Building Government's Data Toolbox:

Practical Uses for Data and Insights

govloop

Introduction

Agencies have long recognized the need for secure, strategic data use, and the growth of AI tools — whether they're like elves working behind the scenes or like media-grabbing movie stars — has made data even more critical. Still, there is a difference between knowing and implementing.

In this guide, the second in our three-part data series, we give tangible guidance on using data to advance your agency's mission. Case studies from the federal government and cities and states nationwide show how data analysis can enhance transparency, promote efficiency and improve outcomes.

No data conversation can ignore AI, so we also interviewed a state chief information officer (CIO) with a long history implementing data-driven reforms that are so vital to AI success. And there's a case study here also: a two-part Mississippi strategy, focused on the cloud and a statewide data exchange, that will transform the state's data use and management.

Used wisely and with forethought, data can help your agency improve its current operations and take advantage of emerging technology. Read on to learn how.

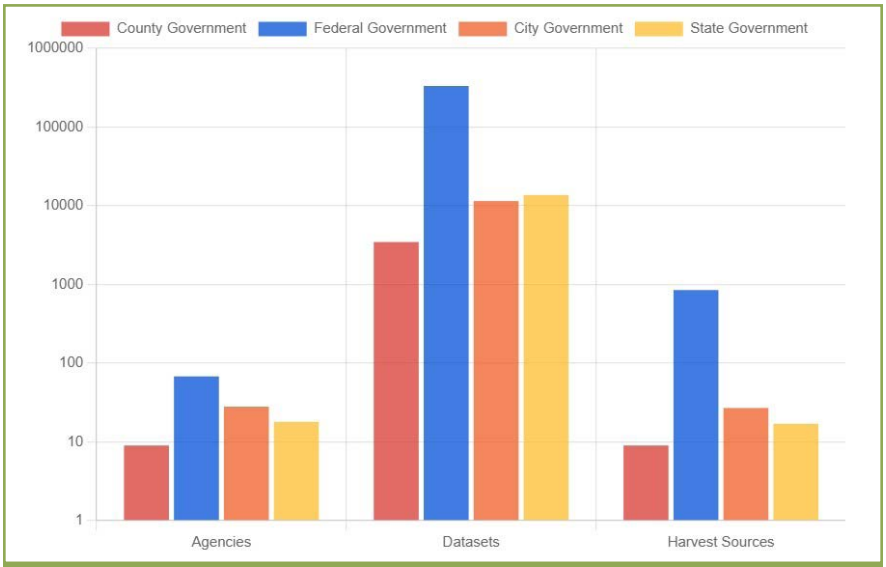
Contents

A Data Building Block

There seems to be an infinite number of data sources for agencies to draw from, but Data.gov is one of the most helpful for organizations at all levels. It's the federal government's open data website, a central repository that offers nearly 370,000 datasets from federal, state, county and city agencies.

You can find the data, tools and resources you need to conduct research, develop web and mobile applications, design data visualizations, and more. Datasets are continually updated.

Key Metrics: Where the Data Comes From

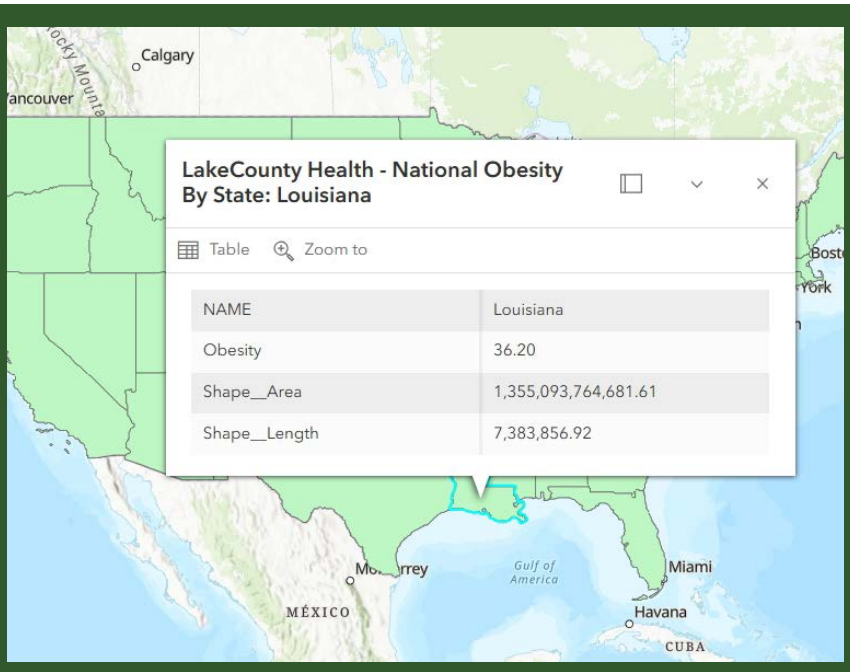


Source: Data.gov

You can search Data.gov information by several filters, including:

- Topic
- Organization/agency
- Level of government
- Popularity
- Date submitted/updated
- Format
- Key word
- Location (e.g., ZIP code)

Data.gov connects to 273,000-plus **geospatial datasets** — that is, data tied to specific locations — which often are found on a sister site, GeoPlatform.gov, as well. The geospatial data covers topics such as the location of wind turbines across the U.S., sites that receive low-income housing tax credits, seismic monitoring centers, elevations of San Francisco Bay wetlands, EPA Superfund sites, and obesity levels per state.



Source: GeoPlatform.gov

A Data.gov Timeline

May 21, 2009:
Data.gov launches with 47 datasets.

May 9, 2013:

- President Obama signs an executive order that makes open and machine-readable data the new default for government information.
- The White House issues the Open Data Policy, which creates a framework to help institutionalize effective data-management principles at all stages of the data life cycle. The goal: to promote interoperability and openness.

January 4, 2019:
President Trump signs the OPEN Government Data Act, which requires federal agencies — by law, not as a general policy — to publish their data online using standardized, machine-readable formats, and to include their metadata in the Data.gov catalog.

March 30, 2023:
Data.gov launches a redesigned homepage.

Most Downloaded Datasets

(as of September 9, 2025)

Electric Vehicle Population Data
Washington State

Lottery Powerball Winning Numbers: Beginning in 2010
NYS

Supply Chain Greenhouse Gas Emission Factors
EPA

Warehouse and Retail Sales
Montgomery County, Maryland

Crime Data from 2020 to Present
City of Los Angeles

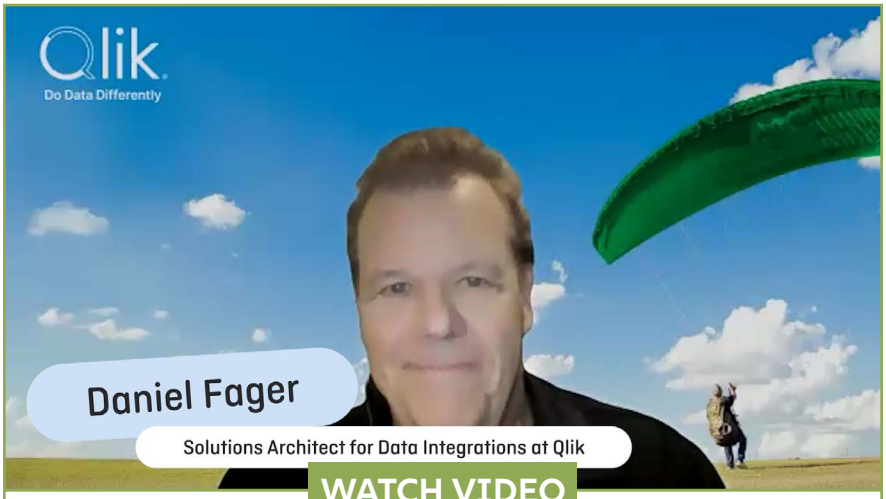
FDIC Failed Bank List
Federal Deposit Insurance Corp.

Border Crossing Entry Data
U.S. Department of Transportation

Data.gov also provides lists of all available open data websites in the U.S. and internationally.

INDUSTRY SPOTLIGHT

Constructing a Strong Data Foundation



Agencies often spend more time wrangling their data than using it. Cleaning and reconciling information consumes valuable hours and can jeopardize data trustworthiness and validity. In addition, incorporating AI-driven analytics makes the need for good data all the more important.

By building a strong data foundation based on secure connectivity to diverse data sources, real-time hybrid pipelines and built-in governance enhanced by AI-driven data quality, organizations can address these issues to produce data that's trusted, up to date and ready to generate insights that support and enhance the mission.

In this video interview, Daniel Fager, Solutions Architect for Data Integrations at Qlik, explains how a modern data foundation can maximize the value of your data.

Topics include:

- The elements that define a strong data foundation.
- Why good data management is more than just breaking down silos.
- How a strong data foundation benefits the mission.

"Agencies need to know that the data is accurate, complete and timely, and that it can be safely shared across programs under strict compliance rules. That gives program leaders confidence to use the data for analytics, policymaking, and AI without second-guessing whether they can rely on it."

