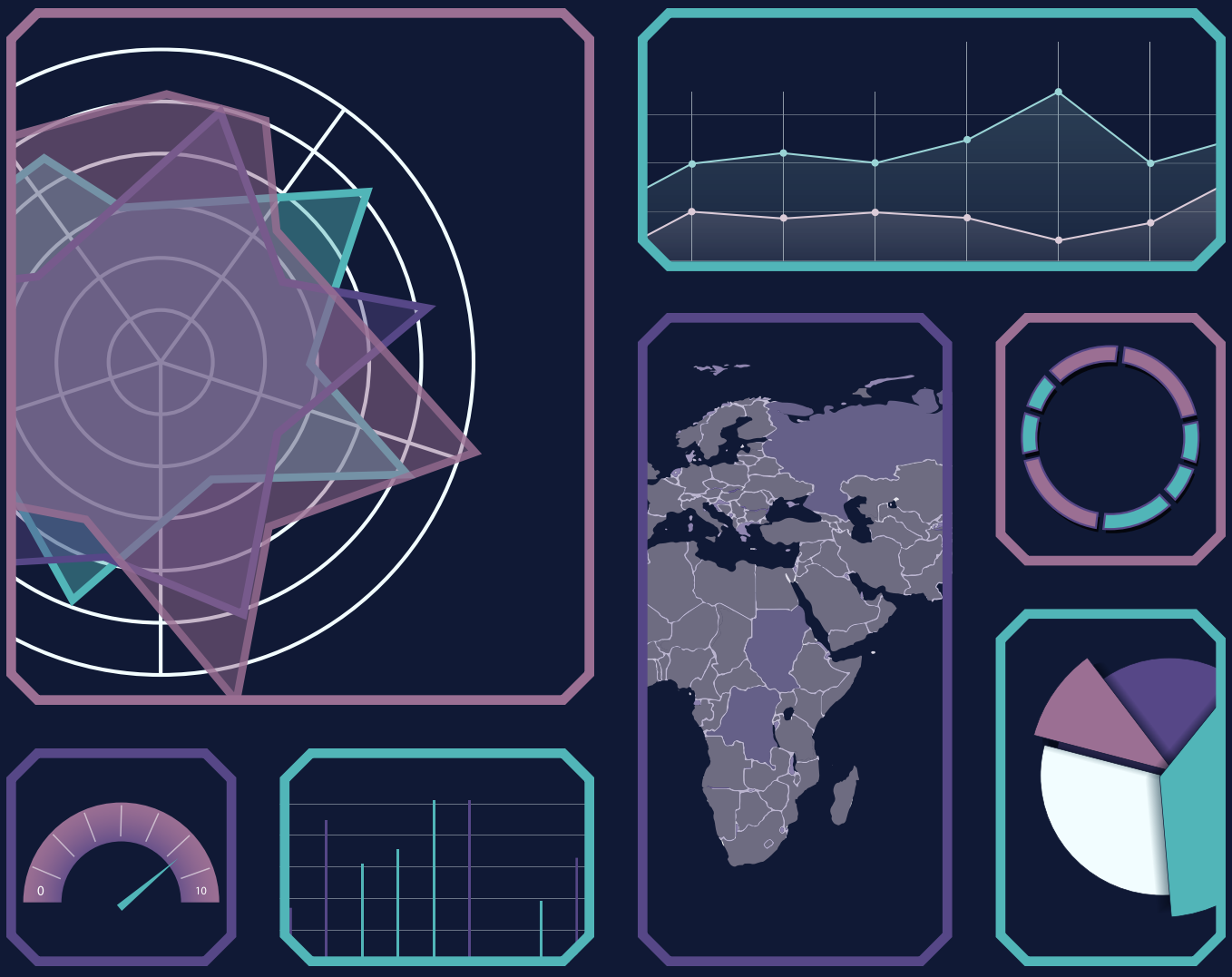
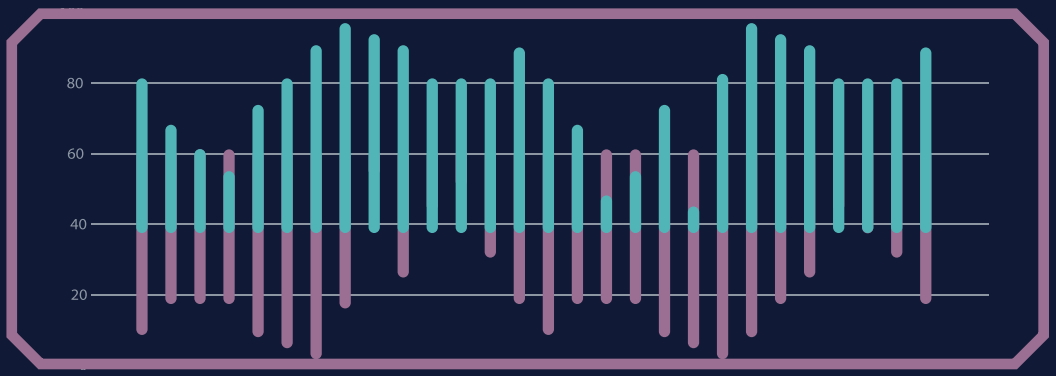


Advancing the Art of Data Analytics

carahsoft

govloop



Carahsoft and GovLoop have partnered to provide resources around the latest federal, state and local tech trends. The goal is to better inform government leaders' 2021 planning and enable their data strategies with the latest technology.

Executive Summary

Agencies understand the mandate: They need to leverage data as a strategic asset.

Within days of President Joe Biden taking office, his administration emphasized the centrality of data in addressing both the COVID-19 pandemic and racial inequity.

Before that, laws and initiatives such as the Foundations for Evidence-Based Policymaking Act of 2018, Digital Accountability and Transparency Act of 2014, and Federal Data Strategy pushed agencies to build their capacity for data-driven operations.

This overall goal has led many agencies to create and appoint data-focused roles, including senior-level chief data officers (CDOs) and first-ever data analysts. (See Page 12.) They have begun advancing data strategies and investing in solutions. But through these mandates and governmentwide initiatives, **how can agencies bring their goals and strategies to life to address their most pressing problems?**

GovLoop and Carahsoft created this guide to help.

- First, we lay out **four basic competency areas** that every organization — federal, state and local — should consider to aptly use data.
- To give you a taste of how agencies are solving real-life issues, we offer **four key use cases** around how government is applying analytics to challenges such as COVID-19 vaccine distribution, data sharing, remote work and autonomous vehicles.
- Additionally, **hear from trailblazers** at the National Library Service for the Blind and Print Disabled and the Census Academy, where data is either being strategically used for the first time or disseminated in new ways.
- Throughout the guide, we offer **best practices and checklists** on how to apply core competencies, from data governance and data skills to data ethics, at your organization.

If you're looking to use analytics for problem-solving beyond checkboxes and requirements, this guide is for you.

– Pearl Kim
GovLoop Staff Writer

Quick Primer: What Agencies Need for Success in Data Projects

DATA GOVERNANCE

Data governance is an accountability framework of roles, responsibilities and processes for data assets. It specifies who has authority over certain data and how that data can be used. "Data governance is the process of setting and enforcing priorities for managing and using data as a strategic asset," according to the Federal Data Strategy's Data Governance Playbook.

The value: It enables agencies to organize their data more consistently so it can be used actionably.

"We've got a lot of pockets across the organization.... They all do data management. They all do data governance. They've always been doing it. It just hasn't been consistent from organization to organization. That has been the biggest thing to start to work toward over the past year with the agency standing up."

