelligence-based research and analysis.



Snowflake's cloud data platform shatters the barriers that have prevented organizations of all sizes from unleashing the true value from their data. Thousands of customers deploy Snowflake to advance their businesses beyond what was once possible by deriving all the insights from all their data by all their business users. Snowflake equips organizations with a single, integrated platform that offers the only data warehouse built for the cloud; instant, secure, and governed access to their entire network of data; and a core architecture to enable many types of data workloads, including a single platform for developing modern data applications. Snowflake: Data without limits. Find out more at [Snowflake.com](https://www.snowflake.com).