— Daniel Fager, Qlik

About ORI

ORI is a certified WOSB and leading U.S. government technology partner and reseller with 35+ years of federal government experience. ORI connects the public sector with industry-leading technologies spanning the data operations life cycle — from preparation and integration to data science and advanced analytics — to deliver comprehensive, adaptive, secure solutions.

Learn more about ORI

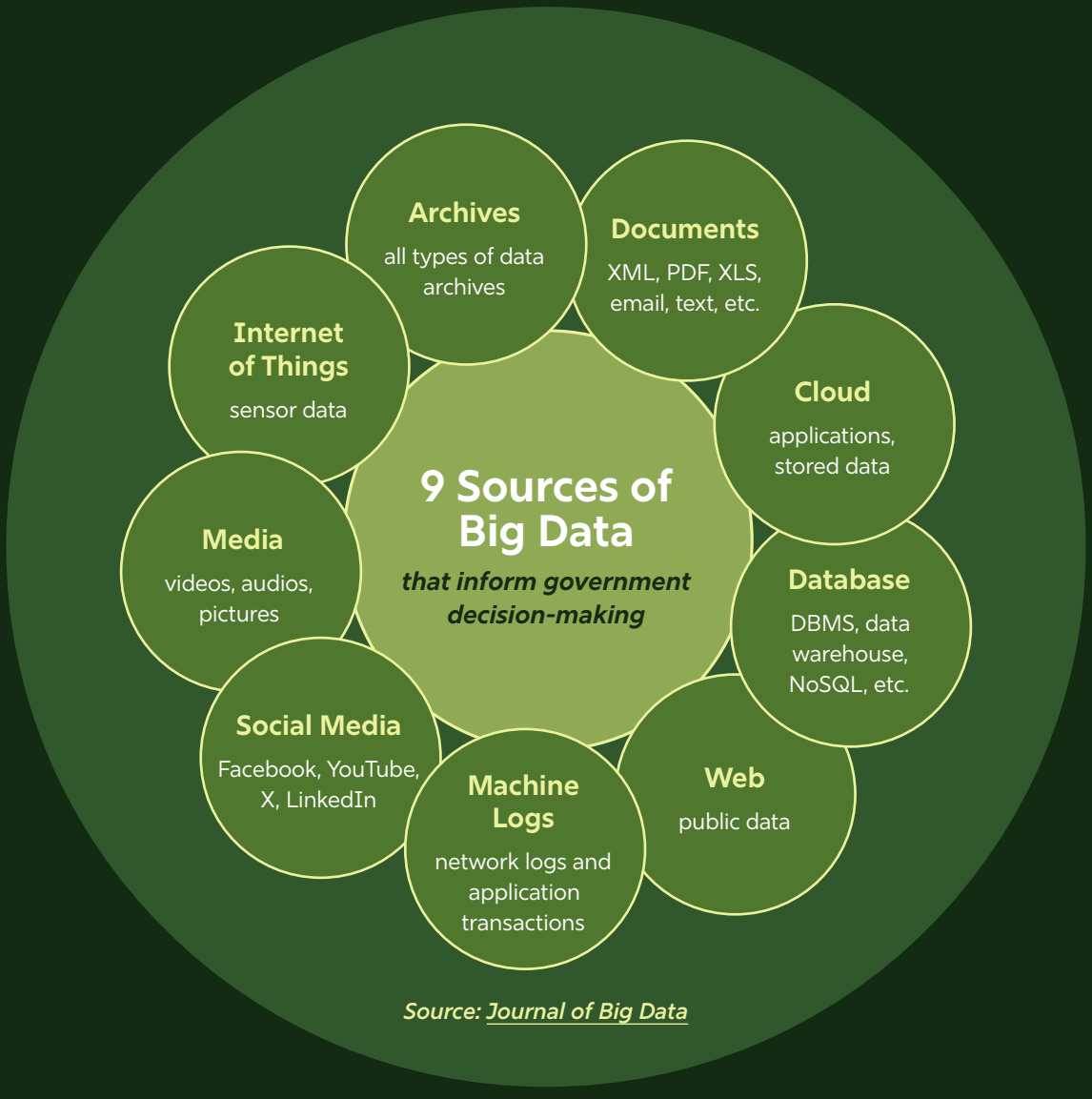
ORI | Qlik

Where Data Drives Mission Success

There are many examples of government creatively using data to address specific issues, such as public safety, workflow management, benefit payments, and funding. Those specific, mission-focused efforts fall into one or more broad categories, such as:

1. **Enhancing Transparency**
2. **Promoting Efficiency**
3. **Improving Outcomes**

In the following pages, we explore case studies in which agencies at all levels used data to realize these objectives.