– **Preston Werntz**, CDO, Cybersecurity and Infrastructure Security Agency (CISA), in a March 2020 GovLoop interview

Check out best practices for governance on page X.

DATA SECURITY

Data security entails protecting sensitive data assets. They include methods such as encryption, which makes data difficult to identify by unauthorized parties, or resilience, which backs up copies of data so organizations can recover it in case of erasure or corruption.

The value: It ensures security at the granular level, particularly when all kinds of information can be stored in many locations: public cloud, private cloud, data center edge locations and more.

"What is it that we're trying to protect? At the end of the day, I say we're trying to protect data. A lot of people come from it from the identity [management] aspect with their solutions, which is perfectly fine.... But it's not always a human being who's accessing data. Data is sitting there at rest when it's not being used. Data is shared system to system. So, I still need to do data protection."

– **Gerald Caron**, Director of Enterprise Network Management, State Department, in an October 2020 GovLoop online training

DATA LITERACY

Data literacy is the skill needed to understand and communicate data. Gartner defines it as "the ability to read, write and communicate data in context." This includes understanding data constructs and analytical methods that are used, in addition to the ability to describe the application and outcome of data.

The value: It is a business enabler. As data and analytics become — and in many cases, already are — core organizational assets, employees must have the basic ability to communicate and understand conversations about data to leverage these assets for mission delivery.

"I think we're getting to a point where, as we become more and more digital overall, there's a need to have some form of basic data literacy training for government employees. I would say that at all levels of government, even if it's a one-hour [training] where you give employees the basics, that can go a long way, in the same way that CIOs are implementing cybersecurity training for all their state employees."

– Tyler Kleykamp, Director, State CDO Network, in a February 2020 GovLoop interview

DATA ETHICS

Data ethics are standards of behavior that keep data acquisition, management and use accountable for these key goals: protecting civil liberties, maximizing the public good, and minimizing risk to individuals and society, according to the federal Data Ethics Framework.

The value: Applying data ethics across data usage can improve public trust. It can also mitigate future risks. One way is by ensuring the effective use of time and resources by reducing the long-term costs associated with ineffective, unreliable data services.

"Remaining a leader in data ethics requires individuals, agencies and cross-agency communities to acknowledge that legal compliance does not guarantee ethical behavior. Therefore, federal leaders and data users must embrace a culture of ongoing discussion, engagement and learning."

– Data Ethics Framework

Check out our worksheet on ethics on page 25.

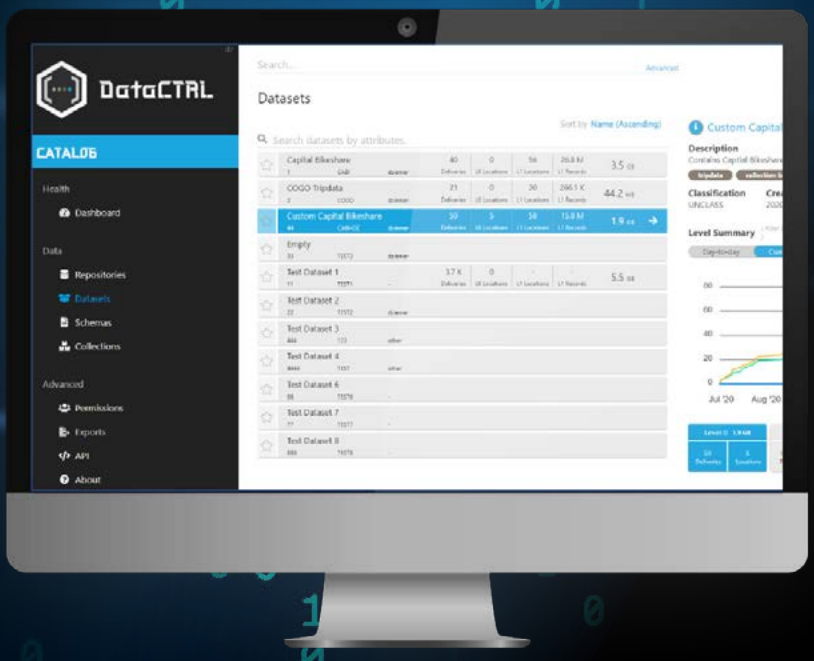
Your **Dark Data** Could Be Hiding Something.

DataCTRL®

Did you know that your organization's data can go dark by not having enough analytic insights for decision-making?

DataCTRL can uncover the data that lurks in the shadows.

Bring all your data into view with scalable workflows of your largest repositories of data – and track that data from input to delivery.



NS2®

SAP NATIONAL SECURITY SERVICES

Bring your data into the light.

sapns2.com/datactrl

Before You Start Data Governance...

An interview with Michael Towles, Principal Business Processes Consultant, SAP NS2

One of the ultimate goals for establishing data governance is to revamp an organization's data culture for the better. And this is an immense task.

That's why before an organization begins strategizing and implementing data governance, there are a few things it can put in order to create a more effective and valuable governance strategy.

"Data culture is valuable, not just to the owner, such as the agency, but to the citizens of this country as well," said Michael Towles, Principal Business Processes Consultant at SAP NS2, a cloud and analytics provider. "We rely on the government for a lot of services, and the data it brings in is really the citizens' data. It's the business of the country to put our best foot forward."

Towles shared three action items agencies can do to prep their data governance strategies.

1. KNOW WHAT DATA YOU HAVE

In discussions with customers, Towles has found that organizations often lack insight into what data they have under their control.

For example, some agencies have an immense amount of dark data that goes back decades. Dark data is the kind of data that doesn't have enough context for agencies to use for analytic insights or decision-making.

You need to know such data exists before you can transform your data culture. Use tools that can help discover and catalog all the data that is under your control. It provides the foundation for building a data strategy.

2. STRATEGICALLY PLAN

When agencies understand the data they have, then they can plan how to turn it into valuable assets for their strategic goals.

For example, data on agricultural crop yields are rich sources of information in and of themselves. But when an agency can analyze them against weather patterns, population densities, even normalized economic models, it can unlock the value of the data in more profound and strategic ways. The agency can begin to predict the future economic impacts of these crops.

"The ability to build new and innovative data models provides value that you otherwise would not have seen," Towles said.

3. BUILD PROCESSES AND PROCEDURES

Once agencies have a strategic plan in place, they can begin to set the regular rhythms of using the data. The processes and procedures should be a well-documented plan that deals with day-to-day technology needs.

Externally, it's important to establish a communications plan that standardizes interactions with data users throughout the data lifecycle.

Internally, it's critical to create procedures that allow data management platform operators to manage services in a reliable fashion, whether that's DevOps or backup and recovery.

Ultimately, these processes should all indicate a measure of trustworthiness to data stakeholders. SAP NS2 can provide a data platform that agencies can rely on, in turn, Towles said. "SAP has always been a company focused on data, and NS2 has always been focused on government."

Not every agency will have the same plans and procedures. But once these steps are prioritized, then refined data governance can really begin.

Best Practices in Becoming a Data-Driven Agency

DO YOUR CO-WORKERS BUY INTO DATA GOVERNANCE?

A key to becoming a data-driven agency is having a sound data governance strategy. But a sound strategy means something only if your co-workers and leaders know it exists.

The Stanford Data Governance Maturity Measurement Tool looks at awareness of data governance in five stages, each moving toward greater support and sustainability.

Which stage best describes your organization?

