KPMG

Better data, better government

Effective use of data and analytics can help governments at all levels deliver improved citizen services and outcomes.

Contents

About the authors

Sid Frank

Sid is a director in KPMG LLP's (KPMG) Information Management, State and Local practice and has been in the data management, data analytics, and business intelligence space for more than 25 years. He has developed and managed data-related solutions for multiple clients in multiple sectors, including federal, state; and local governments; healthcare; financial services; transportation; high-tech manufacturing; and telecommunications.

He is currently leading the independent verification and validation for two data warehouses and reporting systems for a large state. Also, he assisted another large state in developing data analytics processes for voluntary taxpayer compliance, led the development of a capacity planning and forecasting model for foster care children, and program managed the development of a healthcare data and analytics program to improve patient outcomes. He has published numerous articles on data quality and data governance, is project management professional (PMP) certified, and has a MS degree in system engineering degree from the University of Pennsylvania and an executive MBA degree from Temple University.

Traci Gusher

Traci is a managing director and a leader in KPMG's Center of Excellence (CoE) in Big Data and Advanced Analytics. She has 15 years of experience in professional services. Traci is the U.S. firm leader of KPMG's Data and Analytics Strategic Transformation Offering, which is designed to assist clients in establishing sustainable and efficient Data and Analytics capabilities. Traci helps clients articulate their business opportunities and challenges, understand the implications of using data and analytics as part of the approach, and develop a plan to realize benefits. Traci has extensive experience leading workshops and design efforts to assist and accelerate translating business issues and opportunities into data-driven insights and actionable methodologies.

Contributor

Viral Chawda
Viral is a managing director in KPMG's Data and Analytics CoE. He is a recognized leader in the art and science of applying analytics to deliver business value. Over the last 19 years, he has helped large commercial and government organizations with optimization and monetization of their data assets. He has lead experienced, analytics-based client delivery teams. He is recognized as an architect of the strategy for new decision science approaches across a broad range of applications including finance, marketing, risk, operations, portfolio management, and innovation.

He is a creative and structured problem-solver who understands how to leverage the ecosystem of alternative analytic methods and data sources as well as effectively drive change throughout the organization to deliver value. He is a published author and speaker at industry forums and events. Viral holds an MBA degree from Washington University and an MS degree in computer science/data mining from the University of Nebraska.

The case for improved data and analytics

Governments accumulate a wealth of information from the citizens and businesses within their borders. Through their public agencies, states and municipalities collect data of all kinds—economic, demographic, education, health, and more.

As people continue to interact online, with their mobile devices and through other interconnected machines, such as fitness trackers and medical devices, the amount of data available to governments will only grow. Consider that in 2012, International Data Corporation estimated that the digital universe will grow from 130 exabytes in 2005 to 40,000 exabytes in 2020—an increase of more than 300 times.[1]

Deriving value from data and analytics

Today, data is increasingly becoming the most important asset for all enterprises. And governments are no exception. The challenge is turning all of this data into value—that is, using data to address critical issues and needs.[2] In the years ahead, the application of data science—the ability to extract knowledge and insights from large and complex data sets—will be crucial for states, counties, and cities to manage their operations more efficiently and administer their programs more economically while better serving their constituencies.

But the way some governments hold and manage data is keeping officials from gaining the full advantage of that information. At the same time, external forces, such as demands for greater transparency, efficiency, and more cost-effective services, are driving governments to reexamine how best to use their data. Here are examples of a few states and municipalities that have already demonstrated the value of improved data and analytics:

North Carolina formed a public-private partnership group to combine data from multiple sources to predict whether a sudden spike in patients with similar symptoms signals a pending epidemic or even a potential bioterrorist attack.

The group has already demonstrated its effectiveness by the early detection of a rash of stomach virus cases, as well as predicting the severity of a flu outbreak.[3]

Multiple state revenue departments routinely report capturing millions—sometimes even billions—of dollars in revenue by using data and analytics technologies to detect and prevent fraud and improve tax-payment compliance. Similar data and analytics techniques are being used to reduce fraud and abuse in programs such as Medicaid, which in 2015 alone had an improper payment rate of 9.8 percent or $29.1 billion, according to the federal government.[4]