Source: Journal of Big Data

1 Enhancing Transparency

ISSUE: PUBLIC SPENDING

Follow the Money: U.S. Treasury Discloses Federal Income and Outlays

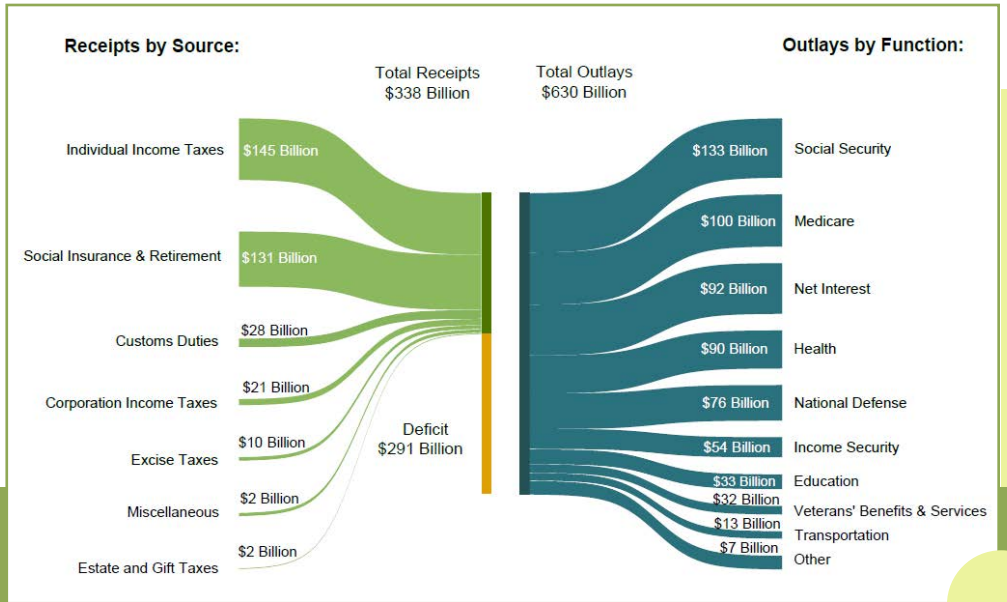
The general public knows that trillions of dollars pass through the federal budget, but people are less clear on where that money goes and when the bills are paid — an uncertainty that breeds confusion and distrust. So to help provide insight into federal revenue and expenditures, the U.S. Treasury Department publishes monthly statements — akin to our personal bank statements — that show federal receipts, outlays and deficits.

For instance, the July 2025 report shows that Treasury received $131 billion in social insurance and retirement–related income and paid out $133 billion for Social Security and $76 billion in national defense, among other statistics.

Treasury also publishes daily statements that explain, in granular detail, how much money more than 150 individual departments and agencies have deposited and withdrawn on a given day. The daily statements disclose the Treasury's General Account (TGA) opening and closing balances as well, comparing them to the TGA balances that month and through the current fiscal year.

The statements are available on Treasury's Fiscal Data website — billed as a "one-stop shop for federal financial data" — which provides educational content, interactive visuals, historical context and running tabulations of the national deficit and debt, among other information.

"Usability is a core value we have embraced from the start of the development of this website," its designers say. "Throughout development, we have integrated best practices and user feedback to deliver a modern, easy-to-navigate site."



Receipts, Outlays and Surplus/Deficit for July 2025

Source: Monthly Treasury Statement, FY 2025 through July 31, 2025

ISSUE: WORKFORCE

Massachusetts Reveals State Workforce Demographics

Government agencies work hard to understand the demographics of communities they serve, but the public sector is less likely to turn the mirror on itself. In Massachusetts, however, an interactive dashboard reveals the overall composition of the executive branch workforce. It tracks the state's progress toward achieving various diversity benchmarks and uses colorful graphics to show demographic details such as gender, race, age group, length of service, and veteran and disability status. Also included: the number of executive branch hires, promotions and separations.

2 Promoting Efficiency

ISSUE: INSPECTIONS

In Chicago, Open Data Improves Food Safety

Like many communities, Chicago has an abundance of restaurants — more than 15,000, in fact — and a proportionally small number of inspectors, just 36, to monitor food safety. Previously, those numbers almost ensured that officials would overlook critical food-safety hazards at restaurants citywide, putting public health at risk.

Open data, however, now helps inspectors pinpoint high-priority safety violations, which saves precious time and resources. To begin modernizing the city's approach, analysts from the Department of Public Health and the Department of Technology and Innovation teamed up with volunteer data scientists from Allstate to collect information from various Chicago data sources, such as food inspection history, weather records and 311 calls.

From there, the group identified nine variables that often forecast critical violations and developed an algorithm that predicts which restaurants are most likely to fail food-safety inspections. The algorithm quickly proved effective. During a pilot phase, for instance, it helped inspectors identify 69% of critical violations early — compared to just 55% revealed early using the traditional method. From a calendar-day perspective, the algorithm helped inspectors find major food-safety hazards nearly eight days sooner across a 60-day period.

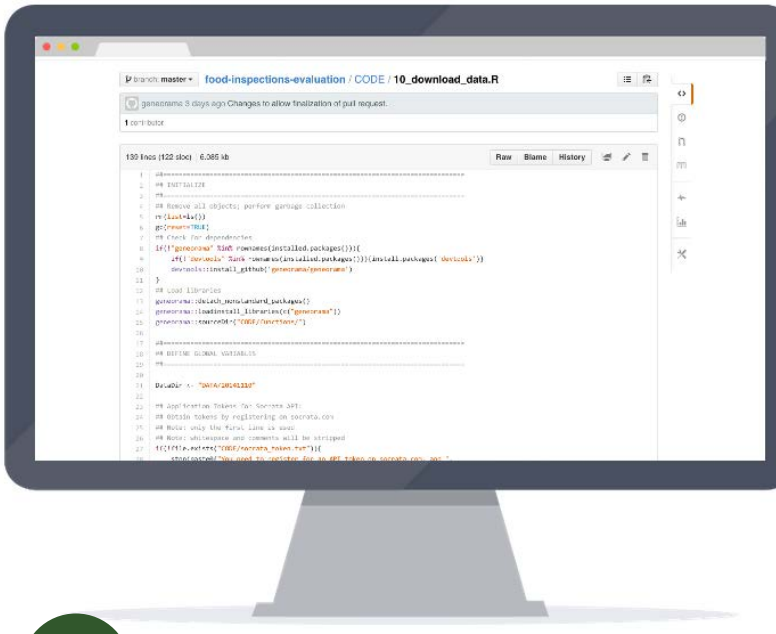
By using open data already available on Chicago's data portal, the research team jump-started the project in two weeks, rather than suffer a longer contractual process to give Allstate volunteers access to non-public city information. The inspection algorithm proved so effective that Chicago made the project code open source, allowing other city departments to use it for their own safety monitoring.

ISSUE: INFRASTRUCTURE MANAGEMENT

Widescale Data Strategy Helps Nuclear Regulators Maintain Facilities

The Department of Energy's National Nuclear Security Administration (NNSA) has weighty responsibilities, including overseeing roughly 44,000 federal and contractor employees and operating a vast, complex portfolio of contractor-run labs, manufacturing and production facilities, and other infrastructure. Each site generally uses its own location-specific maintenance management software — which is fine at the site level, but doesn't give NNSA an overall risk picture.

To address this challenge, the agency teamed with the U.S. Army Corps of Engineers to develop the BUILDER Sustainment Management System. It integrates each site's existing processes into a single, standardized network that helps NNSA make proactive, enterprisewide decisions. The new strategy is similar to looking through the front windshield, an NNSA official said, rather than in the rearview mirror.



Food Inspection Model on Github
Source: City of Chicago

3 Improving Outcomes

ISSUE: LAW ENFORCEMENT

GIS Gives St. Petersburg, Florida, Police Deep Insights Into Crime Data

For years, Florida's St. Petersburg Police Department practiced data-driven decision-making, providing officers, detectives and supervisors with various metrics to use in making patrol assignments or investigating incidents. The problem was that the data was stored in spreadsheets, which made it difficult to translate into meaningful information.

For example, a spreadsheet might show the addresses of houses burglarized in recent weeks, but where was each house in relation to the others? Were the burglaries happening randomly across the city or clustered in one area? An investigator staring at a spreadsheet would have had a hard time finding a pattern.

Today, that data is displayed on a digital map. The department now stores all its data on a central geospatial platform, Esri's ArcGIS, that — beyond making the data easier to manage and access — provides tools for analysis and visualization. For example, the department can track key metrics in visual dashboards, while analysts can create interactive maps focused on specific datasets.

Think about investigating the destruction of parking meters downtown. Analysts mapped each incident to determine where vandals might strike next and sent patrols to that area. The result? Patrol officers caught the criminals. That's how data-driven decision-making is supposed to work.

The ArcGIS platform supports the department's use of the Compstat model of policing, which officials adopted in 2020. CompStat, short for computer statistics, entails regular meetings in which department executives and officers review crime data, identify key trends, decide on the best courses of action and track results.

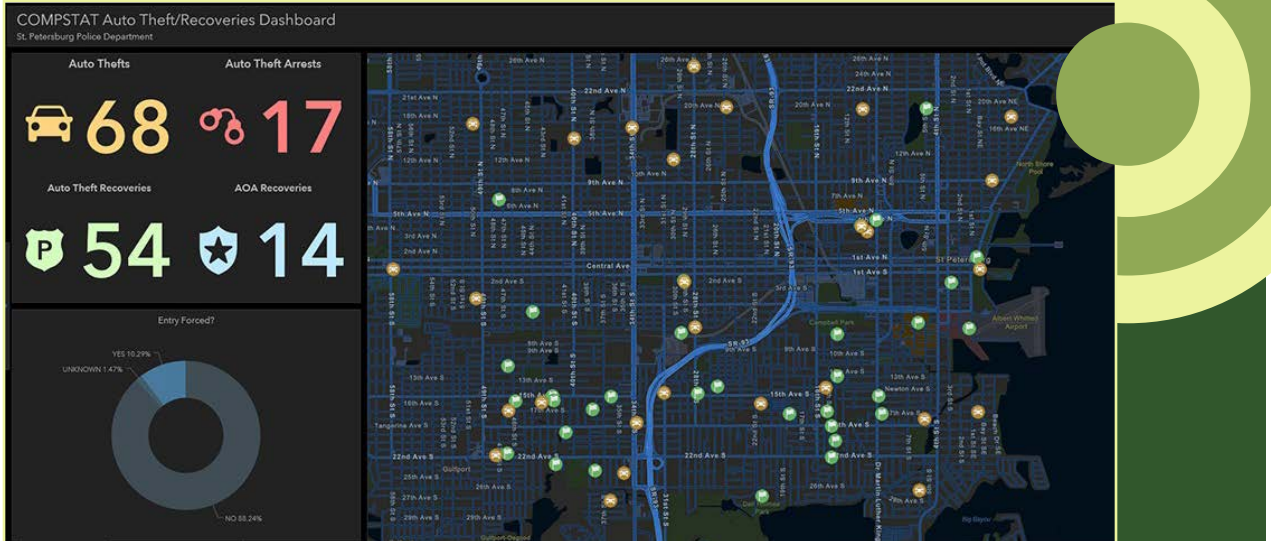
ISSUE: HEALTHCARE DELIVERY

Oklahoma Brings Data Focus to Medicaid Health Outcomes

SoonerCare, Oklahoma's Medicaid program, covers slightly more than 1 million residents, or roughly 25% of the state population. The Oklahoma Health Care Authority (OHCA) runs SoonerCare as a fee-for-service program, with services delivered through about 70,000 healthcare providers. The challenge, always, is to enhance service quality while improving cost effectiveness.

With that in mind, in April 2024 OHCA launched the SoonerSelect initiative, which includes a Quality Assessment and Performance Improvement program that provides metrics and processes for tracking healthcare providers' performance. The program has teeth: OHCA withholds a portion of capitation payments (that is, the fixed, per-patient fee paid to providers), but providers can earn back the withheld funds by achieving certain metrics.

Source: Esri



INDUSTRY SPOTLIGHT

A Modern Solution to Data Storage and Management

Agencies today need fast, reliable, secure access to their data, but their data infrastructure often lets them down. It's too inflexible to handle the explosion of data generated by video surveillance, the Internet of Things, and other sources, and it creates silos that hinder data retrieval and analysis. When time is of the essence — say, when first responders are being deployed during an emergency — agencies cannot wait for their data systems to catch up.

There's a modern alternative: an enterprise data cloud (EDC). While traditional data storage focuses on infrastructure, EDC offers a more holistic approach that lets state and local agencies unify, automate and intelligently control all their data, whether it's on premises, in the cloud, or in a hybrid environment.

In this video interview, Jim Weaver, Pure Storage's National Strategy Advisor for U.S. Public Sector SLED, discusses how agencies can maximize their data's potential with modern, enterprisewide data management. Topics include:

- How agencies can break away from legacy technology
- How an EDC helps agencies deliver outcomes, not just oversee infrastructure
- What to prioritize when considering a modern data solution

"Because technology is going to continue to evolve, resiliency is going to be key, and whether it's increased cyber threats or budget deficits, you need a solution that's going to grow and sail the choppy waves, if you will, as they come."

— Jim Weaver, Pure Storage

About Pure Storage

Pure Storage® delivers the industry's most advanced data storage platform to store, manage, and protect government data at any scale. With Pure Storage, agencies gain ultimate simplicity and flexibility, enabling them to save time, reduce costs, and increase operational efficiency. From AI to archive, Pure Storage provides a modern cloud experience through one unified Storage as-a-Service platform across on-premises, cloud, and hosted environments.

Learn more about Pure Storage

PURESTORAGE®

GOVERNMENT SPOTLIGHT

Deciphering the AI Need for Data

Whether bold and showy or working quietly behind the scenes, AI tools help agencies, and society at large, be more efficient, effective and often more creative. Technology that once seemed fanciful has become ubiquitous. But good AI relies on good-quality data and planning — and therein lies the rub for many government agencies.

"There are a lot of entities that view data as just part of the backend systems of record," said Dr. Craig Orgeron, Executive Director of the Mississippi Department of Information Technology Services (ITS) and state CIO. "I don't know that they view data in a strategic way. And I think the AI conversation does that. It's meant to carry that issue forward."

Trusted Sources, Truer Answers

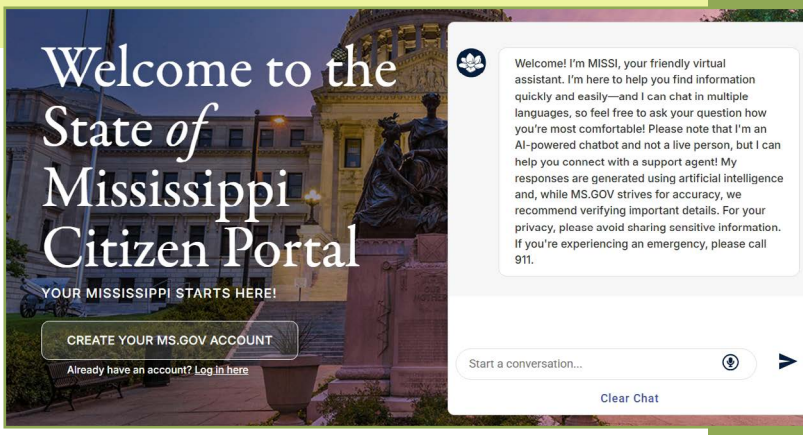
An AI model built on reliable datasets will beat any clever AI model sitting on poor data, Orgeron said. Yet agencies often fixate on generative AI's transformative potential, for instance, but pay scant attention to its underpinnings.

"We're awash in data," he said, "but I've been struck that in the different sectors, it's rarely pristine, and I think there needs to be much more effort in that space if we want these AI models to perform. There's probably more work to do than folks really let on."

That work often includes assigning **ownership**, or stewardship, of priority datasets; using valid **sources of record**; **data profiling**; and **metadata** that makes discovery and interoperability easier. Guarding against **security and privacy** threats and implementing strong data **governance** also are vital to cultivating good data, Orgeron said.

Because think about service delivery, an area in which AI technology has become popular. An AI tool built with low-quality data can tell people the wrong deadlines, incorrectly explain legal requirements and share misleading eligibility rules, among other hazards. But with reliable sources, AI can be a real asset.

For example, Mississippi recently made its chatbot — MISSI — smarter by adding AI capabilities. Now people can more easily learn about applying for licenses and permits, state park closing times, and other straightforward details.



In the future, agentic AI embedded in enterprise-scale software packages will interact with agentic AI in other software, Orgeron said. In other words, an autonomous agent in your enterprise resource planning software could work, without human interaction, with the agentic AI in your customer relationship management system. That would exponentially boost the powers of each agentic AI.

There also will be more opportunities for AI-driven analysis and auditing, summary preparation, and document intake, among other improvements. **"'Do more with less' is an enduring government mantra, and AI tools can be very, genuinely helpful in that space," Orgeron said.**

4 Best Practices

About 80% of AI projects fail — roughly double the failure rate for IT projects without an AI component. Root causes include inadequate training data and misdirected goals, with public-sector initiatives particularly at risk from poor data management. There may be no fail-safes for developing government AI innovations, but Orgeron believes that certain best practices encourage success.

First, **let use cases lead the way**. Identify pain points, then build the use cases to address them. Look for lower-risk use cases that would deliver high value. From a service-delivery perspective, how can you give residents and agency employees straightforward answers faster? What AI changes will lead to visible improvements in how people request and receive services?