1	Limited awareness of the purpose or value of the data governance program.
2	Executives are aware of the existence of the program. There is little knowledge of the program outside upper management.
3	Executives understand how data governance benefits/impacts their portion of the organization, and knowledge workers are aware of the program. Executives actively promote data governance within their groups.
4	Executives understand the long-term data governance strategy and their part in it. Knowledge workers understand how data governance impacts/benefits their portion of the organization. Executives actively promote data governance beyond the immediate group.
5	Both executives and knowledge workers understand their role in the long-term evolution of data governance. Knowledge workers actively promote data governance.

KEY QUESTIONS ABOUT DATA GOVERNANCE

As part of a December 2020 report, the Government Accountability Office (GAO) discussed the importance of making data governance part of an agency's culture. "Such a culture change can result in a shared understanding of the importance of using data as a strategic asset to achieve the mission and improve operations," the report states.

With that in mind, GAO suggested that an organization's CDO or other data leader consider three key questions:

1. How can the CDO ensure that data governance strategies answer the agency's priority mission questions?

CDOs should identify and leverage existing data assets or create new data governance strategies that can help address their respective agencies' missions, top priorities and most important challenges.

2. What are the CDO's key objectives with regard to data governance?

Once CDOs understand their agency's mission, priorities and key data needs, they should define their objectives and key responsibilities to address those needs.

3. How can the CDO communicate the value of data governance?

In addition to identifying priority datasets and the CDO's primary responsibilities and objectives, CDOs need to provide a clear vision and promote the value of data governance agencywide.

THE ROI ON ANALYTICS (LOCAL GOVERNMENT EDITION)

For local governments looking to build an analytics team, the first step is to get funding. That can be a challenge when budgets are tight, so it's more important than ever to talk about analytics in terms of the return on investment.

In a report last year, the Ash Center for Democratic Governance and Innovation at the Harvard Kennedy School highlighted research by Jane Wiseman, Chief Executive Officer of the Institute for Excellence in Government, who found that analytics can deliver the following benefits for local government:

- **Fraud detection cost savings**
- **Efficiency improvements that reduce costs**
- **Accuracy improvements that reduce costs**
- **Efficiency improvements that improve outcomes**
- **Increased revenue capture**
- **Operational changes that increase safety**
- **Increased faith in government due to more transparency**

mongoDB®

Voted the most wanted database by developers for 4-years straight*

The document data model has become the most popular alternative to relational databases.

Learn how a document data model can help developers innovate faster, and uncover use cases where you can put document databases to work.

Learn More

*Stack Overflow, 2020 Developer Survey

How to Develop Strategic Value From Your Database

An interview with Brent Leech, Regional Vice President, Public Sector, MongoDB

In any IT initiative, it's easy to use what has always worked. And this is a reasonable perspective to take in many ways, but it's not always the most judicious.

The best talent wants to work at organizations that use flexible, open source technology. And the best organizations want technology that is secure and low-cost. Older solutions, such as relational database management systems (RDMBS), can't offer all these things. Organizations need to find more modern, effective, attractive and low-cost solutions that can take them into the future.

"It's not just moving away from RDBMS just to move away. You have to move away thoughtfully with purpose, toward a technology that is going to serve you well for years," said Brent Leech, Regional Vice President of Public Sector at MongoDB, an enterprise database solutions provider.

Leech shared three best practices agencies can take to thoughtfully approach their data and applications.

1. DON'T START NEW INITIATIVES WITH OLD TECHNOLOGY

Even if they would like to, agencies can't replatform every legacy application they have. That's why a strategic way to prioritize development initiatives is to start new efforts with the best technology.

Moving toward a modern infrastructure out of the gate will save time and effort in the long run, Leech said, particularly if it works the way people work today.

For instance, relational databases are traditional storage tools that have been around for 40 to 45 years. People use data differently now, and tools like MongoDB's document database solution keep that in mind. It functions as an operational layer over huge data repositories to help agencies extract the most applicable data for their mission, and securely through enterprise tooling.

2. USE FLEXIBLE SOFTWARE METHODOLOGIES AND TOOLS

Software developers cannot predict the future. They don't know what data types will be required over time, or how their database schemas will need to look like as applications evolve.

So, organizations need to provide their employees with a database technology that is adaptable – with flexible schema and data types – allowing them the freedom to change things quickly.

"We store our data in ways that developers, data analysts and data scientists think about data," Leech said. "If you use flexible data technology, your time to mission and reliability will improve, and you can future-proof a lot of what you're developing."

3. SEE LOW-COST FAILURE AS A TOOL

If agencies keep in mind that failures are inevitable, they want to be able to pivot and learn from them quickly.

Investing in large, monolithic technologies that are expensive to procure and maintain isn't the best way to fail fast and fail cheap.

To turn failures into opportunities for growth, agencies need low-cost solutions to keep initiatives light, quick and simple. You don't want unplanned events to hit your budget or timeline in drastic ways.

With MongoDB, you can save up to 90% over traditional database deployments with zero down time, even during maintenance windows.

"Don't take your database technology for granted," Leech said. "Make sure you understand there are alternatives out there."

How Data Analytics Complements Institutional Knowledge

At the National Library Service for the Blind and Print Disabled (NLS)

An interview with David Spett, *Data and Analytics Officer, National Library Service for the Blind and Print Disabled, Library of Congress*

The National Library Service for the Blind and Print Disabled (NLS) is undergoing perhaps the largest modernization effort since its founding in 1931.

Strategic data use is central to the effort. And in 2019, NLS onboarded its first data and analytics officer, David Spett, to help.

NLS has served patrons with low vision, blindness and other disabilities for nine decades. Although data can certainly bring transformation, Spett emphasized that it serves as a "great complement to institutional memory and knowledge."

"Ideally, data will inform our decisions and operations, but it does not fundamentally change the services we provide and the way government operates," Spett said.

In March 2021, GovLoop interviewed Spett to get a peek into his experience helping NLS become more data-informed and mission-effective.

The interview has been lightly edited for clarity and brevity.

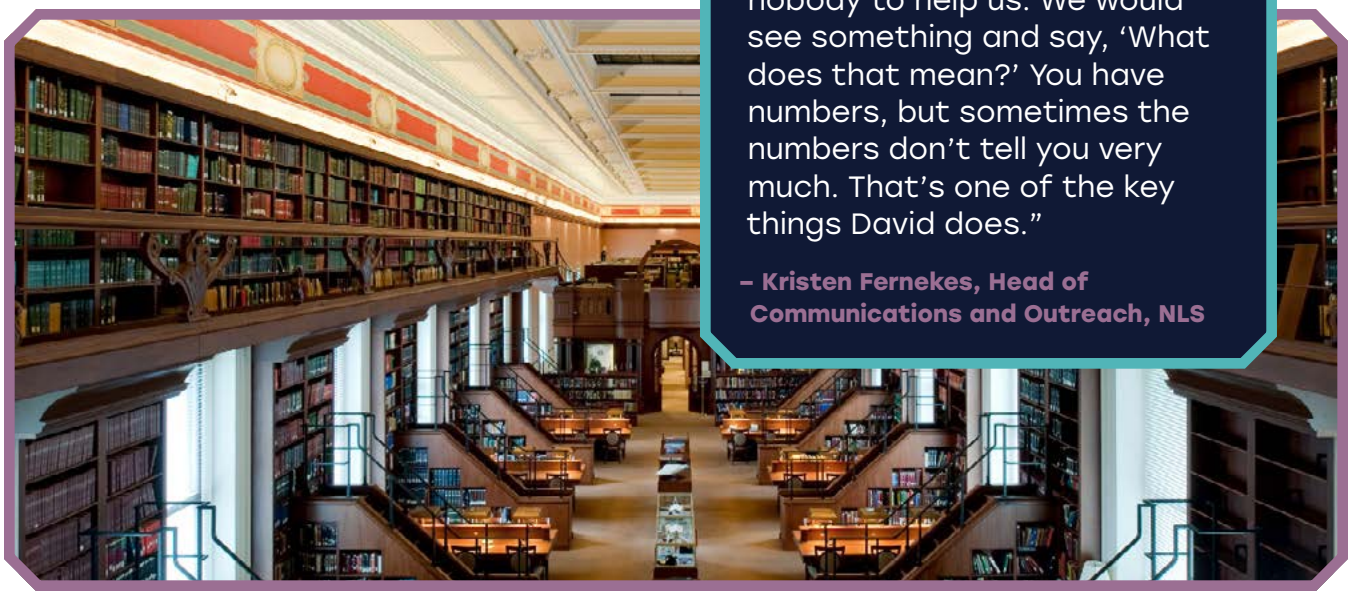
GOVLOOP: So you are in a newly created role. How were you first brought in as the data and analytics officer?