In New York City, the mayor's Office of Data Analytics is working to help the city's agencies deploy resources efficiently. It has created the DataBridge, a central source for all kinds of regulatory information that can be accessed easily by workers in some 40 different agencies. By being able to apply analytics to this data, officials can uncover patterns and relationships allowing them to more effectively assign workers to enforce city regulations. For instance, the city was able to determine which restaurants were likely dumping cooking grease into sewers by comparing restaurants that didn't contract with a carting service for their old cooking oil with the geospatial data of the sewers. Rather than relying on chance for a health inspector to catch a restaurant worker pouring oil into the street, the data office was able to supply a list of likely culprits.[5]

In the area of public safety, Los Angeles, Atlanta, Seattle, and other U.S. cities are using a data technique called "predictive policing" to help curb street crime. Tapping into years' worth of statistics, an algorithm analyzes the data to determine where certain crimes, such as shootings, drug activity, and robberies, are most likely to occur. Software pinpoints problem areas on a map, which is then streamed into patrol cars, alerting officers to potential trouble spots.[6]

---

1 IDC iView, The Digital Universe in 2020, John Gantz and David Reinsel, sponsored by EMC Corp. (December 2012).
2 KPMG International, Going beyond the data, Christian Rast, et al. (June 2015),
3 Government Technology, Bio-Threat Early Warning System Gets Booster Shot, Brian Heaton (August 8, 2013).
4 pamentaccuracy.gov. For more information about reducing fraud, waste and abuse in Medicaid programs, see the KPMG publication, "High-performing state Medicaid program integrity programs: Putting it all together in the "Final Mile." http://www.kpmg-institutes.com/institutes/government-institute/articles/2014/11/high-performing-state-medicaid-integrity-programs--putting-it-al.html
5 GCN.com, How analytics is making NYC's streets and buildings safer, Rutrell Yasin (October 4, 2013).
6 The Guardian, Predicting Crime, LAPD-style, Nate Berg (June 25, 2015).

Even the delivery of basic services can also be improved through data and analytics techniques. In Pittsburgh, GPS trackers on snowplows let residents know which streets are being cleared and help the public works director prioritize snow removal and assign staff accordingly.[7]

The costs of poor data and analytics

When people hear the term data and analytics, they tend to focus and place emphasis on the "analytics" part of the term. Experience and leading practices have shown that as much, if not more attention, should be paid to the "data" part of the term. Data is an enabler of analytics, and poor quality data will lead to poor quality analytics and misleading or erroneous results, as the following examples vividly show:

An audit of Massachusetts' charter schools found some of the data unreliable and overstating the number of students on waiting lists for the schools. The audit also questioned the accuracy of other data reported from charter schools and local districts, making it difficult to judge a charter school's performance.[8]

Pennsylvania's secretary of environmental protection told Congress that fracking for natural gas had had no adverse effect on the state's water supply. But two years later, the department published 248 incidents of tainted well water caused by the gas development. This inconsistency was largely due to the department's six regional offices having vastly different methods of collecting, storing, transmitting, communicating, and dealing with the data and information. A subsequent audit determined that there was a lack of reliable information to effectively manage the program.[9]

New York City's Administration for Children's Services (ACS) struggled with a problem common to many state and local departments and agencies: multiple sources of data, on various platforms, in various formats, of various quality. Analysts would pull from different sources to report on the same performance measures only to come up with different answers. Eighty percent of their time was spent searching and messaging data and only 20 percent of their time was spent analyzing and reporting on the data. They realized they needed to fix their data before analysts, managers, and executives would have confidence in the analytic results.

The Florida Legislative Office of Program Policy Analysis and Government Accountability (OPPAGA) wanted to assess the effectiveness of its Medicaid reform program to decide whether it was worth expanding. Unfortunately, the necessary metrics weren't available. As a result, regardless of the potential of this program, OPPAGA wisely suggested deferring any expansion until the evidence was in. "To date, little data is available to demonstrate that Medicaid Reform has improved access to and quality of care," OPPAGA reported. It also concluded that "little data is yet available on whether Medicaid Reform has produced cost savings or is more cost-effective than traditional Medicaid."[10]

And here's a hypothetical illustration that demonstrates how data can help prevent even tragic outcomes: Consider the typical government caseworker for child services agencies whose job is to ensure at-risk children are protected and don't fall through the cracks. Caseworkers typically monitor these children through scheduled phone calls (perhaps monthly) and home visits (often every four to eight weeks). But what if these caseworkers could be made aware of the trigger events that could put a child at risk? For example, a parent is arrested or admitted to a drug treatment facility or a sibling suspended from school. Governments collect all of this data. If the various state agencies could easily share this information, a caseworker could quickly be made aware of the trigger event and intervene with the family at once.