Next, **keep humans in the loop**. Some people may interpret human involvement as anti-innovative, but Orgeron doesn't see it that way. For example, employees can play key roles monitoring for privacy and security risks. Think of AI as a digital colleague — a "force multiplier for what public servants can do in the near term," he said, rather than a replacement for agency staff.

Build necessary precautions in early. Beyond security and privacy, plan for potential model drift, which is when AI models degrade over time because of behavioral, environmental or other changes. Responsible and explainable AI are less discussed than they used to be, but bias remains very real, Orgeron said, and AI models should be designed to address it.

In addition, **build partnerships**. By bringing together entities — state agencies, industry, higher education and other stakeholders — you can drive progress, create momentum and accelerate safe adoption of AI solutions. For instance, the Mississippi Artificial Intelligence Network and the state's AI Collaborative provide statewide AI leadership, education and information sharing.

"I think that in the future, whether it's agentic or retrieval-augmented models, the [data] that you feed it will be way more relevant to the solutions," Orgeron said. **"That's why bringing the data conversation to the fore is the exact right thing to do."**

A Case Study in Data Strategy

In Mississippi, AI added weight to developing a thoughtful data strategy. "There's been a ton of activity in the last year, setting the stage for a very different experience than the one we've had," said Craig Orgeron, who leads the Department of Information Technology Services and serves as state CIO.

The approach is two-fold.

1. A Cloud Center of Excellence within ITS.
It's a new framework for state agency migration to and management of **cloud computing** and features best practices, governance structures and other guidance. The center improves statewide security, scalability, and cost efficiency of cloud operations.

"I would say that for the last 25 years at least, you had a lot of proliferation of premise-based IT in agencies, and we see that changing," Orgeron said.

2. A statewide data exchange.
This will be a cloud-based platform to facilitate data sharing among state agencies and authorized entities — and will allow for interoperability among programs, data governance and privacy protections.

The **data exchange** project kicks off with a feasibility study, followed by a phased-in implementation plan, a statewide pilot project and full, statewide rollout.

INDUSTRY SPOTLIGHT

Data Resiliency: Knowing and Using Your Data, With Automation and AI

WATCH VIDEO

Government agencies rely on data to make informed decisions across a wide range of functions, from administrative tasks to critical national security initiatives. However, agencies often struggle with redundant, outdated, or irrelevant information, complicating efforts to achieve operational efficiency and cybersecurity.

It frequently is too difficult, though, for agencies to address these challenges on their own, to understand the data they have and achieve data resiliency. That's when agencies need an outside partner to provide critical, independent insight, so they can be more efficient, data-driven and cybersecure and take advantage of artificial intelligence and other technology.

In this video interview, Richard Breakiron, Commvault's Senior Director of Strategic Initiatives in the Americas Public Sector, discusses how agencies can be more effective stewards of their data. Topics include:

- Strategies for becoming more data resilient
- Why automation and AI are critical for data resiliency and cost savings
- Integrating data automation and AI into a zero-trust architecture

"All of us have gone through moves where you say, 'Before I leave, I'm going to clean my old house and not take [everything]. And you still take everything. With data, you need an outside person to say, 'No, you don't need this, or you need to get rid of the 20% of the data that you can't read anyway.'"

— Richard Breakiron, Commvault

About Commvault

Commvault is the gold standard in cyber resilience, helping more than 100,000 organizations uncover, take action on, and rapidly recover from cyberattacks — keeping data safe and businesses resilient and moving forward. Today, Commvault offers the only cyber resilience platform that combines the best data security and rapid recovery at enterprise scale across any workload, anywhere with advanced automation — at the lowest TCO.

Learn more about Commvault

Commvault® carahsoft.

Additional Resources

Below are additional materials to help you develop and implement a sound data strategy.

GovLoop Guides:

- Better Management for Data Maturity
- Leveling Up With Data

Articles and Reports:

- 5 Data Fundamentals Every Leader Should Consider
- The Transformative Power of Data and AI
- The AI Advantage
- The Government & Public Services AI Dossier

Glossaries:

- National Center for Data Services: Data Glossary
- Harvard University: Data Management Terminology

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

govloop

Thank You

Thank you to Commvault, ORI/Qlik and Pure Storage for their support of this valuable resource for public-sector professionals.

Authors

Candace Thorson, Managing Editor

John Monroe, Director of Content

Lauren Walker, Senior Staff Writer

Designer

Kaitlyn Baker, Senior Creative Manager

Kelly Boyer, Motion Graphics Team Lead