SPETT: We are 90 years old, and we have been operating with a unique model since our founding. We provide library materials to Americans and individuals with disabilities. And we generally ship those materials around the country — print braille and audio cartridges via mail.

Now, we're going through an organizationwide modernization that is probably the biggest change in our operating structure since we've been founded. And we are shifting from a predominantly U.S. mail-based delivery mechanism to a predominantly online streaming-based mechanism.

The idea with creating my position was to ensure we had the right data organizationwide to serve our patrons as best as possible, to make sure we are providing them with everything they need throughout the course of this modernization.

> "Before we brought David on, if we wanted to delve into the data we had, there was nobody to help us. We would see something and say, 'What does that mean?' You have numbers, but sometimes the numbers don't tell you very much. That's one of the key things David does."
>
> **– Kristen Fernekes, Head of Communications and Outreach, NLS**



Data touches so much, and you are just one person. So what does your typical workday look like?

I completely agree that data touches so much. And you could argue it touches almost everything we do on a day-to-day basis.

Something I've discovered is that there's almost an unlimited amount of potential work I could do at any given time. There is always more data to analyze and study. And so my day is really about prioritization, because I obviously cannot do everything. When I make decisions about prioritization, I gather input. I look at where I can have the most impact to help drive the most important decisions. I look at what the biggest problems that the organization needs to solve are and, in particular, what our strategic plan calls for.

What impact has data analysis had at NLS so far?

We are lucky to have a good number of staff who have been with the organization for a long time. They have a lot of institutional memory, and they really understand the practices and traditions of our organization and our service model. They also have some preconceived notions about the way things are, and why they are that way.

Sometimes what we find when we analyze rigorously gathered data is that those preconceived notions may be correct, but they also may not be — and the data might suggest that we go in a different direction.

Data is really instructive at serving as a great complement to the institutional memory and knowledge that I mentioned, as we go about making decisions with respect to the modernization.

What has the data shown so far? Has it challenged some preconceived notions?

One of the questions we're working through now is how much and what kind of assistance our patrons will need to adapt to a streaming model. Will they need hardware, software and training, or just a subset of these different items?

Data has suggested to us that, actually, a greater proportion of patrons than we thought already have a smartphone or some Wi-Fi. This tells us that patrons are likely to need a little less of a leg up — that we can simply provide them with an app or materials to stream. The data we gathered enabled us to understand that we may not need to provide some of the hardware we thought we needed to provide in the past.

How do you try to marry data analytics with institutional knowledge, as you mentioned, in your work?

I find that working collaboratively is really helpful. My work is improved when I have the opportunity to hear from and obtain input from my colleagues, particularly those with the organizational knowledge and history of [NLS]. That can both inform my analysis and help me better understand what data-gathering processes may be ideal and what may not be ideal.

Now that you've been in your role for about two years, what are some effective ways you've learned to communicate using data?

I find it's really helpful to share information in multiple ways, often through narrative and visualization. I like tables, graphs and charts, but I also like to provide narrative explanations of those visualizations. We have a number of blind staff, including NLS' director, so it's really important that whenever I provide a visualization, there's also a narrative.

It's important as well not to just share the data itself, but to take it one or two steps forward and explain what it means. What are the implications for our current decision-making and our organizationwide strategic goals?

What do you want employees or leaders who don't have explicit data-related roles to know about data?

What I would say to employees who don't have data-related roles is that, ideally, data will inform our decisions and operations, but it does not fundamentally change the services that we provide and the way government operates. I would look at data and data-informed decision-making as an evolutionary process, not a revolutionary process. It's not something to be afraid of, but something that can help us all be better at our jobs.

alteryx

Make Breakthroughs That Matter

Drive mission-critical decisions with analytics automation

Free Public Sector Starter Kit

How Automated Analytics Can Fuel Digital Transformation

An interview with Andy MacIsaac, Solution Marketing Director, Alteryx

Deriving data insights through analytics is no longer just nice to have – it is critical for the operation and improvement of government, especially in times of crisis.

But there are many challenges to deriving insights, whether that's having the data available to analyze or having the right skills to analyze the data.

Automation can help. For years, agencies have strategically adopted digital technologies to transform their business processes and services. But according to a 2018 McKinsey report, only a third of initiatives have succeeded.

Those who succeeded were able to flourish because they converged these three pillars for digital transformation: data democratization, process automation and people upskilling.

Analytic process automation (APA) meets all three needs at once by addressing the whole analytics lifecycle through one unified platform.

"We're helping states rebuild critical unemployment systems with analytic automation, enabling agencies to deliver on their mission and get critical support to citizens," said Andy MacIsaac, Solutions Marketing Director at Alteryx, an analytics solutions provider.

DEMOCRATIZING DATA

Essentially, to democratize data is to make it available for analysis. But this can be challenging, considering the sheer amount and variety of data types and sources.

APA can automate the huge amounts of data that need to be ingested, and a data-agnostic platform can enable organizations to input various data types, everything from traditional local files and applications to information collected on bots and sensors.

APA platforms allow the greatest amount of data to be leveraged and streamlines the process to prepare it for analysis.

AUTOMATE PROCESSES

Second, APA focuses on delivering high-value insights by streamlining low-value tasks.

From data prep and blending all the way to predictive and prescriptive analytics, APA automates and unifies analytic steps by providing users with over 260 automation building blocks that cover the full spectrum of the analytics lifecycle.

"Mundane tasks like compiling data and joining spreadsheets is not high-value work," MacIsaac said. "It's important work – it has to happen – but if you can automate those processes, then your data team and analysts can focus on what they have a passion for, and that's uncovering insights."

UPSKILL EMPLOYEES

With automation freeing up employees' time, they can focus on building their skill sets as well.

Data skills are in high demand but low supply. So, how can you grow these skills through the creative talent you already have within your organization? This is the spark for cultural change to build a digital-ready workforce.

A code-free or code-friendly platform with guided models and community forums can invite even those without technical data skills to engage in analytics and learn as they go.

"What we enable is our data workers to unleash their creativity and problem-solving skills. And when that happens, organizations are going to reap the benefits of their data and analytics capabilities," MacIsaac said. "That's the benefit of APA through the Alteryx platform."

Trending Use Cases for Data & Analytics

The ways to use data are almost endless. Any organization, any team, even any individual can use data to help solve a multitude of problems.

Here are some examples of ways government is using data and analytics to address the most pressing issues today.

RACIAL EQUITY

North Carolina Rolls out Vaccines

Race and ethnicity data collection plays a critical role for government to equitably distribute COVID-19 vaccines. Governments and health care providers can carry out racially equitable distribution only when they know who is receiving vaccines by race and ethnicity.

In January 2021, only 17 states were publicly reporting vaccination data by race and ethnicity, according to KFF, a health policy analysis organization. The number has grown to 41 as of March 2021, which still means not all states are reporting or collecting this information.

North Carolina is one of the 17 states that started tracking race and ethnicity data early on. As of March 2021, it reported having less than 1% of its vaccination data missing this information.

How? The state made race and ethnicity a required field **for vaccine providers to complete.**

- Providers already needed to use the state's management system to acquire vaccines. Adding a mandatory data field to complete registration helped the state collect complete data on people who received shots to guide equitable vaccine distribution.
- Bloomberg reported that the state has one of the smallest differences in vaccination rates between Black and white residents, though it is struggling to vaccinate Hispanic residents at the same level.
- **The bottom line:** It wouldn't have known this without the data.

"We are...holding ourselves and vaccine providers accountable for ensuring that underserved and marginalized communities have access to vaccines," said Mandy Cohen, Secretary of the North Carolina Department of Health and Human Services.

PUBLIC HEALTH

CDC Modernizes Mortality Reporting

COVID-19 was the third leading cause of death in 2020, a report from the Centers for Disease Control and Prevention (CDC) found in March 2021.

- The gruesome statistic is the kind of critical information that public health decision-makers need. And CDC was able to acquire this information partly through "improvements in timeliness" around gathering mortality data across state, local and federal levels.

Mortality analysis is compiled using death certificate data, which is first gathered at the local level in medical examiners' and coroners' offices. Then, it is funneled up through states to CDC's National Center for Health Statistics (NCHS).

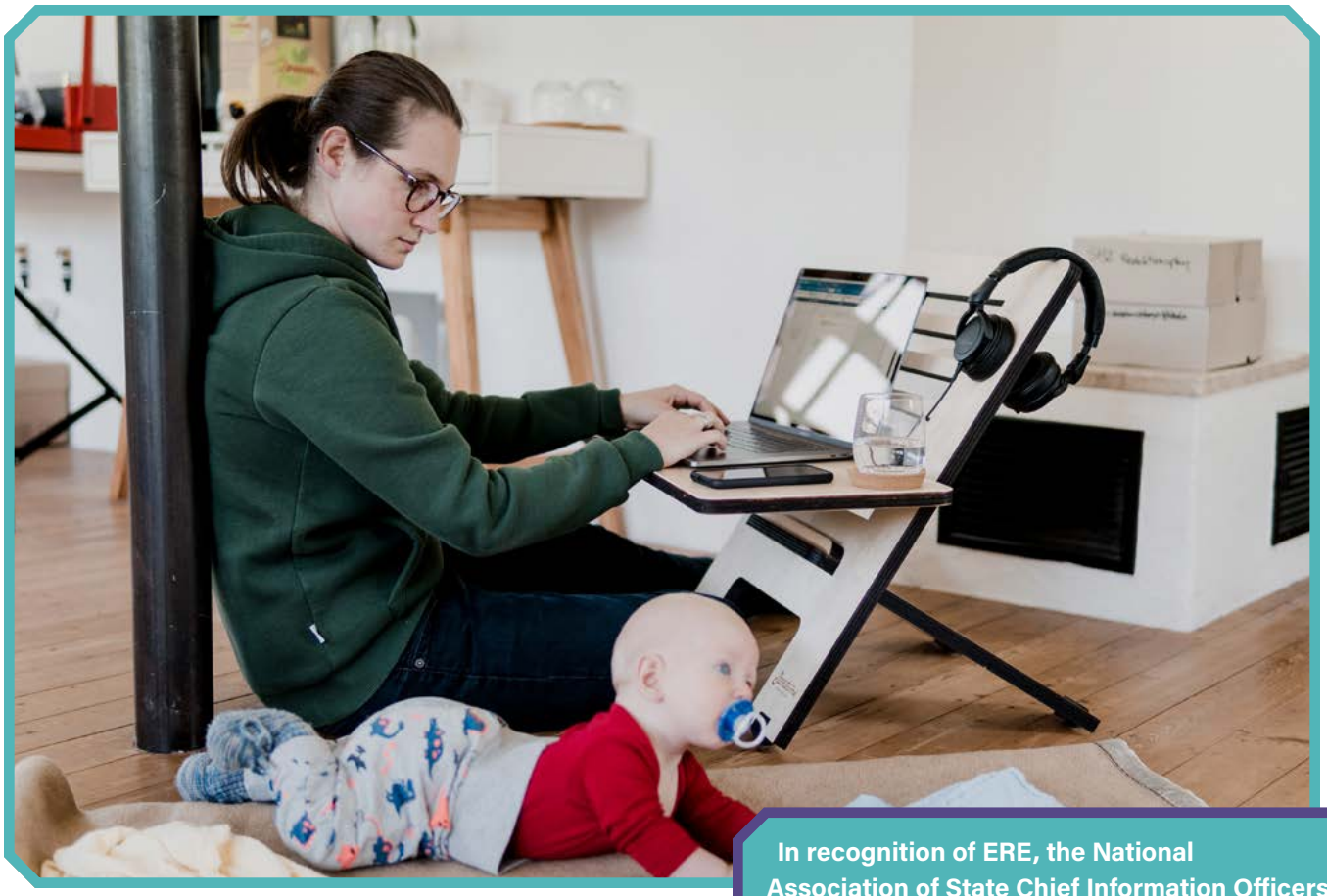
The decentralized nature of reporting makes receiving accurate, timely data at the federal level complex.

- **Electronic systems are often not designed to easily exchange information** across government levels due to paper-based processes and legacy technology.
- Plus, the need for technological interoperability reflects a **larger challenge of getting people to communicate and collaborate** with one another, as Rhonda Smith from Washington, D.C.'s Vital Records Division pointed out.

That's why in 2018, NCHS created an informal working group to improve mortality data sharing.

- The **"implementer's community"** consists of stakeholders across CDC and from several jurisdictions, including Washington, D.C., who meet to develop, test and track "new and more interoperable approaches to sharing important mortality data."

As a result, several jurisdictions have improved their data collection and sharing processes. In turn, CDC receives more timely and accurate mortality reports in times of highest need.



In recognition of ERE, the National Association of State Chief Information Officers (NASCIO) awarded Indiana the State IT award for Innovative and Emerging Technology.

REMOTE WORK

Indiana Builds Secure Collaboration Environment

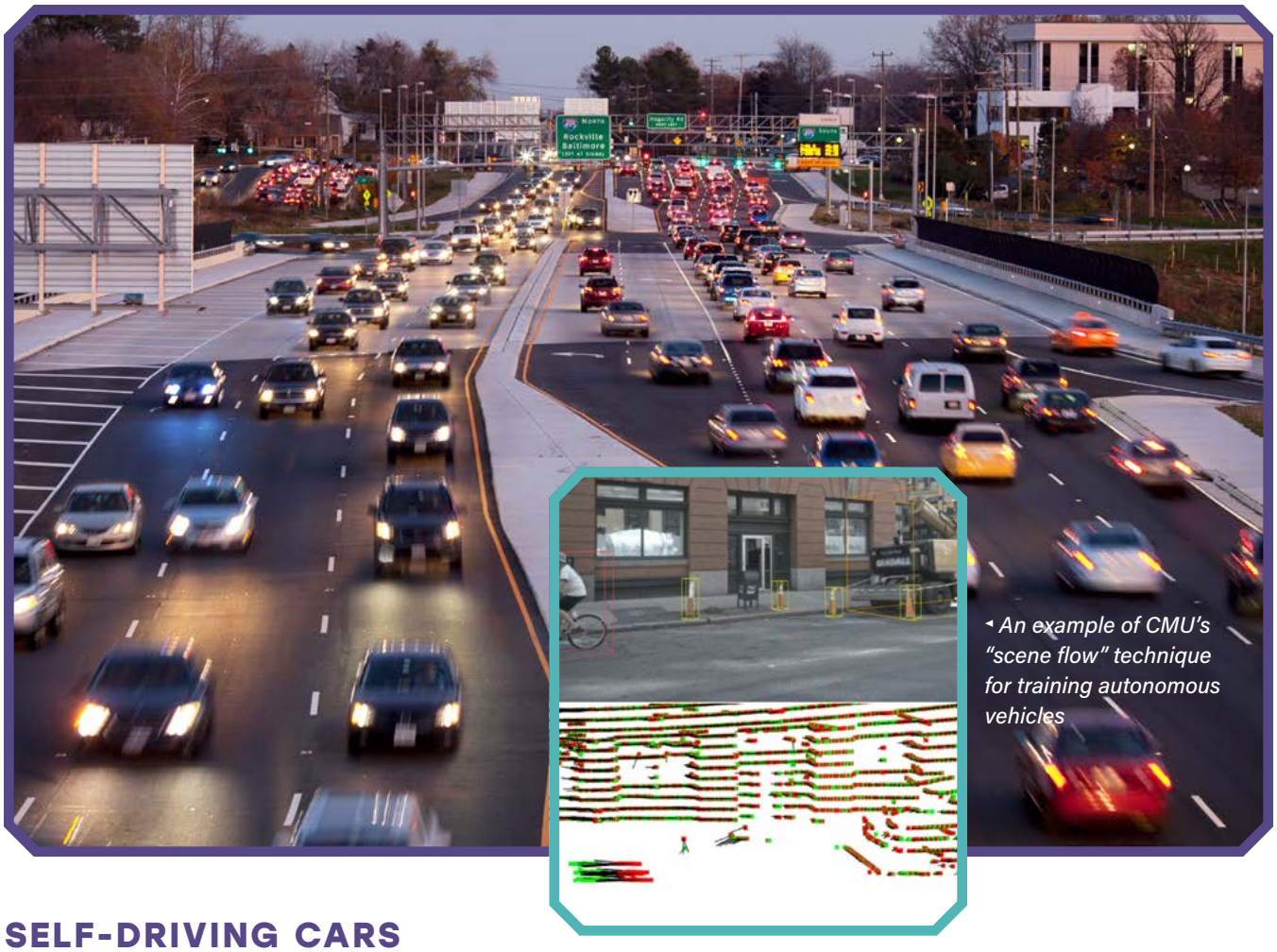
Data silos were an obstacle to collaboration long before the pandemic. Many agencies have struggled to create a single source of truth — one that ensures that everyone in the organization makes decisions based on the same data.

Remote work exacerbated the situation. Now, data was not only scattered across systems, but employees working from home did not necessarily have access to all of those systems.

Serendipitously, Indiana launched the **Enhanced Research Environment (ERE)** just days after many offices began shutting down in March 2020. ERE, in the works since early 2019, has played a pivotal role in the state's response to COVID-19.

- More than 60 leaders from multiple state agencies and research organizations interact daily to derive insights through common tools and data, drawing on more than a dozen data sources.
- They can work with data without removing it from the environment, ensuring its privacy and security.
- ERE is part of the Indiana Management Performance Hub, which provides analytics solutions tailored to address complex management and policy questions for the state.

"In the spring, the state of Indiana was finding ways to act quickly and collaborate to provide the best response and outcomes for Hoosiers," said Indiana CDO Josh Martin. ***"ERE allowed key experts and decision-makers to come together and work with the same data, from the safety of their own homes and offices."***



An example of CMU's "scene flow" technique for training autonomous vehicles

SELF-DRIVING CARS

Data Drives Autonomous Vehicles

Government, academia and industry are going all-in on autonomous vehicles. The Transportation Department (DOT) expects automated driving systems to improve road safety, and the National League of Cities and other urban experts see the potential also to improve traffic management, increase the efficiency of freight and deliveries, and even reduce the costs of transportation for people who are homebound.

Data is the key. To understand and respond to its environment — for example, to safely avoid a collision with a car changing lanes — an autonomous vehicle must continuously gather and analyze data in real time. Experts say autonomous vehicles could produce as much as 4 terabytes of data an hour as they move through a city.

Here are some of the efforts underway:

- The Intelligent Transportation Systems Joint Program Office, part of DOT, is supporting numerous "connected vehicle" pilot projects in various cities, including New York and Tampa, Florida.
- Vermont's Agency of Transportation has developed a permitting process for organizations looking to test autonomous vehicles on state and town highways. In February 2021, the agency hosted a virtual forum on autonomous vehicles.
- Researchers at Carnegie Mellon University (CMU) have developed a new methodology for using traffic data to train autonomous vehicles to navigate more effectively.

"The sheer amount of mobility data being collected daily has transformed data from a planning tool into an integral piece of infrastructure," stated a report from the National Association of City Transportation Officials.

Collibra

Use trusted data to drive government decision making

with the industry's only FedRAMP certified Data Intelligence cloud

Learn more at Collibra.com/public-sector

How to Drive Mission Value with Your Data Governance

An interview with Aileen Black, Senior Vice President, Public Sector, Collibra

When it comes to data governance, the best defense is not necessarily a strong offense.

Data governance, the policies and processes that define how data can or should be used, ought to be focused on supporting an organization's mission and its operational goals.

Unfortunately, many government organizations take a defensive approach, focusing governance efforts on compliance. This creates rigid organizational practices that leave agencies unable to adapt to changing circumstances.

The solution? Adaptive data governance. This flexible governance structure not only ensures compliance but also helps accomplish mission objectives amid change through empowering collaboration.

"Adaptive data governance shifts the onus from IT to a place of collaboration with business users, to enable people regardless of technical background to communicate around data to achieve outcomes with data in the context of constraints imposed by regulation, security and privacy," said Aileen Black, Senior Vice President, Public Sector at Collibra, a leading data intelligence company.

Black outlined three ways that adaptive data governance can drive value for agencies.

1. GROW

Data governance doesn't just impact data programs. It impacts the whole agency. To grow mission capabilities, the governance strategy must be flexible enough to meet the changing needs of the agency and its people.

"Adaptive governance enables leaders to select the appropriate governance style for different business scenarios," Black said. This gives you the flexibility to be innovative and forward-thinking. It's less about checking boxes and more about growing business value for the mission by using governance as a tool.

2. SAVE

When agencies can connect data governance practices to mission outcomes, they are likely to be more cost-efficient. With data governance, agencies can more easily identify the right data and analytics to ensure target outcomes.

Ensuring the data-to-business connection, adaptive data governance helps the IT and agency business teams align around the governance strategy and create a common language. When teams are united, agencies can extract more value from data from the beginning. IT operations and maintenance costs are reduced as technical staff spend less time simply managing data, resulting in agencies becoming more efficient in data operations.

3. PROTECT

Adaptive data governance helps agencies protect data and increase trust when data breaches and security issues are top of mind.

It's impossible to completely avoid risk or prevent all security incidents. For the best outcome, risk and security must be embedded in the process. Adaptive data governance is best suited to do this because it is context-aware. It can address risk and security and fulfill mission objectives instead of the traditional strategy that governs all three separately. It leads to the most effective decision-making, which in turn increases trust.

"Collibra Data Governance allows organizations to create a shared data language to facilitate agile decision-making," Black said. "With Collibra's secure, scalable platform and automated governance and stewardship tasks, agencies can trust in their data and realize all of the benefits of adaptive data governance."

Critical Data Skills: A Checklist

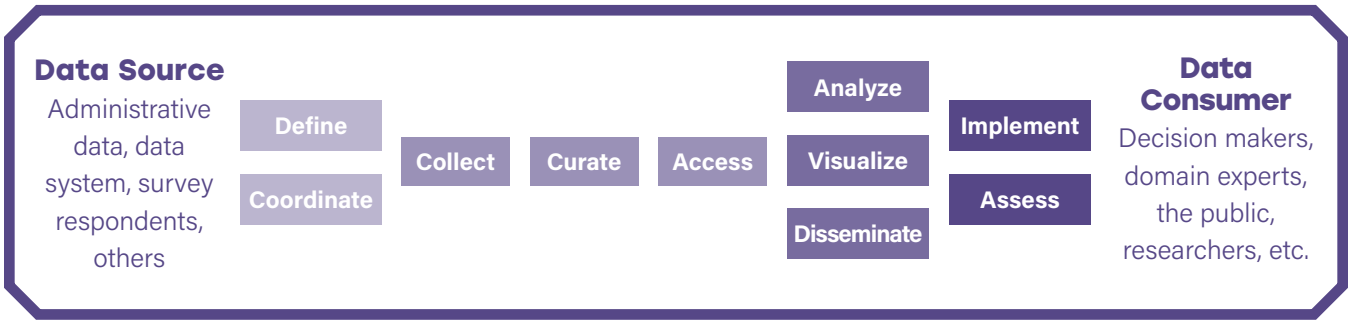
To carry out a data strategy, an organization needs to ensure that its staff has all the necessary skills. Here is how GSA defined the various data roles and their associated responsibilities for the Federal Data Strategy:

- ☐ **Define:** Identify agency and stakeholder needs for data of sufficient quality for intended uses.
- ☐ **Coordinate:** Assess the ability of data resources and infrastructure to meet agency and stakeholder needs.
- ☐ **Collect:** Organize, plan, and execute data collections and acquisitions to meet agency and stakeholder needs.
- ☐ **Curate:** Organize, refine and maintain agency data resources with sufficient quality to meet agency and stakeholder needs.
- ☐ **Access:** Identify and develop multiple data access methods for agency staff and stakeholders.
- ☐ **Analyze:** Optimize the ability of staff and stakeholders to use agency data to generate insights.
- ☐ **Visualize:** Present data insights for consumption by leaders and stakeholders.
- ☐ **Disseminate:** Provide multiple avenues for release of data and insights.
- ☐ **Implement and Assess:** Maximize the use of data for decision-making, accountability, the public good and continuously improving the data process.

HOW TO FILL THE GAPS

Is your organization facing a data skills gap? GSA suggests the following strategies:

- ▷ Use the agency's data governance body to analyze whether another part of the agency (e.g., the analytics office can provide analysis support to other units) can fill gaps.
- ▷ Coordinate and share lessons learned in addressing skills gaps through governmentwide councils such as the CDO Council.
- ▷ Convene subject-matter expert meetings to discuss topics ranging from developing workforce planning, upskilling and reskilling to improving recruitment processes, job analysis documentation and other human resources initiatives.
- ▷ Bring the voice of the employee to the table in data skill discussions and define data skills that all staff should have versus those appropriate for more advanced staff. Then, customize the data skills training program based upon the level of competency different groups of employees need.
- ▷ Identify successful change management strategies for managers to help them better understand how data can enhance the work of the organization.



Cutting 'Data Gems' to Train Census Data Users

At Census Academy

An interview with Alexandra Barker, Co-Founder, Census Academy, and Supervisor, Data Dissemination and Training Branch, Census Bureau

When Census Academy first launched in May 2019, the creators knew it would be a success. They didn't know it would be a hit.

"I remember the first six months, we couldn't even manage the amount of feedback and thank you notes we were getting," said Alexandra Barker, Supervisor at the Census Bureau's Data Dissemination and Training Branch, and unofficially, co-founder of Census Academy.

The video hub for Census data training is aimed for general data users at all levels, particularly adult learners who often need to use Census data for their jobs.

In April 2021, GovLoop interviewed Barker to hear the story of how Census Academy was established, how it is helping people today and how it can help your agency now.

The interview has been lightly edited for clarity and brevity.

GOVLOOP: Why was Census Academy created?

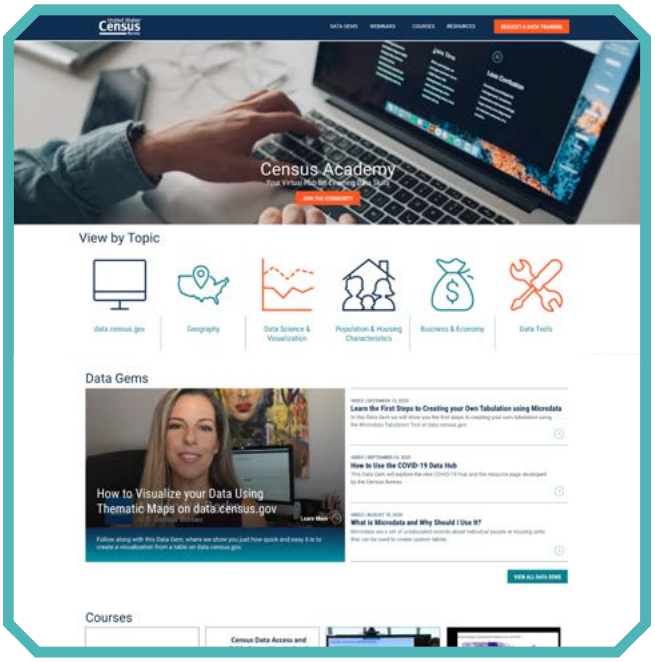
Barker: Before, we only offered data training to the public via webinar or in person, and these had to be provided based on request. So, you had to know we existed and you had to be knowledgeable about our services. And an organization could request one of our trainings, but it was only available at that time for those guests of that organization. It simply wasn't accessible to everyone who needed it at all times.

Think of a grant writer with a deadline at midnight and they need help, say, two hours before the clock. There's no workshop for them, no presentation. A person like that is more likely to start Googling information. So we felt there was a gap there. We needed to put content for that user, who may not even visit our website but is in need of our data and in need of our help. That's how we came about to develop the concept of Census Academy.

Is there any memorable feedback you've received from users since the launch?

To me, this is actually a challenging question, because there was just an outpouring of feedback. We knew based on our research, as well as our experience working with data users for so long, that it would be successful. But as soon as we launched, I remember the first six months, we couldn't even manage the amount of feedback and thank you notes we were getting. I remember a congressional office in Massachusetts said, "Well, it was about time," — but they also said, "Thank you, this is really good. We're going to forward to our constituents."

Aside from all the "thank yous" and "it's about time," we got a lot of ideas. It felt that people were excited and jumping onboard — "Do this!" "Can you do that?" We were not prepared to manage the amount of ideas that were sent to us from users and stakeholders on data gems. Data gems are one of our most popular products, for your information. The amount of information for data gems was incredible. We still receive a receive a ton today.



How did you end up in your role, founding Census Academy and being passionate about data education?

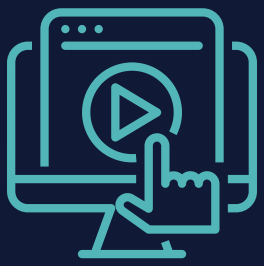
When I entered journalism school in 1999, I actually hated numbers. I was so glad I wasn't going to have to touch math anymore. But when I saw the power of numbers in telling a story that could change things, I became fascinated. That's why I say I'm a data nerd now.

Getting my first job at the Census Bureau as a media specialist was a way for me to start bringing my journalism background — loving to tell stories — together with numbers. And from that point on, when I became a data dissemination specialist and I started working with the users day to day, I learned about the challenges regular users had in accessing our data. And I became really invested in bringing data education to diverse audiences. Every day, I was working with a different group — one day, I was in a school, the next day in a tribe, the next day a congressional office or business. I became invested in making sure we had a solution that was scalable, efficient and effective for all the users we had.

What do you want government employees who don't have data-related roles to know about data?

Numbers can tell a powerful story. Regardless of your role, data can help any staff demonstrate a need, make a business case, make informed decisions and so much more. For any professional, it's a critical skill to have.

HOW AGENCIES CAN USE CENSUS ACADEMY FOR THEIR NEEDS



Use it as a model for teaching data skills or any complex information.

"Our research demonstrated that about 75% of users who are adult learners wanted to learn through videos. Short videos, a few minutes with the content I need at the time I need. It's called microlearning. Using that to teach data or any complex content internally or externally is valuable."

Offer it to employees who help perform community assessments and grant writing.

"We have a lot of federal agencies that provide grants, and they have requirements for those grants. And many of those requirements require census data. When agencies are working with community assessments for developing the requirements for their grant, it's important that their staff is familiar with census data — what's available and how to access it."

Collaborate to create a course or set up your own program at your organization.

"We were approached by the Department of Labor to help them with the development of their own Census Academy. They feel they have a lot of information and data they need to share with users. It's been about a year we've been in consistent communication with them, holding hands to try to get them off the ground."

Worksheet: Applying the Federal Data Ethics Framework

As part of the Federal Data Strategy, GSA released the Data Ethics Framework in December 2020 to help encourage ethical data decision-making throughout the federal government.

Although the framework does not provide mandates, the guidelines are imperative for mitigating risk and serving federal customers and constituents. For complete details, recommendations and use cases, check out the framework in full.

To help you start applying these tenets in your everyday work, fill out the worksheet below. The prompts are based off of the framework's recommendations.

---

1. Uphold Applicable Statutes, Regulations, Professional Practices & Ethical Standards

List the legal authorities, professional codes of conduct and ethical standards that apply to your agency.

2. Respect the Public, Individuals and Communities

List 2-3 ways your organization's data activities might impact the public, both individuals and communities.

List 1-2 procedures to mitigate harm when negative consequences are unavoidable.

3. Respect Privacy and Confidentiality

Check off what your organization is doing:

- ☐ Providing employee training on appropriate handling of sensitive data
- ☐ Monitoring technological advances that increase or minimize the risk of identification of individuals or entities represented in public datasets
- ☐ Regularly updating disclosure protection protocols to mitigate privacy and confidentiality risks to the greatest extent possible
- ☐ Making clear the tradeoff between confidentiality and granularity of data in public data releases
- ☐ Providing tools to enable users to evaluate the impact data protection measures have on results obtained from data
- ☐ Establishing protocols for notifying data providers and other relevant stakeholders if there is a breach that potentially impacts their privacy
- ☐ Providing mechanisms for accommodating victims of those breaches
- ☐ Supporting and implementing mechanisms that limit risks such as disclosure limitations and controlled access to data
- ☐ Complying with applicable legal authorities that govern the protection and use of sensitive data
- ☐ Establishing or have established policies and procedures to prevent re-identification of sensitive data made public
- ☐ Maintaining the minimum amount of sensitive data necessary
- ☐ Adhering to data sharing and use agreements

4. Act With Honesty, Integrity and Humility

How are you developing a culture of honesty, integrity and humility within your organization?

5. Stay Informed of Developments in the Fields of Data Management and Data Science

How are you developing a diverse workforce to support the policies, oversight and governance structure for any large-scale system that learns from data? Jot down 1-2 ways.

6. Hold Oneself and Others Accountable

Jot down how you stay accountable to the following data stakeholders, as it pertains to your organization:

» Individuals and communities providing data as respondents or research subjects

» Those directly impacted by data use (e.g., recipients of services and benefits)

» Members of the public who rely on data products

» Data consumers, including customers or clients requesting data (e.g., agency program managers, Congress members and their staff, etc.)

7. Promote Transparency

How are you nurturing a culture that supports appropriate, transparent reporting of data activities and products? List 2-3 ways.

Conclusion

In many cases, governments have been the primary purveyors, dominant distributors and colossal collectors of data for decades. They have invaluable assets at their fingertips — the challenge lies in applying them in new, modern ways.

We hope this guide can help you start tackling that challenge. Now that you understand key concepts that every successful data effort requires and found inspiration in the case studies and Q&As with thought leaders, you're ready to take the next step. Start by filling out the worksheets, answering the questions honestly, to help get your organization closer to its data goals.

ABOUT CARAHSOFT

Carahsoft Technology Corp. is The Trusted Government IT Solutions Provider®, supporting Public Sector organizations across Federal, State and Local Government agencies and Education and Healthcare markets. As the Master Government Aggregator® for our vendor partners, we deliver solutions for Cybersecurity, MultiCloud, DevSecOps, Big Data, Artificial Intelligence, Open Source, Customer Experience and more. Working with resellers, systems integrators and consultants, our sales and marketing teams provide industry leading IT products, services and training through hundreds of contract vehicles.

Visit www.carahsoft.com, follow @Carahsoft, or email sales@carahsoft.com for more information.

THANK YOU

Thank you to Alteryx, Collibra, MongoDB and SAP NS2 for their support of this valuable resource for public sector professionals.

ABOUT GOVLOOP

GovLoop's mission is to inspire public sector professionals by serving as the knowledge network for government. GovLoop connects more than 300,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington, D.C., with a team of dedicated professionals who share a commitment to the public sector. For more information about this report, please reach out to info@govloop.com.

govloop.com | @govloop

AUTHOR

Pearl Kim, Staff Writer

DESIGNER

Kaitlyn Baker, Creative Manager

carahsoft

Carahsoft's data analytics solutions are available through its reseller partners on a variety of contracts including Carahsoft's GSA Schedule 70, SEWP V, ITES-SW2, NASPO ValuePoint, NCPA, OMNIA Partners and numerous State and Local contracts.

Learn more at carahsoft.com/data.

See the latest innovations in Government IT from Carahsoft's vendor partners at carahsoft.com/innovation.

govloop

1152 15th St. NW Suite 800
Washington, DC 20005

P: (202) 407-7421 | F: (202) 407-7501

www.govloop.com
@GovLoop