Given these examples, the benefits and value of improved data and analytics are clear, and there is a limitless range of further potential applications of data and analytics in government. However, to successfully launch a data and analytics improvement program and reap its benefits, government officials must be cognizant of both the drivers and the potential barriers for initiating and executing the program in their state, county, or municipality.

---

7 The Center for Digital Government, Big Data and Analytics, 2015.

8 Boston Globe, State Audit finds fault with some charter school data, James Vaznis (December 19, 2014).

9 Government Technology, How Big Data in Government Falls Dangerously Short of the Mark, Katherine Barrett and Richard Green (June 24, 2015).

10 Governing, The Management Challenge of Bad Data, Katherine Barrett and Richard Greene (October 2009).

Drivers and barriers

Governments launch data and analytics improvement initiatives to achieve one of more of the following objectives:

– Enhancing service delivery and providing value to citizens through innovative combinations of data and related analytics

– Achieving better results and outcomes through the delivery of these services

– Spurring economic development

– Improving the efficiency of usage and allocation of resources

– Improving citizen access to data and providing more transparency to and accountability for performance

– Being compliant with federal reporting requirements, such as the Digital Accountability and Transparency Act

– Reducing fraud, waste, and abuse

– Ability to fund programs by monetizing data assets

Achieving these objectives requires overcoming policy, organizational, cultural, resource, and technical barriers, such as:

– Multiple, distributed data silos in various formats and various degrees of data quality

– Lack of willingness among agency or department personnel to share data, some based on valid legal or policy reasons, some based on rigid historical management methods or other perceived objections to sharing data

– Antiquated data management infrastructures

– Lack of data governance organization, processes, or tools and lack of sponsorship to implement data governance

– Lack of understanding for the "art of the possible" in the use of advanced analytics and big data to drive valuable insights

– Rapidly changing development and availability of big data and advanced analytics tools, both proprietary and open source

– Changing priorities

The chief data officer

Organizations are realizing the great potential value of improving government data and analytics capabilities and understanding that an executive commitment is required to promote a data-sharing culture and to launch, execute, and sustain data and analytics improvement programs. Consequently, a new executive position has emerged, the chief data officer (CDO).

The CDO is a relatively new position both in the government and commercial sectors, but it is increasingly being recognized as an essential role in growing and maintaining institutions' data and analytics capabilities. Colorado became the first state to appoint a CDO in 2010. And New York City became the first local government to appoint a CDO a year later.[11] More than a dozen CDOs are serving in state and municipalities across the United States and the number is expected to grow.[12]

A CDO's responsibilities typically include:

– Acting as a data and analytics evangelist and champion

– Promoting data sharing both internally and externally; facilitating data-sharing agreements

– Managing open data policies and developing guidelines to help agencies and departments identify and prioritize meaningful information and data sets

– Coordinating cross-department/agency data and analytics initiatives

– Facilitating standardization of enterprise data and analytics assets; promoting data and analytics standards in departments and agencies

– Enabling enterprise data governance

– Developing data as a service; providing transparent, easy access to all consumers of data independent of data location, format, or platform; making data "analytics ready"

– Leading the development of an enterprise data strategy and plan

Given the relative newness of the role, many CDOs are still finding their way to make the most powerful impact in their organizations. These CDOs say that one overarching goal is to find the most reliable data to make data-driven decisions and more effectively meet the needs of their agencies' clients.[13]

As states, counties, and large municipalities increasingly look to data and analytics to help achieve more efficient delivery of services and better program outcomes, the CDO will be a critical force behind developing a government enterprise strategy for delivering effective data and analytics programs. But overcoming the challenges and barriers previously cited will be a large and time-consuming undertaking.

The chief analytics officer

There is also another newly emerging role of the chief analytics officer (CAO). While the CDO is focused on the "data" side of data and analytics, the CAO is focused in the "analytics" side of data and analytics: how to use analytics—tools, processes, and methods to exploit the data and achieve state and local government objectives.

---

11 Government Technology, Which States and Cities Have Chief Data Officers, Steve Towns (June 2013).
12 Ibid.
13 Federal Times, Data officers put business first, not IT, Aaron Boyd (July 7, 2015).

