Q&A

Executive Viewpoint A Conversation with SUZETTE KENT

SUZETTE KENT
U.S. CIO Suzette Kent

The federal CIO discusses how the government's new data strategy informs cybersecurity, emerging technology and citizen services

The federal data strategy is an ambitious, comprehensive approach to managing, using and protecting data. Why is it so important to tackle those challenges now?

We're really excited about the data strategy, and there are a couple reasons "why now?" and "what's in it?" As part of our big focus on technology modernization, we're looking at legacy applications and transforming them and, in many cases, extracting embedded databases. So it's the right time to ask questions such as: What do we want to do with the data? How do we make it more useful? How do we make it more accessible? From a work effort standpoint, modernization and addressing how we plan to use data in the future go together.

A second reason is we're very focused on protecting citizen data, so cybersecurity and privacy are very important. It's essential that we do the right thing to protect citizen data. Some of those concerns are addressed in the Year-1 Action Plan through creating data inventory tools, increasing infrastructure maturity and building workforce skills.

The last big reason, and probably the one I am most excited about, is that data is the foundation for our future-state technology, including much of the automated technologies, because of the way our data-driven society works and the fact that we're collecting more data everywhere. We have the opportunity to use new technologies to leverage data and information in a way that we never had before — to make citizen services better and improve the quality of the policy and program decisions we make.

We have a very direct mandate to protect citizen privacy but also to use data to stimulate the economy and do great work serving missions. A strong data foundation is how we best position ourselves for the future.

What are the biggest potential obstacles to implementing the strategy, and how can agencies overcome them?

We've had a great set of exchanges around the data strategy with academia, private sector and agencies. It's one of the areas where we've had some of the broadest public engagement. One topic that is top of mind, and probably not shocking, is how we build the skills inside federal agencies. In the academic world, some of the data science disciplines are newer and some of the advanced technologies are newer, and those are skills that we want to bring into the government and build up at agencies where people already understand the mission.

Another challenge has to do with priorities. It's a simple business decision. There are a lot of priorities at every agency, so how do we ensure that the data strategy makes it to the top of the list? The way we drive priority is to use data to solve important mission-serving questions. We align the data work around something that's delivering tangible results, whether it's for citizen services or other mission-serving purposes. The way we get enthusiastic people working in this space is by using the data we have to solve tough questions that are closely linked to mission.

What impact will the strategy have on the government's cybersecurity efforts?

Cybersecurity is an area where we've been leveraging data as a strategic asset for some time, whether it's looking at pattern identification or sharing threat information. The industry and federal agencies have been making investments in data and advanced technology for many years. Now we can expand the focus beyond cybersecurity to everything that we do with data — not just examine the protection of data, but how it is used, how we protect individual privacy, how we look for bias in data and how we decide in which situations to use multiple years' worth of data. The strategy gives us the opportunity to expand the set of disciplines we already have in the cybersecurity arena.

We're also looking at using more commercially available software and broader technologies. That gives us the opportunity to use more common data nomenclature, which helps us improve the speed and quality of the information that we can share.

How does the data strategy mesh with governmentwide initiatives to embrace emerging technology, such as artificial intelligence?

That's one of my favorite questions to answer, and I've used this analogy before with people who ask, "How do these two things go together?" They are critically linked. Think about artificial intelligence or automation as an engine of some type. If you don't have the data, which is the fuel, your engine's not going to work very well. It's like having the world's fastest plane but no fuel.

Responsible, ethical use of robotic process automation, natural language processing, image recognition or machine learning relies on high-quality and usually very large volumes of data.

The really interesting questions are solved by analyzing relationships among multiple datasets. We're trying to identify preventative steps in disease management — that requires large volumes of data. To make progress on autonomous vehicles, we need weather data, road data, vehicle data and lots of sensor data.

The power that those technologies can bring is directly related to the quality of information that you can put side by side with them.

What steps can agencies take now to begin moving toward the strategy's objectives?

There are a lot of activities in the Year-1 Action Plan that agencies should have already been taking steps to achieve, and many are doing so. Having a data protection toolkit, an ethics framework for how to leverage data, an understanding of where resources and tools are — those are some things that many agencies were already doing. We're just putting more structure around those activities.

Improving our federal inventory, adding standard data catalogs and being more proactive across government to identify opportunities to increase staff skills are some other things agencies can and should be doing.

They're also asked to look at their own level of maturity for how they manage and use data so they can identify places they need to make investments.

The most powerful of the agency priority items is understanding what questions they want to answer. Why are you making an investment in using data, and what part of the agency mission does it drive? If there's any place they can start, it's right there because that's how you tie the mission, the resource dedication, the people investments and all those things back into answering the question: What are we doing for mission, for citizen services or as stewards of taxpayer money?

We update strategy.data.gov continually, and under the Proof Points tab, we share agency success stories — whether they're operational, outcome-based or skills-building — to show different actions toward delivery of the strategy.

Are there other important elements of the strategy that you'd like to highlight?

When I look at the data strategy, most of my comments are focused on what we do inside the federal government. My responsibility from both a policy and oversight perspective involves working with agencies to improve the way we use federal data to provide the services of government.

But another part of what the agencies are doing, and it is embedded in the strategy as well, is how we might make datasets available publicly that drive other types of activities. This year we prioritized using information to drive AI research and development in the American private sector. We also have priorities around sharing geospatial data, and we're prioritizing financial data inside the government both for transparency and economic benefit.

The discipline absolutely supports what we're doing inside federal agencies, but there's also benefit to the private sector, the states and others who are doing research in some very specific areas where we're putting additional investment and resources. ■