



public **cio**

Special Report

Has data
analytics
delivered on
its radical
promise?

Big Data ^{or} Big Hype?

NASCIO is the premier network and resource for state CIOs and an advocate for information technology at all levels of government. Visit NASCIO.org for resources that identify and promote leading practices and innovations that foster government excellence through quality business practices, information management and technology policy.

Publications

Stay informed on issues impacting state IT through NASCIO publications, issue briefs, surveys and toolkits.



Topics include:

- Cybersecurity
- Governance
- IT Workforce
- IT Accessibility
- Enterprise Architecture
- Data Management
- Procurement

Awards Library

Browse hundreds of innovative information technology initiatives submitted for the NASCIO State IT Recognition Awards. The Awards Library contains full nominations from 2001 to present searchable by year, category, state, recipient level or keyword. Look to fellow states for ideas, inspiration and best practices at NASCIO.org/Awards/SIT.



Mobile Apps Catalog

Is there an app for that? Visit NASCIO.org/apps to explore mobile apps created by state government to provide information and deliver services to constituents.



- Browse over 350 apps by state or category
- Download apps to test functionality and explore content and generate ideas

Have an app? Visit the Catalog to submit your app for inclusion.

Member Profiles

The CIOs from every state, the District of Columbia and two territories call NASCIO their professional home. NASCIO.org provides up-to-date contact and profile information and quick links to important state information at your fingertips. Additionally, you can also find contact information and quick links to our more than 100 corporate and associate members.



**Follow
NASCIO**



Table of Contents

4 Growth Spurt
A broad look at whether the practice of data analytics is on the right track.

6 Connecting Data with Policy
The tools and techniques of data analytics in the public sector.

10 Enter the Chief Data Officer
The emerging discipline shows signs of maturity.

14 Warehousing Data
A new approach to dealing with the deluge.

19 Survey Says
Can AI help tap the potential of analytics?

20 CDOs and CIOs: A Vital Relationship
How the two roles work together.

22 The Growth Challenge
Making smart decisions in light of increased demand.

28 Getting Results
How to put analytics to work.

If you live in New Orleans, the city will give you a free smoke alarm and install it as well. But in 2014, few residents took the time to schedule an alarm installation. Following a tragic fire in which five people died, including three children, in a home that lacked smoke alarms, the city's fire department decided to get more aggressive about installing the safety devices, but had little information on which houses among its 400,000 residents needed them.

The fire department turned to the city's analytics team and the Office of Performance and Accountability to find an answer. Taking U.S. Census

housing survey data and combining it with very granular data variables, such as a proportion of poverty level, age of structure and length of tenancy, the city was able to come up with a risk assessment for the likelihood of a fire fatality at the individual census block group level. The result: a map-based risk assessment of individual blocks in New Orleans.

Since 2015, the New Orleans Fire Department has installed thousands of fire alarms in the city's highest-risk structures. The effort was considered such a success that it went national. With more than 386,500 residential fires per year in the country and with

fire fatalities twice as high in homes without alarms, cities that use analytics to assess the risk of a fire fatality have increased the impact of fire alarm safety in a much more effective manner.

Other examples of predictive analytics have sprouted throughout the country. Cities can use data analytics to forecast when best to inspect a restaurant; county welfare agencies use predictive tools to assess whether a child under their care is at risk of maltreatment; and states are using analytics to spot the likelihood of fraud in unemployment claims and tax returns.

And it's not just predictive efforts. States and localities are using analytics

Growth

to figure out the most efficient routes for school buses; when the best time is to de-ice a road during a snowstorm; where the best opportunities are to reduce heating costs in government buildings. The list goes on.

The story of data analytics in state and local government is not new, but it has changed. The number of projects continues to grow. The number of analytics teams and chief data officers also continues to expand. The range of tools, both proprietary and open source, is also on the upswing.

But more doesn't always mean better. Questions have been raised about the effectiveness and impact

of analytics and whether its use is isolated to a few active jurisdictions. Should government focus its attention on open data or does it make sense to put resources into selective data sets for internal use? Then there are the questions about the best place for analytics teams to operate. Should they be part of the IT department, or do they work better reporting to the mayor or governor's office? And what about the technology that enables analytics? Is cloud computing the best infrastructure? Are open source tools as good as or better than proprietary tools? What about artificial intelligence? Could it make data analytics more

accessible to agencies and workers? Finally, there are rising concerns about the ethics of analytics: Are we in danger of running government based on algorithms? Will the privacy of citizens be compromised in the pursuit of better policies based on data analytics?

This report will try to get to the bottom of these questions as well as others by investigating the projects and processes underway around state and local government. It will include discussions with experts and practitioners to find out where data analytics is headed and the best practices that can help move it forward.

Spurt



Connecting Data

Consider the following:

- In 2016, there were an estimated **1.2 million violent crimes** committed nationally. The same year, **7.9 million property crimes** were reported to the FBI's Uniform Crime Reporting program by 17,000 law enforcement agencies nationwide. Also in 2016, federal and state prisons held **1.5 million prisoners**, according to the Bureau of Justice Statistics.
- In 2016, there were **64,000 drug overdose deaths**, according to the National Center for Health Statistics.
- Each year, state governments spend well over **\$500 billion** on social services, including welfare and Medicaid, according to the U.S. Census.
- Estimates vary widely, but spending on smart cities technologies conservatively could reach **\$23.5 billion in five years**.

At first glance, these facts appear to be random data points, but they all have one thing in common: They represent the four most important policy areas in state and local government, both in terms of overall costs and what drives investments in data analytics, according to a 2018 survey conducted by the Center for Digital Government (CDG is part of e.Republic, Government Technology's parent company).

Now consider this:

- **Chicago**, one of the first cities to test an algorithm that could predict crime, is in the process of opening more than a half-dozen mini fusion centers in neighborhoods where crime is high. The centers will be equipped with data analytics tools, as well as other technologies to proactively prevent crimes before they happen. Other cities, as diverse as Santa Cruz, Calif., and Eden Prairie,

Minn., are using predictive crime tools and techniques. Meanwhile, county jails are using data analytics to reduce prison populations and cut costs.

- The **Massachusetts** Department of Public Health uses analytics to sift through vast amounts of data to spot patterns in the use of opioids and to devise an early warning system about hot spots in the state where possible overdoses and deaths might occur.
- **Allegheny County, Pa.**, Department of Human Services has built a data warehouse and is using analytics to improve decision-making, including the ability to predict when a case worker should respond to possible cases of child abuse.
- A growing number of cities, such as **Dallas and Columbus, Ohio**, far outside the upper tier of coastal elite urban centers, have doubled down on using smart technologies and the Internet of Things (IoT), pairing them with analytics to improve everything from economic development to transportation.

ta with Policy

Like the private sector, the use of data analytics in state and local government continues to grow. But where the private sector is looking to find ways, directly or indirectly, to improve the bottom line, government is focused on making policies more effective while improving public-sector decision-making.

A few years ago, the focus was on testing the concept of using large amounts of data and examining them with either statistics or algorithms to see what answers could be found. Now, the process has become more formalized, more disciplined, with the focus on linking the results with policies.

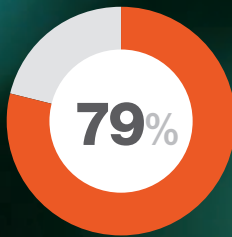
“What’s important to effective analytics is the connection to important policy questions,” said Jane Wiseman, the founder of the Institute for Excellence in Government and a senior fellow at the Ash Center for Democratic Governance and Innovation at Harvard Kennedy School. In the

past, governments have often had the desire to go after an interesting methodological process, rather than an important policy outcome. “But having a sophisticated algorithm is not what matters,” she said. “The question is whether you are plowing the streets more efficiently; are you making people’s lives better?”

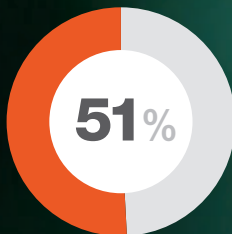
Across all sectors of government, data analytics is growing at a rate of more than 10 percent annually, according to IDC, the research firm. That’s certainly reflected in the Center for Digital Government’s survey on data analytics, which shows that state and local governments are investing in dashboards, data warehouses, data mining, predictive modeling



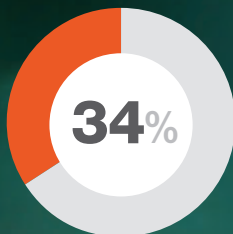
Data Analytics Technologies in Use



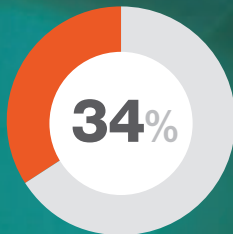
Dashboards/Portals



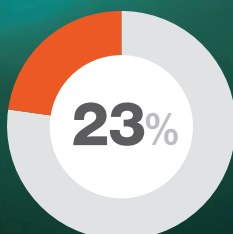
Enterprise Data Warehouse



Data Mining



Predictive Modeling



Master Data Management

“A lot of important things are not measured, or measured poorly with a lot of error.”

Lee Branstetter, professor, Carnegie Mellon University's Heinz College of Information Systems, Public Policy and Management

and master data management technologies to get the job done.

Defining good analytics isn't easy. There is a broad range of analytic efforts underway, using varying amounts of data to tell the story or to create the right model. But it starts with data.

“Government organizations need substantial amounts of data to improve processes,” said Lee Branstetter, director of the Center for the Future of Work and professor of economics and public policy at Carnegie Mellon University's Heinz College of Information Systems, Public Policy and Management.

“Analytics can be extremely sophisticated, involving artificial intelligence or machine learning algorithms, or it can be as simple as collecting information on some outcome, using statistical tools,” he said. “Given how expansive this is, it's not surprising how varied the analytics experience can be.”

Certainly, the scope and function of analytics varies broadly in state and local government. According to CDG's survey, 39 percent of respondents characterized it as a “guide to daily decision-making and problem-solving,” while 37 percent said the use of data analytics is widespread across different departments; another 36 percent reported having a dedicated employee or department for conducting analytics; less than 7

percent said their government entity is not using or does not plan to use analytics.

An in-depth research report produced by the Pew Charitable Trusts found that states have in recent years turned increasingly to administrative data to help make strategic data-informed decisions. Analytics, in particular, has become the foundation for uncovering important insights, revealing the root cause of a persistent issue, diagnosing breakdowns in systems, highlighting obstacles and predicting future phenomena, according to the report.

Those findings mirror, to some degree, what CDG has uncovered. Specific benefits of analytics are more concentrated, with 64 percent of respondents calling more informed decision-making the most important reason to use it, while just under 19 percent said it benefits citizen services, and 12 percent emphasized its ability to increase internal efficiencies.

While the scope, function and benefits of analytics in state and local government appear robust, the actual effectiveness of data analytics is another story. Only 14 percent of respondents to CDG's survey believe their government's analytics is highly effective, while 53 percent rate it as somewhat effective. More than 20 percent found analytics to be somewhat ineffective or were totally neutral on its impact.

CHECKLIST: BEST PRACTICES FOR BUILDING INTERNAL ANALYTIC CAPABILITY FOR HHS PROGRAMS

Changing policies, large caseloads and uncertain funding make it more important than ever for health and human services agencies to gain actionable intelligence from their data. Analytics and analytic services equip HHS agencies to flexibly and effectively serve clients despite the shifting landscape.

Analytics provide state and local government leaders with strategic insights from their data that identify cost, efficiency and quality improvement opportunities. Analytic services go a step further, helping states build the analytic capabilities they need quickly so they can proactively change outcomes and drive measurable results, including program and payment reform and improved health outcomes.

✓ Understand Your Existing Capacity and Capabilities

Asking and framing the right question is the single most important step in the analytics process. What are the critical business challenges that you need analytic services to solve? Where are the operational and capability gaps and areas for improvement within your organization — and are these gaps mainly technological or process and staff capacity related?

✓ Leverage a Partner

Short-term engagements with third-party vendors can help agencies launch analytic efforts, but these engagements work best when they're more of a partnership and less transactional. It's critical to work with a partner who understands your agency's specific needs and challenges and collaborates to help build your internal capacity and support your agency throughout your cultural shift.

✓ Monitor Resource Allocation

Sometimes, the answer isn't more money. Rather, it's shifting funds to areas where spending has the greatest impact. Analytics help organizations identify high-cost drivers, and assess whether they've made the best resource investments to proactively drive positive changes — whether it's directing more resources to preventive care programs that lower utilization, or launching pilot projects in high-risk pools that represent a small portion of the care population but the largest share of the costs.

✓ Develop a Sustainable Analytic Culture

Cultivate your analytic culture, specifically the technology, processes and people within your organization. Develop an analytic vision for your statewide programs that can become an integral part of your organization's mission. What are your goals? Is it building predictive models to measure and monitor progress, or is it using prescriptive analytics to make clinical and policy decisions? Also develop a process for data governance and continuous program improvement that involves input from key stakeholders and data modeling to enhance care quality and lower costs.



For more information on how we can help your state, visit optum.com/government. To download the full handbook, visit optum.com/hhsanalytics



“The office of the CDO has become more sophisticated, viewing themselves as a strategic asset.”

Jane Wiseman, founder, Institute for Excellence in Government



Enter the Chief Data Officer

To help states and localities make the best use of data analytics, they have created a position in government that, until recently, didn't exist. Chief data officers have become the focal point in government's embrace of all things data. Less than three years ago, just a handful of cities had CDOs. Now, there are at least 20 in city government; and 18 states and the District of Columbia have appointed a chief data officer or similar position, according to the Pew Charitable Trusts.

Their growth reflects government's shifting attitude toward analytics. "A few years ago, the CDO was rare and in the early stage of development," said Wiseman. "Now, the office of the CDO has become more sophisticated, viewing themselves as a strategic asset; they have become

more proactive than reactive when it comes to data projects."

As a sign of their growth, Harvard's Ash Center has set up the Civic Analytics Network for city CDOs where they can collaborate on shared projects that advance the use of data visualization and predictive analytics to solve important urban problems, as well as to discuss issues around data governance and infrastructure.

Still, state and local governments have been slow to embrace the new role overall. According to CDG's survey, only 32 percent of the respondents reported having a CDO or equivalent. Those jurisdictions that have a CDO say that just over 50 percent report to the chief information officer. But

DAVID KIDD



**More than
90 percent
of survey
respondents
say chief data
officers are very
or somewhat
influential
when it comes
to advancing
data policies.**

DAVID KIDD

what stands out is that more than 90 percent of the respondents say CDOs are very or somewhat influential in advancing data policies. That's a lot of influence in a very short period of time.

One of the early pioneers in bringing predictive analytics to the forefront of government operations is Tom Schenk Jr., Chicago's chief data officer since 2014. He put Chicago's analytics on the national map with the development of a predictive model for restaurant inspections, which has been adopted by several other cities. Since then, Schenk and his team of analysts have created models for predicting E.coli levels in Lake Michigan and West Nile virus outbreaks, and now have a project underway that will predict the risk of lead poisoning in low-income homes where children are present.

But Schenk is perhaps best known for accelerating the impact of open data. In 2012, Mayor Rahm Emanuel signed an executive order requiring the city to have an open data portal. Initially, the goal was to put as many open data sets as

possible on the portal, according to Schenk. But when he took over as chief data officer, Schenk sought to impose some regularity on a project that made hundreds of data sets available to the public in a few short years. "I've been working on improving our workflow and management around it, while allowing the program to continue to grow," he said.

As it continues to spread, open data can have a significant impact, especially at the local level. "Open data is the oxygen of analytics," said Wiseman. "Remember, analytics requires large volumes of good, quality data. When the public can look at the data, spot what's missing and push for more of it, government benefits."

Still, state and local governments remain somewhat conflicted around open data. According to the CDG survey, 21.5 percent of respondents believe open data is a foundational principle, and that all data that can be open should be available to the public. An equal number report that open data is optional, not mandatory in

their organization, while another 15.5 percent say their government doesn't have an open data initiative in place.

Another city that has been aggressive about open data is Pittsburgh. Under the direction of Mayor William Peduto, the city has made strides in offering up its data for public consumption. It has partnered with Allegheny County and the University of Pittsburgh to launch the Western Pennsylvania Regional Data Center, what some consider to be the country's first regional open data platform. The city has a Department of Innovation and Performance, and last year it released a progress report that showed how open data has been used to confront the opioid addiction problem in the area, improve street safety for bicyclists and pedestrians, help convert vacant lots into community spaces and help decision-makers address inequality problems in the region.

Unleashing the Power of Data

Advanced analytics can only yield new insights with the right approach to enterprise data management.

For almost two decades government has pursued the goal of effective data analytics, but has achieved mixed results. Data is an asset that gives agencies the potential to better protect citizens, improve policies, implement new public services and more. The end result? Citizens have greater faith in government. By taking better advantage of data, the public sector can significantly improve the delivery of repeatable and consistent outcomes.

With the expanding role of data, some state and local governments are adding new data-centric positions, such as chief data officers, and find themselves needing next-generation enterprise data management platforms. These platforms can accelerate data discovery, streamline data onboarding, and balance self-service data preparation with data governance, enabling agencies to improve mission delivery.

For organizations seeking to transform data into trusted business assets, new technologies such as data lakes, machine learning and artificial intelligence (AI) offer innovative means for generating trusted, timely and accurate information.

Informatica, the leader in Enterprise Cloud Data Management, accelerates data-driven digital transformation by unleashing the power of data and turning more sources of data into trusted business assets. With Informatica's unique AI-driven approach to data management, agencies can quickly and easily catalog, ingest, integrate, prepare, enrich, relate, secure and govern data — in the cloud or on-premises.

There has never been a more critical time to invest in enterprise data management. Without a strategic approach to data management, data becomes increasingly difficult to find, master, govern and protect, which can have profound negative consequences.

Informatica's solution paired with AI capabilities enables government leaders to keep pace with the ever-changing data landscape, while also capturing and processing information in a way that yields results consistently. At the same time, agencies can ensure they are complying with data governance and security mandates. As AI becomes an increasingly mandatory technology, Informatica's solutions will enable public sector organizations to better interpret data for both analytic and predictive purposes, ultimately leading to better policies and better social outcomes.

Why data management?

- Taking data out of silos makes it possible to discover meaningful relationships in information across agencies
- Self-service preparation tools enable analysts to rapidly turn data into insights
- Automated processes and procedures drive sustainable operations

Learn more at [informatica.com](https://www.informatica.com)



Informatica™

Warehous Data



ing

Michael Schnuerle has been chief data officer of Louisville, Ky., for just over one year. He is in the Office of Civic Innovation and leads a data governance group of approximately 40 people spread across multiple departments. But besides him, there's just one other person in the city with a data science background whom he works with. "We don't have a staff of full-time analysts; we're not there yet," he said.

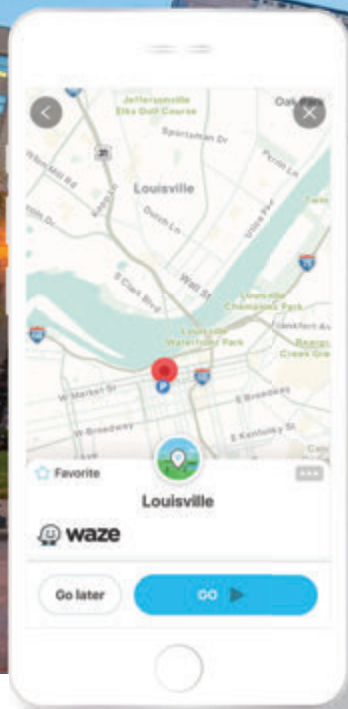
What Schnuerle's data team lacks in resources is made up for by the scope of work underway. For the data governance group, he runs an academy that helps with training with the newest tools, as well as coaching on best practices. The city has a robust open data program too, and it has led to some mutually beneficial data sharing that serves the need to improve mobility and street safety in Louisville.

"We publish road closure data on the open data portal, which Waze [the road

navigation app] can consume. They add the data to their app, allowing drivers to avoid closures," said Schnuerle. In return, the city gets a "firehose" of data every two minutes from Waze, which includes valuable information on traffic jams, potholes and other vehicle speed-related data for the entire city. The Waze data is correlated with the city's traffic sensor data, which can be used to determine the impact of a traffic signal or lane change. "We can look at the Waze data before and after we make a change and analyze how the change improves how people move through a specific corridor, which can reduce accidents," he said.

The data collaboration with Waze is just part of a much more ambitious data analytics initiative that will drive one of Louisville's bigger policy efforts: better mobility. For two years, the city's Office of Civic Innovation has worked with the Traffic Department to bring together data that will be





Traffic data from Waze is part of the “firehose” of data Louisville wants to dump into a data warehouse. Advanced analytics could then help the city make the most of limited transportation funds.

generated by connected vehicles and smart streetlights, for example.

“A big initiative is to create a traffic data warehouse, where we can dump disparate data from the city and external vendors and partners, regional and state transportation agencies, to do advanced analysis on how best to use our limited transportation funds to optimize everything from intersections to major transportation corridors,” said Schnuerle. “The data will help us figure out what devices we need, then to connect those devices through our city-owned broadband network, which will generate as well as feed data into the warehouse.”

Louisville is one of several governments that see data warehouses as the platform on which data analytics sits and can act as a secure and effective environment for running data-driven policies. Another example is Allegheny County’s Department of Human Services, which uses a data warehouse for its various health and welfare analytics efforts.

Boston is also in the process of building a data warehouse to support analytics, but on a much grander scale. There, the plan is to create a warehouse platform that will support self-service

analytics for all city departments, as well as for the city’s analytics team.

The person behind this project is Andrew Therriault, the city’s CDO since 2016. He oversees what may be the largest analytics team in local government, with nearly 20 staff and external partners conducting work. Despite its size, the number of analysts isn’t enough to handle the growing demand from city departments, according to Therriault. “Long term, it’s clear that we don’t see ourselves doing all the analytics. I would need at least 10 times the people to do all the work,” he said.

That’s where the data warehouse initiative comes in. Rather than keep adding more analysts, Therriault wants to make city departments more self-sufficient. “It’s best to have people in the departments understand the data and the processes rather than have a huge analytics team handling all of the work for them,” said Therriault.

Putting a city’s data in one place is a unique proposition. But with the growth in IoT and big data, cities find themselves with more data than their existing infrastructure of servers is made to handle. “We now have terabytes of data,” said Therriault, a volume that was unheard of just a few years ago and is complicating how analytics can get done. “This makes it hard to have a holistic plan for how we collect, organize and manage our data going forward,” he said.

Therriault leads a data infrastructure working group within the Civic Analytics Network, which is exploring what technology platform and database solutions are best for a city’s unique environment. For Boston, the focus is on building databases that let the analysts go after very broad questions. For example, Mayor Marty Walsh asked Therriault’s team about the number and type of small businesses in a certain neighborhood. Despite the simplicity of the question, such a database of information didn’t exist. “There were different types of information, such as business licenses, but the data was incomplete,” said Therriault.

“We now have terabytes of data. This makes it hard to have a holistic plan for how we collect, organize and manage our data going forward.”

Andrew Therriault, Boston CDO

Accelerate Data Projects by **More Than 70 Percent** with NetApp and SAP

State and local governments need to work faster and smarter to expedite data projects, reduce costs and unburden IT teams. One of the most resource-intensive tasks for IT teams is supporting big data workloads and real-time analytics. NetApp and SAP solutions deliver uninterrupted real-time data access, multitenancy and accelerated project ramp-up to propel your organization forward and deliver quicker data intelligence.

NetApp's unique data management features and tight integration into SAP's application toolset — including SAP LaMa and HANA studio — allow organizations to streamline and automate SAP system copying in all project phases, from design and testing to production and data backup.

Prototyping and project deployment — Create SAP system clones in less than 10 minutes, while using less than 10 percent additional disk space. Because there is no impact on the production environment, organizations can create system copies anytime — meaning no more work at nights or on weekends.

Data migration and cloud integration — Move SAP data and systems to an on-premises, cloud or hybrid environment that makes the most operational and financial sense for any given scenario — quickly, efficiently and without disrupting production.

Backup and disaster recovery — Protect data in seconds by using NetApp to seamlessly replicate data and systems to offsite backup locations or cloud providers. Those same backups can then be leveraged for system copy functions regardless if it's an offsite backup location or a cloud provider.

Testing and training — Take advantage of rapid provisioning and fast refreshes to set up and maintain testing and training environments more efficiently. Actively use disaster recovery images or offsite backups for testing and training — at no extra effort, cost or impact.

Advanced operations — Take SAP environments to the next level by using NetApp to enable self-service provisioning, one-click operations, sophisticated automation and other advanced services.

NetApp is a founding member of the SAP Co-Innovation Lab (COIL). With more than 20 years' experience in running SAP solutions on its platform, NetApp provides the foundation that state and local governments need to deliver SAP projects faster, more efficiently and with less risk. Leverage NetApp and SAP and learn how to accelerate your applications and project delivery and unlock the insights in your data.



Boston has what may be the largest analytics team in local government, but it's not enough to keep pace with demand, leading the city to propose an enterprise data warehouse for analytics.

Once an enterprise data warehouse is created with the right kind of data structures in place, it becomes easier to perform both simple and complex analytics. It will also allow city departments to create their own dashboards using much richer sets of data from various sources that can be shared — and that have been properly prepared by the city's analytics team.

Boston's proposal to build an enterprise data warehouse for analytics is unique. More commonly, warehouses can be found at the state level that support a specific type of analytics. For example, many states have modernized their unemployment insurance (UI) systems, letting workers file for claims online, but also increasing the opportunity to go after fraud and other types of benefit leakage problems, like improper overpayments.

Utah's Department of Workforce Services modernized its UI technology in 2006. The new system included a data warehouse that has allowed the agency to amass data to monitor and evaluate possible fraudulent activity. In 2015, the state paid out \$200 million in unemployment benefits, with a fraud rate of 1.3 percent, well below the national average.

Medicaid is another major part of state government that has benefited from the use of a warehouse and analytics. Indiana's Family and Social Services Administration, which oversees the state's Medicaid program, has used its data warehouse to analyze health-care information and to help identify more than \$150 million in potential savings. For example, the state looked at anonymized and aggregated data and uncovered

evidence that certain emergency room visits could have been better and more cheaply handled through a primary-care visit.

Historically, state and local governments have built data warehouses on the premises. But the trend now is to put them in the cloud. Boston, for example, is looking at a variety of data warehouse options, but a cloud solution is at the forefront, according to Therriault. Whether to build a data warehouse in the cloud or on premises often comes down to policy. Many state and local governments have regulations or laws that require data to reside within a certain geographic area, to avoid having data stored in servers that may reside in another country.

One factor in favor of using the cloud is time: Setting up a data warehouse can happen much faster in the cloud than building one from scratch in a government data center. Some might also argue that cost is another advantage of the cloud. However, costs for a cloud solution can mount, depending on the amount of data stored in the cloud, the type of performance needed, such as hot or fast access versus cold or sporadic access, and what type of software functionality is required. These kinds of cost variables have made data warehousing as a service somewhat unpredictable and subject to inflation as data storage and usage mounts.

When it comes to data warehouse software functionality, CDOs appear to be agnostic about using proprietary or open source, like Apache or Hadoop. For data science and analytics, government analytic teams like to use rich and versatile programming tools, such as Python and R, as well as simpler software, like Excel. Data visualization tools, both proprietary and open source, are also widely used.

Survey Says

Even though **data analytics** remains an emerging field, there's already interest in adding a new component: artificial intelligence. The technology has moved quickly from the fringes into the mainstream of government with new applications popping up nearly every week.

But can AI have an impact on analytics and, more importantly, how could it benefit the practice? The Center for Digital Government asked this question, and government executives responded in the following way:

48%
believe AI will help generate predictive analytic reports that aid work activities.

47%
think AI can generate real-time analytic reports for decision-makers.

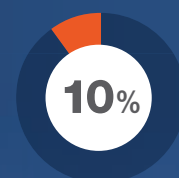
46%
say AI can identify trends, opportunities and rules that would be too time-consuming for employees to identify on the job.

41%
believe AI will automate tasks that would normally be done by an employee who is not a data scientist or analyst.

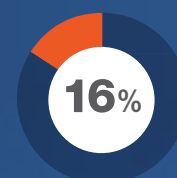
40%
say AI will improve citizen-facing services.

25%
think AI will automate tasks that would normally be done by a full-time dedicated data analytics employee.

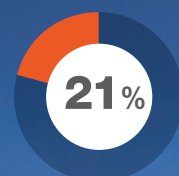
But when it comes to putting these ideas into action, state and local governments have barely begun to use AI and analytics:



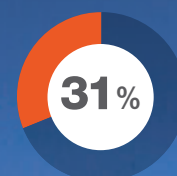
currently use AI



plan to use AI



have discussed using AI, but have no plans



have not discussed using AI





CDOs & CIOs: A Vi

Creating the right kind

of infrastructure for data, software development and analytic tools requires the understanding and support of the government IT department. At first, it would seem natural that a CDO report to the CIO, to ensure success for any analytics program. But according to the Center for Digital Government survey, 46 percent of respondents said the CDO does not report to the CIO.

"I'm agnostic whether a CDO reports to a CIO," said Jane Wiseman. "But the CIO needs to be a champion because they own all those giant

enterprise systems that generate huge amounts of administrative data." More precisely, the CIO can lay the foundation for a great data analytics program by helping to ensure the quality and availability, as well as transparency and accessibility, of the data the CDO needs.

That describes the symbiotic relationship Maksim Pecherskiy, chief data officer for the city of San Diego, has with his CIO. Pecherskiy reports to the mayor, but his relationship with the city's CIO is both close and critical. "We are mutually dependent on each

other," he said. Using a metaphor, he explained the relationship this way: "IT's role is to make sure the infrastructure is in peak form — the lights are on — whereas we want to know where the lights are located, how many of them there are and how much energy they are consuming," he said. "I can't do that if the lights are off."

In San Francisco, Joy Bonaguro, the city's chief data officer, also reports to the mayor's office. She has a team of five analysts who operate like a small startup within the city, helping departments choose the right analytics



tal Relationship

project, validating that it's addressing the right kind of question, making sure it aligns with the mayor's vision, and then working closely with the staff to implement the data project and to ensure it is sustainable.

Because city government in San Francisco is highly federated, it's not unusual for key agencies to have their own IT department, while the city also has a central IT office run by CIO Linda Gerull. Given that organizational structure, Bonaguro works with what she calls "institutional partners" to address three core data

challenges: knowledge of and access to data, governance, and data usage.

The city's CIO aids the analytics team with addressing its first challenge: accessing data. "We have a program around data services and architecture. Within that, we manage the city's open data program and set strategies around data infrastructure and investments," Bonaguro said. "For that to work the city CIO is a key institutional partner."

In Louisville, CDO Michael Schnuerle said that working directly with the city's CIO was critical to his efforts to put analytics into the

cloud. Like most cities that venture down this path, some of the data and infrastructure will remain on premise on servers, while other parts will be available as a service. "But what we are missing is that interoperability between our network and the cloud," he said. It's a problem other cities are struggling with, and it's one in which the support of the CIO is badly needed. Schnuerle called the effort "transformational" and one that requires the direct support of the city's CIO.



The Growth Challenge

The position of chief data officer

is Maksim Pecherskiy's first job in government. Working in the San Diego Performance and Analytics Department, under Mayor Kevin Faulconer, has exposed Pecherskiy to the kinds of complex policy problems that cities have to navigate in order to run things faster, better and cheaper. Prior to his current position, Pecherskiy was a Code for America fellow in Puerto Rico and has worked on building large Web application projects for international organizations. But like many CDOs working in an environment that is still somewhat new and growing fast, Pecherskiy has been operating with few guideposts on how to run an analytics operation within a large city.

The mission of his analytics team is to focus on three objectives: open data, data delivery and analytics. So far, they have worked on how to automate the extraction, transformation and loading of data from various sources into the city's open data portal. As for data delivery, Pecherskiy said, "We help people discover data and get it;

we also try to make it easier to use and deliver it to them, like pizza."

More seriously, Pecherskiy has been challenged with launching and sustaining analytics as the third of his objectives. The reason is that early efforts to provide analytics meant trying to avoid some of the city's bureaucracy, but that led to problems. Now, there is an agreement that details what everybody is responsible for, the goals and the timeline. Pecherskiy has the right to cancel or re-prioritize a project if the client isn't meeting their responsibilities. "It's important that everybody have some skin in the game," he said.

Bringing cohesion and structure to government analytics has become important as the practice grows and demand for services often exceeds what the teams can deliver. Every chief data officer interviewed for this report mentioned the surge in demand from departments as the benefits of data analytics to improve performance, predict outcomes and improve decision-making become clearer every day.

"IT's role is to make sure the infrastructure is in peak form — the lights are on — whereas we want to know where the lights are located, how many of them there are and how much energy they are consuming."

Maksim Pecherskiy, CDO, San Diego

Interestingly, in San Diego, Pecherskiy's team will act as a consulting firm for an agency's analytics work that it can't handle on its own. "We'll act as a consultant for the agency, sit in with them and the vendor, to make sure they are doing the right thing," he said.

While growth is good, it has forced CDOs and their teams to become more critical about the types of projects they take on, enforcing the criteria that is needed before a full-scale analytics effort gets underway. In San Francisco, the analytics team has set up a solicitation process that includes using specific language to gauge whether or not a city agency has a question that data science can solve, or whether it might be a performance management effort, or just might need a dashboard for better data visualization.

"We look for a variety of projects in terms of size and scope, but all with value," said Bonaguro of San Francisco. Once a project is deemed worthy, the next step is to evaluate its scope and, more importantly, whether the answers to the data science question will result in a service

change that can be implemented in a way that makes it sustainable.

Examples include a dashboard for the Department of Environment to identify new clients for a light subsidy program the city runs. The dashboard gave the department a tool to filter and find the right clients, with the outcome consisting of a pipeline of new clients for the program. In another project, analytics helped San Francisco's Women, Infants and Children program find out why their client base was falling. "We did modeling that identified key issues that were driving up dropouts," said Bonaguro. "The program is using those findings to change their client experience."

But data analytics programs face far more fundamental challenges than having a process in place that ensures the right kind of data science project gets implemented. Data quality remains a huge issue, including the collection of data itself. "A lot of important things are not measured, or measured poorly with a lot of error," said Branstetter. Cities and states need to invest in high-quality data if they hope to take advantage

of what analytics can do, he added. "That means getting the public to understand the need for government to invest in quality data collection."

Schnuerle said he struggles with data quality issues. A good example is that Louisville is just one of nearly 80 cities and towns in the metro area. Trying to find consistent, quality data around, say, car collisions, can be a battle, with small police departments collecting and organizing data in different ways. "If we want to get the big picture on car collisions, we pull data from the state police," he said. But even here, there can be problems with the accuracy of a geo-coded collision provided by state police. "We end up having to do a lot of data cleaning."

Another challenge is what Branstetter calls the "statistical understanding" barrier. "Executive decision-makers know the problem they are trying to solve, but don't understand what statistical technique or combination of data and statistics can help them," he said. "They don't know where to start."

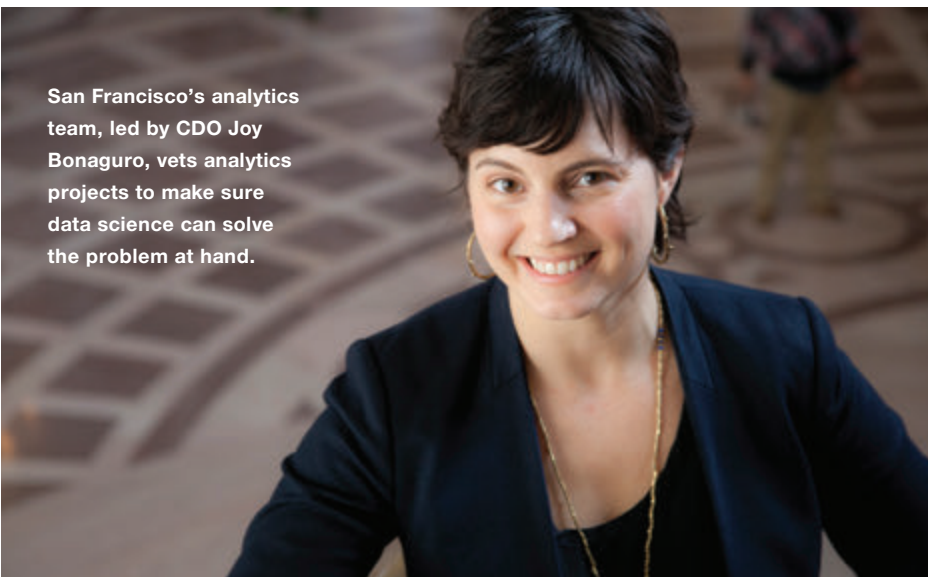
That's a concern because statisticians are used to getting a fairly well-defined problem to tackle, but they don't necessarily understand the city's problem. "So, it's a lack of statistical expertise within city government on the one hand, and a lack of familiarity among statisticians of public-sector problems," he said.

CDOs confirmed that finding talent with an understanding of public policy issues is an acute problem. CDG's survey found that lack of qualified staff was the biggest impediment to incorporating big data and practicing analytics in government, followed by lack of funding.

Someone who is at the center of this problem is Alex Engler, program director for computational analysis and public policy at the University of Chicago. Engler recruits and trains promising

San Francisco's analytics team, led by CDO Joy Bonaguro, vets analytics projects to make sure data science can solve the problem at hand.

JESSICA MULHOLLAND



DEMYSTIFYING EMERGING TECHNOLOGY TO MOVE BEYOND THE HYPE

Artificial intelligence (AI), digital twins and bots of every kind have exploded into the discourse of many government technology leaders. It can be enjoyable to discuss the possible benefits of emerging innovations, but how does an organization discern which technologies are more than just buzz? SHI's breadth of resources help IT leaders move beyond the hype and make sure they get the most value out of modernizations.

PROMOTE EMERGING TECHNOLOGY AWARENESS

It's all too easy to get caught up in the excitement around a particular emerging technology — it's more difficult to understand the advantages it could bring your organization. For example, what areas of AI might yield the greatest returns? Are chatbots right for your organization? Will digital twins allow you to be more effective and efficient? The volume of new technology is growing daily and is ever-changing. Most of these technologies rely on a foundation of data currently residing with organizations.

Public sector organizations should formalize a process to promote emerging technology awareness and adoption. As

part of this strategy, consider leveraging SHI resources. SHI employs experts who not only understand emerging technologies, but who can also anticipate the challenges public entities will face when they introduce potentially disruptive changes into their environment.

DON'T TRUST THE HYPE, PROVE THE VALUE

After identifying data and emerging technology uses that could be beneficial to your organization, your staff will want to work to prove the value. But running proofs of concept can be time-consuming and costly. SHI's Customer Innovation Center (CIC) allows organizations to test emerging technologies from multiple vendors in a secure environment.

It's necessary to identify specific innovations that will provide the greatest return to your organization. By proving the value ahead of procurement and implementation you can ensure the value proposition for your organization isn't just hype. SHI can help demystify emerging technology and identify potential value for your organization through our broad expertise and partner relationships.



For more information, visit: www.shi.com

public policy students in the arcane world of data science and analytics. But there simply are not enough students interested in a career that brings together data science and public policy. “The number of computer science graduates who go into public policy is unbelievably low,” he told *Government Technology* in 2017. “The public sector is about 30 percent of the economy, yet less than 10 percent of computer science grads go into public service.”

There are no simple answers to this problem. Cities (and counties) lucky to be near a university that practices data science and public policy can often tap into these programs by offering students challenging projects to work on. For example, Allegheny County and Pittsburgh both take advantage of the student talent at Carnegie Mellon University’s Heinz College of Information Systems, Public Policy and Management. The city of Chicago can tap into Engler’s program at the University of Chicago. Los Angeles has UCLA and USC in its backyard.

But what about smaller jurisdictions? Michael Schnuerle knows he’s unlikely to see an increase in funding to cover the costs for the

number of analysts he’d like to have on staff. So the options come down to keeping himself informed and up to date through networks and outside collaboration. “It’s important for CDOs like me to network with other peers and organizations like What Works Cities and the Sunlight Foundation to expand what they are able to do with limited budgets and limited time,” he said.

Schnuerle has also turned this idea into action. Recently he launched an initiative called the Open Government Coalition, which helps cities collaborate on projects without having to get involved with vendors or spend funds on expensive technology. One of the first projects that the partnership has tackled is formalizing how cities can collaborate with Waze and its rich traffic data, as Louisville has done. So far, 19 cities are on board and five are contributing code that can easily process Waze data in the cloud. “It came together quickly,” said Schnuerle. “It’s a good model and network to use on other projects as well.”

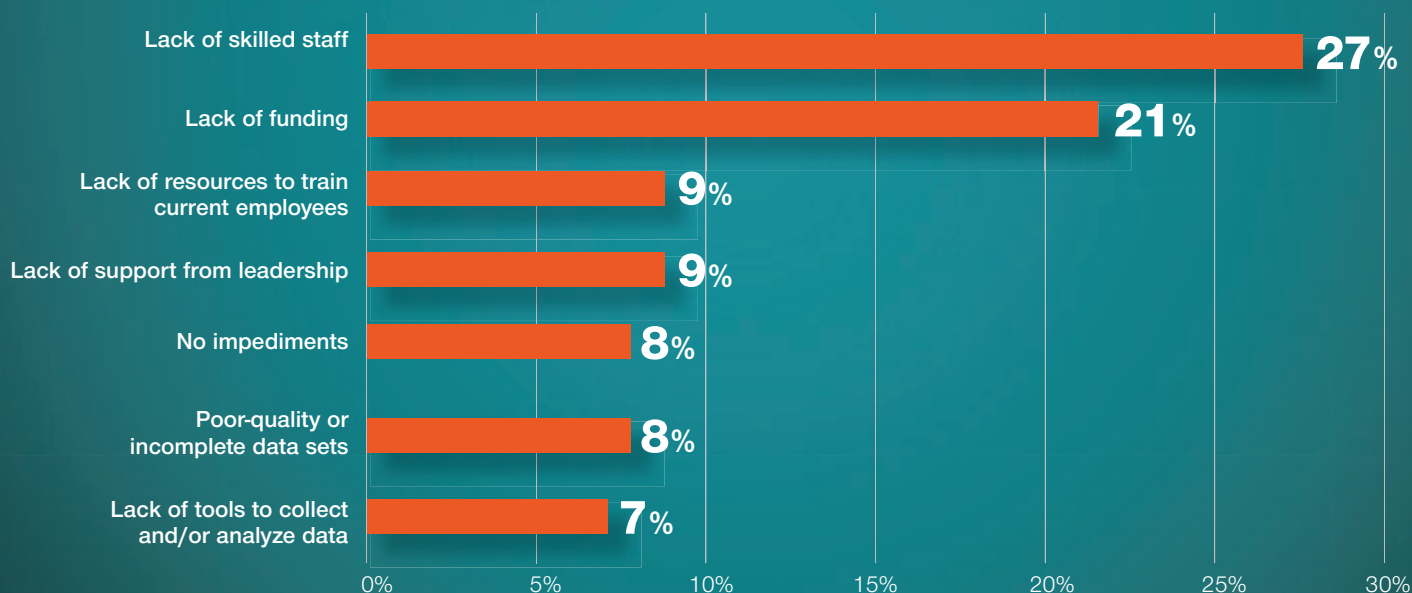
Jurisdictions also must confront issues around technology, such as getting permission to use certain kinds of software tools that might not be

sanctioned by the IT department. But the other major challenge for CDOs and data analytics is the issue of ethics. Good analytics projects often sit on top of vast amounts of data, and the more granular and varied it is, the better the outcomes. But in government, that can be a problem.

“Data analytics isn’t just a project between a city department, the chief data officer, and his or her analysts,” said Branstetter. “This is about data, some of which is coming from citizens, some of it may be sensitive and personal.”

Facebook’s lack of an ethical playbook, which it only recently began to address, has exposed the kind of ethical problems around data that have been cropping up more regularly in the private sector. “But government is supposed to be accountable when it comes to how it uses different kinds of data,” said Branstetter. “As we collect more data and do more with it, and as government makes more decisions to do more with it, there needs to be a process of consultation and education, a democratic process that brings people along.”

Biggest Impediments to Using Data Analytics





A Data Platform for the Cloud Era

Data is everywhere today, and organizations have much to gain by putting that data to work. But legacy mainframe and client/server era technology and processes simply can't cope with the explosive growth in data and the need for real-time analysis and agility. Simply put, today's cloud model of computing demands a new model of infrastructure.

Pure Storage's Data Platform for the Cloud Era was built to bridge the gap between traditional and new stack modes of IT and agile application development. We put data to work with a platform that supports today's volume, velocity and variety of data. Our platform enables governments to accelerate new data-driven applications, derive valuable insights, boost agility and simplify hybrid cloud development. All of this leads to substantially reduced costs and complexity.

The Pure Storage Data Platform for the Cloud Era is:

- EFFORTLESS:** Reliable storage that is always on, always fast and always secure.
- EFFICIENT:** Consolidate all data, whether block or file, structured or unstructured, into efficient all-flash storage that takes 10X less space, power and cooling.
- EVERGREEN:** Buy storage once and upgrade it as needed, without any disruption. Our Evergreen Storage Subscriptions mean governments can take advantage of software, hardware and flash innovation with no need to wait for a refresh. This can result in a six-year total cost of ownership that is about half the cost vs. competition.



For more information, visit WWW.PURESTORAGE.COM/CLOUD



Getting Results

Boston, like a number of other top-tier cities, has seen its population start to grow again after decades of decline. While growth is good, it has also brought more traffic to a city that has a very dense network of streets. That became a particular problem for the city's emergency medical services (EMS) drivers. Ambulance response times have been increasing, while the size of the EMS fleet stayed the same. For help, EMS turned to Boston's analytics team for some answers.

Looking at the data, analysts realized certain kinds of emergencies related to substance abuse, mental health and homelessness were taking far more time to resolve than others. These people needed much more help, but of a different kind. So, EMS developed a new program

that wrapped social services around emergency management. And rather than send a fully staffed ambulance, the city began dispatching smaller vehicles with a staff person trained to deal with the intricate, long-term care issues that this vulnerable population needed the most.

Boston's data-driven response to an emerging challenge showcases the impact that data analytics can have in government. It can help decision-makers solve complex problems that cut across a number of government programs and policies. But it takes resources — individuals trained in analytics — to figure out where the answers lie and a well-vetted process to help implement the needed changes.

Just as importantly, data analytics needs a flexible infrastructure, either

in the cloud or within a dynamic data center, that can collect, store and manage the enormous amounts of data needed to advance analytics, as well as the right kinds of software tools that can provide meaningful information. Whether or not a chief data officer reports to the CIO, the IT leader needs to be a champion that supports the endeavors of the analytics team.

Last, government executives — mayors, county commissioners, governors — need to understand that data analytics requires an investment of people, time and money in order to have a sizable and sustainable impact on problem-solving, better service delivery, improved resource management and increased effectiveness.

ANALYTICS ECOSYSTEM

TERADATA CONTEST SHOWS THE POWER OF SYNERGY

A key premise in both business and government is that if you give people real data and ask them to solve a real problem, you'll get an effective, real-world solution.

Teradata put this idea to the test last fall in a chain of events that brought together civic decision-makers, student participants and federal open data providers to address challenges around transportation and other civic issues.

The outcome offers an object lesson in the art of the possible. A design-athon featuring executives from Ohio agencies came up with a list of challenges. Students showed that, armed with good data and a rich analytic ecosystem, they could deliver practical solutions.

A team of Ohio State students used data analytics and visualization to develop a "worker commutability" score as a new way to reframe mobility choices in Columbus, Ohio. Their special achievement lay in bringing together disparate data sets in the service of a practical problem.

Collaborating with the Census department, students tapped richly detailed data to find out where people live and work. They plugged into private sector data to calculate travel times and drew from cell phone big data to understand travel patterns. Taken separately, each data source adds to the city's store of knowledge. Combined through the power of an analytic ecosystem, the data offered new and actionable insights.

LESSONS LEARNED

Teradata draws three key insights from this round of development events.

- ✓ **SYNERGY COUNTS.** Because the events brought together public and private sector players, competitors were able to tie technical solutions to real-world problems.
- ✓ **DATA IS OUT THERE.** Organizers were surprised to discover the depth and breadth of publicly available open data. Census data in particular was critical to this civic improvement, and yet that is just the type of data that typically goes underutilized.
- ✓ **AN ANALYTICS ECOSYSTEM IS KEY.** The winning team needed to draw information out from its customary silos and combine it in new ways. Teradata's analytic environment creates a common ground on which data culled from wide-ranging sources can be compared and combined, unlocking new analytic capacity and creative potential.

When different data sets are brought together, civic leaders are provided new measures, visualizations or ways to recognize underlying patterns. It creates a compelling means to deliver data-driven civic improvements.



Find out how we can help your organization become safer, smarter and more engaged at: www.teradata.com/government

Acknowledgments



The Center for Digital Government, a division of e.Republic, is a national research and advisory institute on information technology policies and best practices in state and local government. The Center conducts e.Republic's annual Digital Cities and Counties Surveys; the biennial Digital States Survey; and a wide range of custom research projects. For more information, contact Executive Director Teri Takai at ttakai@erepublic.com.
www.centerdigitalgov.com



Public CIO, a division of e.Republic, is an award-winning platform dedicated to technology thought leadership in federal, state and local government. Through print, online and a portfolio of events, Public CIO provides CIOs and key enterprise leaders with career critical insights on leading and navigating the innovative trends creating efficiencies, driving collaboration and impacting government services.
www.public-cio.com



Tod Newcombe

With more than 20 years of experience covering state and local government, Tod previously was the editor of *Public CIO*, e.Republic's award-winning publication for information technology executives in the public sector. He is now a senior editor for *Government Technology*. Tod is also the author of several books on information management.

Sponsors





GOVERNMENT TECHNOLOGY'S

public cio

YOUR EXECUTIVE UPDATE

PUBLIC CIO WEEKLY

News, commentary and
thought leadership for
the C-level IT community.

SIGN UP TODAY AT

www.govtech.com/newsletters

© 2018 e.Republic. All Rights Reserved
100 Blue Ravine Road, Folsom, CA 95630
916.932.1300 phone | 916.932.1470 fax