Based on industry leading practices and years of delivering successful data and analytics engagements for a range of governments and commercial organizations globally, there are some initial steps a CDO can take to promote success.

1. Create, validate, and communicate a vision for the CDO organization

While technology is intrinsically tied to data and data management, the CDO needs to look at technology as just "one, albeit vital, component of a healthy and functioning data ecosystem within his or her organizations."[14] Rather, the CDO's focus should be on identifying data, ensuring it is reliable, and determining how that data can affect the outcomes of the enterprise's various services.

Therefore, one of the first and fundamental steps for the new CDO is to create a written vision statement that aligns data and analytics planned activities with executive and legislative objectives. For example, if the administration's goal is open data, the vision statement would revolve around helping to improve data access and transparency.

Once the vision statement is crafted, it's crucial to have the document validated by whomever the CDO reports to—ideally, the governor or another executive.

"Keeping and maintaining executive support is the only way that the CDO can be successful," said Viral Chawda, a managing director in KPMG's Data and Analytics CoE. "You want to make sure that you're not just heading off doing something that you think is great, but the governor or agency heads think isn't important. CDOs must make sure whomever their sponsor is, buys into it."

Once the vision is validated, it should be communicated using the enterprise's Web site and other forums for communicating to agency and department leaders.

2. Form a working group of members who can assist with developing the data and analytics strategy

Once the vision statement has been communicated, the CDO should form a working group consisting of representatives from the different agencies and departments (organizations) who can help shape the data and analytics strategy and plan. The members of this group should have the following traits:

– Understand and appreciate the value of data and analytics in improving service delivery both on an organizational and enterprise level

– Have a perspective on potential analytical use cases that could provide value to the organization and those use cases that cannot be executed either due to data challenges or other resource or capability gaps

– Ability to facilitate the identification and usage of the data and analytics assets within their organization

– Provide insights into their organization's data-sharing culture and data-sharing challenges

– Be a data and analytics decision maker, representing their organization

"The CDO is going to need people who are passionate about data and analytics and who will be passionate about the vision," Chawda said. "Ideally, these people will be able to not only provide great ideas but also be authorized to make decisions."

How does a CDO find potential members of the working group?

First, chief information officers (CIO), if they have the institutional knowledge, should be able to provide some suggestions for candidates for the working group. Second, the agencies may already have individuals aligned with data functions, either as a data steward or a technical liaison. Third, CDOs should definitely do a "walk around" to identify those professionals involved in IT and data issues.

Additionally, the CDO must be aware of the challenges faced by the people in the trenches—administrators, caseworkers, and the like. Some CDOs have said that their work involves the "decidedly low-tech themes of people, process, and culture."[15] To that end, CDOs should spend time speaking to not just to IT, but to program and policy people as well. Using insights gained from those conversations, CDOs can work to establish cooperation among agencies. For example, a state CIO, while discussing objectives with two agencies, realized that they both had similar geocoding-mapping requirements and that early benefits could be realized through geocoding standards and software technologies.
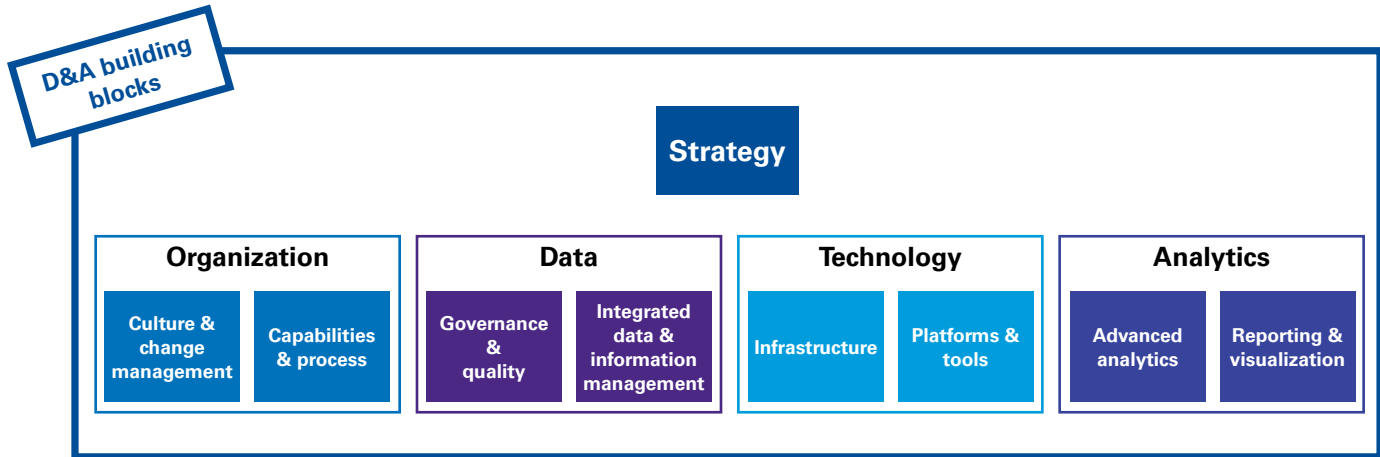
---

14 Nextgov.com, The Rise of the Chief Data Officer, Jack Moore (March 18, 2016).
15 Ibid.

3. Construct initial inventory of data assets

Before an effective data and analytics strategy and plan can be developed, the CDO will need to identify, at least at a high level, the data assets in place, under development, and planned for development.

Data assets can be characterized by four interacting dimensions:

D&A building blocks

Strategy

Organization	Data	Technology	Analytics
Culture & change management / Capabilities & process	Governance & quality / Integrated data & information management	Infrastructure / Platforms & tools	Advanced analytics / Reporting & visualization

Organization – What organizational components are in place for data governance, data management, and data analytics? What are the corresponding capabilities and skill sets?

Data – What are the available data/information assets, whether they be internal or external, original source, or derived analytic products? What processes are in place to govern, manage, acquire, control/secure, standardize, enhance, curate, exploit, access/distribute, and analyze the data? What memorandums of understanding (MOUs) are in place for sharing data?

Technology – What technology infrastructure and tool sets are in place to support the efficient and effective execution of the data and analytics processes?

Analytics – What techniques are being used to extract value from the data? These can include a range of capabilities from statistical analysis to predictive analytics to visualization.

When performing the initial inventory of the data, the CDO may want to consider the following data asset metadata:

– **Data asset title** – An identifying name for the asset.

– **Data description** – A broad description of the type of data and its usage. For example, mapping data (GIS) showing the locations of all the highway bridges in the state.

– **Data security classification** – What is the associated level of sensitivity? Is the data restricted to certain people within an agency and department or can the data

be shared with other departments or agencies or even exposed to the public? What is the reason or source of the classification? Does data need to be anonymized before it can be accessed or distributed?

– **Data accessibility** – If the data can be shared, are the mechanisms in place to make it accessible? Is it in a standardized format? What types of tools are available to access the data?

– **Data ownership** – What agency or department is considered the data owner? Who is the owner point of contact?

– **Data consumers** – What agency or departments are currently using the data? Are there MOUs in place or required to allow access?

– **Data structure** – Structured, semistructured, unstructured.

Also, as the data and analytics asset is being inventoried and cataloged, an assessment of its classification as an enterprise-wide asset should be made, considering:

– Does it or can it **support executive/legislative objectives?**

– Does it or can it **support multiple department/agency objectives?**

(It is important to recognize that this is an "initial" step in inventorying data assets, and that for these assets to be sustainable, this process needs to be periodically, if not continuously, executed and update).

4. Gather and understand existing data and analytics initiatives

Before a data and analytics strategy and plan is constructed, it's also important to identify and understand ongoing or planned data and analytics initiatives. This would not only have an impact on the future data asset inventory but could also represent opportunities for synergies. For example, if two agencies are separately gathering geographic information science (GIS) data to track and analyze usage of physical assets, the two efforts may be pooled, reducing acquisition and installation costs, and setting up a standard GIS interchange protocols. There may also be synergies gained if two or more agencies are purchasing the same research content.

5. Begin identifying and prioritizing candidate opportunities for exploiting data and analytics assets

Once these preliminary steps are taken, and the data is inventoried and existing projects noted, the CDO, with the support of the working group, can begin the process of prioritizing candidate data and analytics improvement projects. CDOs can use the following criteria to help prioritize and select initial data and analytics projects.

– Alignment with executive and/or legislative objectives.

– A use case that clearly demonstrates value.

– How quickly can results be demonstrated and value delivered?

– Can existing projects be combined or leveraged?

– Can the project's results be leveraged by multiple agencies?

– Is there a clear business case for the project, considering value, schedule, life cycle costs, and risks?

– Will the results of the data and analytics project be sustainable or are you seeking a one-time answer/insight?

For the results of the data and analytics project to be sustainable, the necessary infrastructure (organization, data and information management, technology, and analytics) needs to be in place. Therefore, the CDO needs to develop a plan that balances infrastructure needs with early demonstrable results. A too heavy focus at the onset in building out all the infrastructure means agencies may have to wait too long to see the value from the effort. On the other hand, a single-project approach that delivers results quickly, but lacks a solid infrastructure will not be sustainable in the long run.

"So, you need to come up with a plan that balances multiple capabilities—that contributes to building the data infrastructure, but also demonstrates early value," Chawda said. "If not, your plan won't in the long run because you won't have the right pieces in place."

6. Select and execute a pilot project

This is the critical final step. A well-thought-out and designed pilot will demonstrate the benefits of data and analytics and prompt buy-in from agency and department officials. Lessons learned will ease the application of the program to other agencies and departments. As an example, consider NYC's ACS INFO project. The agency initially targeted their data and analytics project towards standardizing and aggregating data, as well as enhancing their reporting and analytics capabilities for their Department of Policy and Planning. When the Finance Department realized the value of the project, they subsequently jumped on board and are now using the same data and analytics asset to improve their claims processing as well as detecting potential fraud.

A few caveats

These steps won't necessarily be sequential. A CDO may come into office and immediately receive a mandate from the governor for a specific project—for example, open data or another policy initiative. In that circumstance, the CDO can postpone developing the overall vision statement, moving right to data inventory and determining what information is currently shared with the public and what isn't.

Also, some of these steps can be done simultaneously, such as forming the working group, inventorying data, and identifying existing data initiatives.

In addition, CDOs should remember that objectives and priorities can change suddenly, either because of budget challenges, changes in administration, or new federal mandates, to name a few. In those cases, the data and analytics initiatives need to be revisited, reevaluated, and revised accordingly.

Final thoughts

The data collected and available to governments at all levels is growing explosively and will continue to do so. The goal is to turn that data into useful information to help agencies better carry out their missions with less waste, more efficiencies, and better outcomes—and even potentially short-circuit situations with potentially tragic consequences.

Effective data and analytics can lead the way for government officials to glean the full benefits from their data. Under the leadership of a capable CDO, government agencies can begin to break down silos to better share information, establish common standards to make data easier to access, and create greater transparency for citizens, businesses, and other officials to find the data they need in a format they can use.

By developing a successful data and analytics initiative, CDOs can provide a strong beginning to a government's data and analytics efforts, and ultimately lead to a better government to improve the quality of life for all its citizens.

How KPMG can help

Data and analytics, at its core, is a new way of solving problems with insight and innovation. KPMG is specially qualified to help governments and their agencies transform their organizations into data-driven, decision-making organizations.

Our understanding of government processes, regulation, and data is deep. KPMG has a long history serving government through our audit and advisory services, including data analytics. Our team includes industry specialists and subject matter professionals in healthcare, financial capital markets, financial management, grants, and federal and state regulations who understand the data at a very granular level. They are supported by our large team of data scientists. Our approach to data analytics is a business-led, collaborative approach. Our experienced team focuses on developing an agile, incremental approach to analytics to quickly deliver value and generate business results.

KPMG has strategically invested in data and analytics, unifying our resources into a CoE. In addition to our robust team of data scientists, the CoE houses preconfigured analytics technology platforms and a variety of proprietary and third-party data sets. It provides our teams with advanced capabilities in big data, predictive analytics, optimization modeling, and analytics technologies.

Contact us

John Druke
Principal, Data and Analytics Lead,
State and Local Advisory
T: 315-727-6513
E: jdruke@kpmg.com

Sid Frank
Director, Information Management,
State and Local Advisory
T: 770-833-0983
E: sidfrank@kpmg.com

Traci Gusher
Managing Director, Data & Analytics
T: 267-256-1617
E: tgusher@kpmg.com

Viral Chawda
Managing Director, Data & Analytics
T: 832-535-8712
E: vchawda@kpmg.com

Jeffrey C. Steinhoff
Managing Director, KPMG Government Institute
T: 703-286-8710
E: jsteinhoff@kpmg.com

Visit the KPMG Government Institute at www.kpmg.com/us/governmentinstitute.

kpmg.com/socialmedia

Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates.