



GenAI in federal government

Five things agency leaders need to be preparing for now

Few technologies have transitioned from obscurity to implementation as rapidly as generative artificial intelligence (GenAI). Were we writing this article just a few months earlier, we might be explaining what GenAI is and the many benefits it can offer federal agencies. Today, most are past the “what” and the “why” phase and are now looking at the “how.”

Recognizing federal agencies’ keen interests in adopting GenAI, the Office of Management and Budget (OMB) released Memorandum M-24-10 in March 2024. At its core, the memo established “requirements and guidance for AI governance, innovation, and risk management” so that agencies adopt AI responsibly. However, in a true testament to the sea change in capabilities ushered in by GenAI, the memo went a step further by mandating federal agencies identify and remove barriers to AI adoption.

In this same bullish spirit, the memo also set an aggressive 60-day deadline for most federal agencies to appoint a chief artificial intelligence officer (CAIO) if they hadn’t already. As we’ve seen, this newly defined role has largely been filled by incumbents from IT and data operations. As a result, CAIOs will likely have to wear many hats, juggling OMB M-24-10 compliance and GenAI implementation efforts alongside their other executive duties, including managing any pre-existing modernization efforts.

Compounding the challenge, CAIOs face the remarkable pace at which AI technologies and capabilities are evolving. This means CAIOs are finding themselves in the unusual yet exciting position as early adopters. Historically, governments have always relied on the private sector to blaze the trail ahead, to work out the kinks, discover where the landmines were buried, and provide models for optimum deployment efficiency and value. But not in this case, where even GenAI developers and solution providers are paddling as fast as they can to try to keep up with the flow of developments given how quickly the technology continues to evolve.

Why modern government is important

Government agencies in the US must modernize in order to keep up with changing user needs, regulations, and health and public safety requirements. Leaders of modern governments rethink business processes and service delivery models to more effectively achieve their mission. This article is one of a series that features how modernizing affects the government workforce and the user experience, improves security and public trust, and accelerates the digital journey. KPMG team members offer insights intended to help guide governments in their modernization efforts to encompass all processes, technologies, policies, and the workforce so each works together to create connected, powered, and trusted organizations.

Although it’s still early days, at KPMG we’ve worked with several federal agencies as well as commercial clients on GenAI strategy, experimentation, and full-scale implementation. We have heard the questions being asked and many of the same concerns arise. To help agency leaders get out ahead of these issues, here are five things we believe they should be thinking about now to lay the groundwork for OMB M-24-10 compliance and successful GenAI adoption.





1 Start building an AI-ready culture now

GenAI has the potential to automate a huge percentage of our tasks because AI systems have gotten so much better at understanding and replicating the way we naturally communicate, a skill that cuts across nearly every domain of work. Unlike previous technological advancements, AI is projected to have at least as much impact on white-collar jobs as blue-collar jobs.¹ AI will accelerate a shift in ways of working in different ways and to different degrees across the spectrum of income and skills. In other words, we'll all feel the effects.

This reality has not gone unnoticed by employees. In a recent study conducted by the American Psychological Association (APA), close to two out of every five workers reported worrying that AI might make some or all of their job duties obsolete in the future.² Employees who are insecure about their continued employment display lower levels of performance, commitment, well-being, and trust in the organization.³

The same APA survey revealed that 66 percent of workers worried about AI said their employer thinks their workplace is a lot mentally healthier than it is. At the same time, nearly half of private sector executives responding to our [Q1 2024 KPMG GenAI quarterly pulse survey](#) said that they planned to measure ROI on their GenAI investments through employee satisfaction.⁴ Unless this disconnect between the perceptions and AI aspirations of executives and the AI-based fears of employees is addressed, the result may be an organizational culture that forms a major barrier to the success of any AI effort. Employee buy-in is an essential ingredient.

You can't start soon enough on building an AI-ready culture even if you haven't yet identified your first GenAI project. Our culture strategy colleagues at KPMG have a [detailed, multistage approach](#) to help build belief and buy-in at the tactical level for individual use cases as well as at the strategic level for building an AI-ready culture across any organization. But at its core, this work requires one essential thing: listening to employees and involving them throughout the entire process so that any GenAI implementation is done *with* them and not *to* them.



¹ Source: "AI Is Starting to Threaten White-Collar Jobs. Few Industries Are Immune," Ray A. Smith, The Wall Street Journal, February 12, 2024

² Source: "Worried about AI in the workplace? You're not alone," Michele Lerner, American Psychological Association, January 30, 2024

³ Source: "Understanding and Exploring the Concept of Fear, in the Work Context and Its Role in Improving Safety Performance and Reducing Well-Being in a Steady Job Insecurity Period," Diego Bellini et al., Sustainability Journal, 2022

⁴ Source: "KPMG GenAI Quarterly Pulse Survey: The path to sustainable returns," KPMG LLP, March 22, 2024

2 Focus on the right challenges

It's clear that GenAI can have a significantly positive effect on agencies. In many cases, it can help improve employee efficiency and effectiveness, helping them deal less with bureaucratic or mundane tasks to become more customer- or mission-focused, and improving job satisfaction by making work less tedious and more interesting and compelling. It can enable agencies to offer new and improved services or provide insights that had otherwise remained elusive.

But how to get from here to there?

How do you select the right technology provider? Which capabilities do you want to enable? For whom do you turn it on? What training will be required? What impact will it have on the budget? How do you measure performance, risk, and ROI? What impact will it have on the skills needed in new hires, the agency's culture, its relationship with unions? How does it change contracts? What data will it be linked to? How are data sources evaluated? What controls will you need surrounding it? How do you secure it? How do you ensure it delivers accurate responses? What parameters define its ethical use? How are potential biases or unethical uses identified and addressed? The list goes on.

These challenges aren't unique to government. Private sector chief executive officers (CEOs), for example, say ethical challenges are their top obstacle to successfully implementing GenAI, not budget or technology issues.⁵ These concerns are well-founded, too, since most GenAI models, including large language models (LLMs) such as ChatGPT, are well-known to convincingly fabricate information when they lack specific knowledge. Known as "hallucinations," these fabrications are notoriously hard to detect because the models have been trained to be helpful assistants that exude confidence and authority instead of self-doubt and ineptitude. The fact that humans tend to attribute a great deal of authority and trustworthiness to software means we're even more likely to mistake fiction for fact when it comes from an AI system.⁶

Consider, too, that many GenAI model outputs are nondeterministic. This means an input that created a hallucination at one point in time might not create a hallucination when used again—or it might result in a different hallucination. This functionality is a stark departure from how other AI/machine learning models operate and is one of the most challenging technical aspects of engineering trust into GenAI systems.

The good news is that researchers have been hard at work with these issues for well over a year now. And the research has shown that LLMs have a remarkable capability now

known as in-context learning that allows them to "learn" from information provided to them as an input without an explicit training. This remarkable capability has led to the development of a novel hallucination mitigation technique called retrieval augmented generation (RAG). In RAG, you combine LLMs with more traditional machine learning techniques and a specialized search database to pull in relevant, contextual information from myriad data sources based on a user's input.

For instance, say a user asks an LLM about their company's health benefits. This is private information that no LLM was ever trained on. Because of this, an LLM will either tell you it doesn't know about your company's health benefits—or it will hallucinate details in an effort to be perceived as helpful. RAG can prevent the hallucinations by *retrieving* information from the company's internal documents and then *augmenting* the original question with relevant information to improve the LLM's *generation*. In essence, RAG makes question/answer like an open-book test for your LLMs. However, RAG isn't a cure-all. LLMs might still hallucinate, especially if they are provided with irrelevant or voluminous retrieved content.

RAG also creates a new issue: how do I give an AI model access to trusted, high-quality data? That's why many organizations, both public and private, have had to quickly adjust their data and technology modernization roadmaps. Government organizations, perhaps more so than any private entity, sit atop vast, untapped troves of structured and unstructured data. This data can reduce risks in GenAI systems, but it can also increase risks if the data isn't of a high enough quality. The old adage of "garbage in, garbage out" most definitely still applies to GenAI.

All of this is to say that, in all but the most trivial of use cases, GenAI isn't a snap-in solution. It's an enormous collection of things agencies must account for—across people, processes, technologies, and data—far faster than they are used to. It adds a new dimension to what was already a vast ecosystem of emerging tools and technologies that must be thought through, coordinated, and transformed. The good news is that these things have been thought through, and there are solutions designed to help any organization address them.

For example, [Trusted AI](#) is our framework for designing, building, deploying, and using AI systems in a safe, trustworthy, and ethical manner. It's designed to help organizations map, measure, manage, and govern AI-related risks and challenges, including reputational, compliance, security, privacy, and even value risks, where insufficient resources, capabilities, or technologies might diminish the value organizations can derive from their GenAI solutions.

⁵ Source: KPMG 2023 US CEO Survey

⁶ Source: "Artificial intelligence and illusions of understanding in scientific research," Lisa Messeri and MJ Crockett, *Nature*, March 6, 2024



3 Understand what the technology really can and can't do

Along with GenAI's rapid appearance has come an incredible amount of confusion and hype, where some believe it can do everything from writing award-winning novels (which it actually did⁷) to solving complex global issues, while others view it as a tool that, although powerful, has its limitations and requires careful handling to avoid ethical and practical pitfalls. The latter camp is the right place to be and getting there starts with understanding GenAI and its relationship to more traditional AI and machine learning techniques.

As its name implies, the focus of GenAI is on generating new content—primarily text, images, and video—designed to be indistinguishable from human-generated content. There's a common misconception that GenAI is "AI 2.0" or that it has replaced more traditional AI models, but they are really designed for two very different purposes. Put simply, "traditional AI can analyze data and tell you what it sees, but generative AI can use that same data to create something entirely new."⁸ More advanced AI systems attempt to leverage both.

A good starting point for an even deeper understanding of what GenAI can do is by looking more closely at the most popular subset of GenAI: LLMs. Broadly, there are seven use case categories for LLMs:

- Generate
- Summarize
- Rewrite
- Extract
- Search/Similarity
- Cluster
- Classify.

Because of the general-purpose nature of most LLMs, domain expertise can help you derive domain-specific use cases from these seven use case categories. For instance, you could use the *classify* capability of an LLM to identify fraud in financial transactions—a finance use case—or to identify unacceptable terms and conditions in a contract clause—a contracting use case. Of course, more traditional machine-learning techniques can also—and have been—performing these functions. The point is that it's important to not get caught up in the GenAI hype and automatically dismiss what might be a more ideal solution if it doesn't happen to include GenAI. Just because we have calculus, for example, doesn't mean we no longer need algebra. In many of our conversations with agency leaders, we help them match the right tool for the job at hand to help them avoid wasting precious time, effort, and resources heading down the wrong path.

⁷ Source: Sonja Anderson, "ChatGPT Helped Write This Award-Winning Japanese Novel," Smithsonian Magazine, January 24, 2024

⁸ Source: Barnard Marr, "The Difference Between Generative AI And Traditional AI: An Easy Explanation For Anyone," Forbes, July 24, 2023



4 Focus on the use cases and the risks, not on the technology

While a broad understanding of the technology is essential, it's more important to understand that you don't implement GenAI technology; you implement GenAI use cases. And the risks only come to the fore in light of the use case's business context.

Technology is never enabled for technology's sake. It's always designed to drive a business outcome and help further the agency's mission in some way. Matching—or more likely, being able to adapt—a provider's technology to your specific needs becomes the key challenge. That abstraction between what the technology providers are offering and where GenAI can help improve mission delivery adds complexity to the decision and implementation processes.

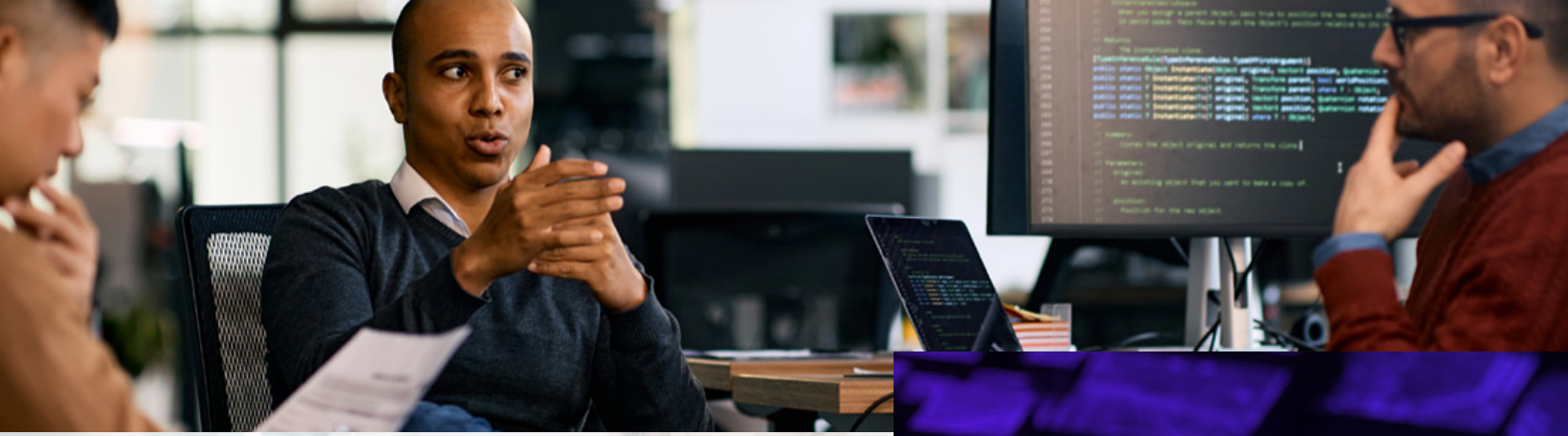
Importantly, if agencies focus solely on the technology and not the context of each use case, then they'll fall short of OMB M-24-10's requirements for risk management practices. Indeed, as The National Institute of Standards and Technology (NIST) says in their AI Risk Management Framework, "AI risk management can drive responsible uses and practices by prompting organizations and their internal teams who design, develop, and deploy AI to think more critically about context and potential or unexpected negative and positive impacts."⁹

While much has been written about the risks associated with AI and GenAI in particular, it's worth recapping the highlights to see how the business context ought to inform risk management:

- **Ethical complexities:** Use of GenAI raises ethical questions about consent (especially when using personally identifiable information for training purposes), autonomy, and the potential for the technology to make decisions that could significantly impact individuals' lives and well-being.
- **Data dependence:** The reliability of a GenAI model is dependent on the quality of the data it uses and the methods used to train the model. Erroneous, incomplete, and/or out-of-date data could lead to flawed outcomes. For use cases that have a bearing on human rights or safety, which is almost the exclusive purview of the government, the stakes are high.
- **Bias potential:** If the data a model is trained on contains biases, then the content it creates is likely to contain biases, too. This could inadvertently perpetuate socioeconomic, racial or gender disparities in program benefits or services.
- **Lack of transparency:** GenAI doesn't inherently provide explanations for decisions it makes. This opaque "black box" nature of some GenAI models can be a barrier in instances that require openness and accountability in decision-making processes. There are ways to back into transparency, but they might not prove robust enough for use cases that impact human rights or safety.
- **Inadequate measures of performance:** How do you know if your GenAI solution is delivering as promised? Standard test, evaluation, validation, and verification (TEVV) methods don't necessarily apply due in large part to the nondeterministic nature of model outputs. Each use case will likely require its own approach to TEVV.
- **AI overreliance:** Overreliance on GenAI could potentially lead to the devaluation or atrophy of human judgment and expertise, which are crucial in nuanced and complex decision-making processes.

Each of these risks will look different within the business context of a specific use case. For instance, consider how a junior software developer's overreliance on a coding assistant could stymie their career growth and even harm the project they are working on. Without years of on-the-job experience, it may be hard for a junior developer to fully understand every code suggestion generated by an AI. Should they be pressured into tight deadlines in the absence of a more senior software developer's oversight, they might lapse into the practice of accepting code suggestions that "just work" without bothering to understand why or how the code works. Over time, overreliance on their AI coding assistant could rob them of valuable learning opportunities. In the worst case, they unknowingly introduce a critical bug into their team's codebase that they are completely ill-equipped to debug due to their lack of understanding. Understanding the risks associated with specific use cases like this is crucial to prioritizing AI projects and identifying the quick wins.

⁹ Source: "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," The National Institute of Standards and Technology, January 26, 2023



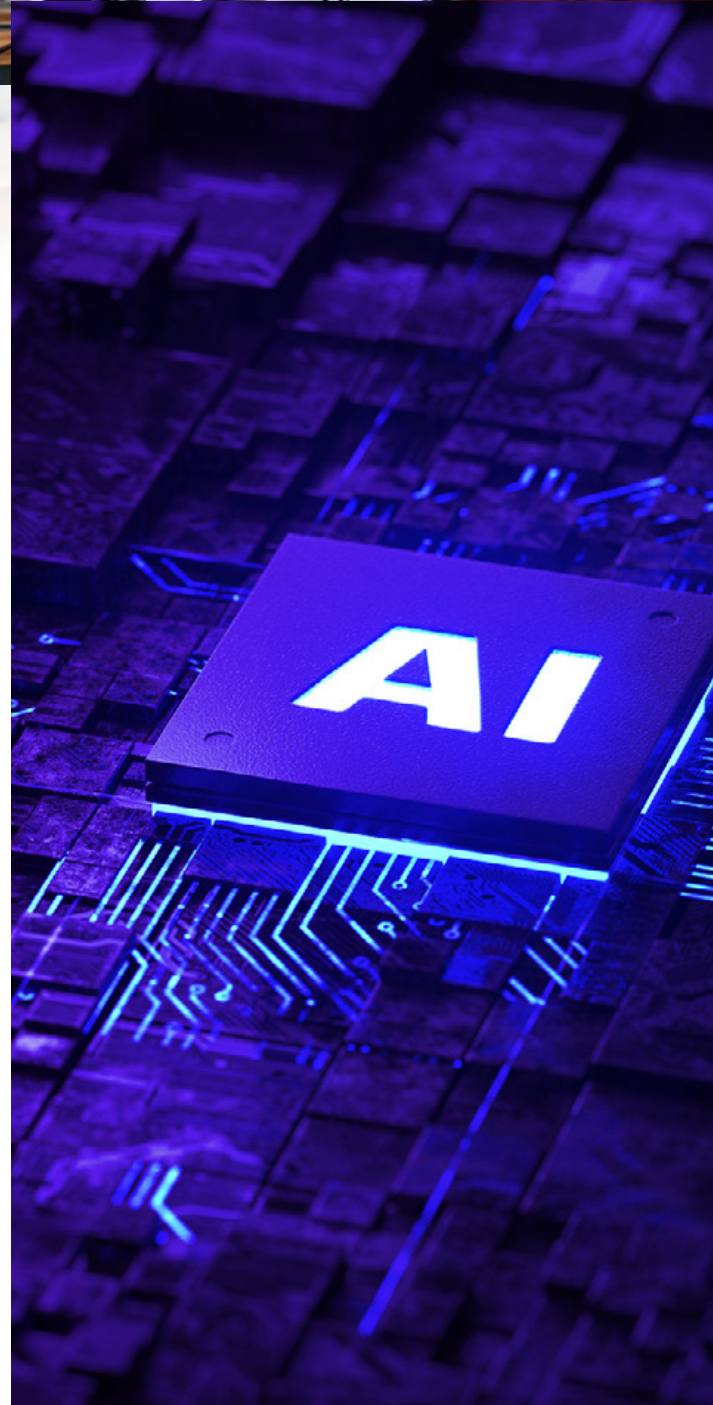
So where do you get started?

From our own experience in rolling out GenAI to our 30,000 US professionals as well as from our experiences with various federal and commercial clients, it's helpful to differentiate between Everyday AI and Specialized AI. Everyday AI covers the most common, foundational tools and apps in our daily lives. Specialized AI includes advanced tools and techniques designed to radically reshape professions. Everyday AI is the low-hanging fruit, and it generally manifests in one of three forms:

- **Enterprise platform AI:** Enterprise platforms with embedded, turnkey AI features
- **Productivity apps:** AI embedded into daily-use productivity apps, such as document editors, meeting apps, and even code editors
- **White-label chat:** Controlled access to a private GenAI chatbot for freeform content generation.

The range of what you can do with these everyday capabilities is wide. And their implementation serves as the launching point for tomorrow's bolder and more advanced use cases.

To help you identify those use cases and simultaneously address the talent dimension of AI readiness, a careful audit of existing job requisitions may be the ideal place to start. There are many benefits to this approach. Even as it creates entirely new jobs, technology has a habit of significantly changing and even eliminating existing jobs—we no longer have travel agents or elevator operators, for example. AI is no exception. And, undoubtedly, there's lots of cruft that has built up in job requisitions over the years, too. A careful reexamination of position descriptions within your agency not only will help you prepare your talent for AI but also may help identify GenAI use cases through an examination of your employees' roles and responsibilities.





5 Stay ahead of cost and budgeting issues

OMB M-24-10 requires each AI use case to have an impact assessment. Part of this assessment is “the intended purpose for the AI and its expected benefit.” Calculating that expected benefit requires factoring in the costs to the agency of the GenAI capability. This becomes particularly important for agencies that operate through working capital funds or other budgetary authorities as they might need to consider service fee increases or other chargeback mechanisms to defray their new AI costs.

Just calculating the costs could be a major challenge. For example, a software provider might offer access to GenAI features on a per-user-per-month basis, but that may not be for every GenAI feature the app might provide. In other words, it might be not a blanket “for this amount you can use it all” but rather, “for this amount you can use a pointed set of functionality.” If you navigate to the plans and pricing page of OpenAI’s ChatGPT, you’ll see exactly this type of licensing model.

To project licensing costs—and avoid paying a licensing fee for users who don’t need a particular set of GenAI functionality—you need to carefully identify and track which use cases should have which sets of GenAI functions enabled. How are you going to keep track of this potentially complex use-case/feature/user/cost matrix, especially when it might involve a dozen or more licensing options, and grow to hundreds or thousands of uses cases and tens of thousands of users?

Unfortunately, contract and licensing costs are the easy part of costing an AI use case. Quantifying your risk exposure is the harder part, and federal authorities such as the Government Accountability Office, OMB, NIST, and others have not been very prescriptive on how agencies ought to do this.

Although useful as a heuristic, qualitative risk registers won’t serve you well when you are trying to perform an apples-to-apples comparisons between competing AI use cases. This makes portfolio governance at the CAIO level very difficult even when you can soundly quantify more traditional costs such as licensing.

It was in response to these very challenges at one federal healthcare agency that we at KPMG developed a probabilistic risk model that can quantify and account for the complex, potentially interdependent risk factors of an AI system. By enhancing understanding of the complex, multidisciplinary risk factors at play, as well as their relationship to undesirable impacts, the methodology ultimately helps stakeholders uncover the right set of risk mitigation strategies and monitoring plans and to then incorporate their costs into a more holistic assessment of a use case’s ROI.

Given the austere budgetary environment established by the Fiscal Responsibility Act of 2023, avoiding budgetary surprises may be more important than ever for the successful adoption of AI.





How KPMG can help

KPMG has worked with federal, state, and local governments for more than a century, so we know how public sector agencies, education institutions, and healthcare organizations work—and we’re helping them design and implement real AI use cases today.

We’re experienced, nimble, and flexible. We understand the unique issues, pressures, and challenges government organizations face on the journey to AI adoption. We’ll meet you where you are on that journey and help advance your progress with no agenda other than to see you succeed. We’ll help you leverage the investments you’ve already made to help maximize their value—not try to sell you something new.

We offer clarity and insight. As a trusted advisor, we can help you make sense of everything going on in the highly dynamic world of AI that can impact your mission, from regulatory mandates to emerging technologies. We can help align your efforts with leading practices from both the private and public sectors, and help keep you moving forward quickly with confidence and conviction.

We see the big picture. We can help you anticipate and adapt to the wide-ranging impacts AI can have on your organization, including budgets and financial controls, business processes and operating models, and employee growth and retention. We can help you understand your data—where it comes from, what controls are required, how to help maximize value locked in it, and how to share that value across organizations. We can help you harness the power of AI ethically and responsibly with trusted AI principles and governance models for managing risk.



We can help you from strategy through implementation.

Unlike business-only consultancies, our more than 15,000 technology professionals have the resources, the skills and experience, the battle-tested tools and methodologies, and the close alignment with leading AI technology providers to help achieve your vision quickly, efficiently, and reliably. And unlike technology-only firms, we have the business credentials, subject matter professionals, and public sector experience to help you deliver measurable results, not just blinking lights.

About KPMG

KPMG has worked with federal, state, and local governments for more than a century, so we know how agencies work. Our team understands the unique issues, pressures, and challenges you encounter in the journey to modernize. We draw on our government operations knowledge to offer methodologies tailored to help you overcome these challenges and work with you from beginning to end to deliver the results that matter.

The KPMG team starts with the business issue before we determine the solution because we understand the ultimate mission. When the way people work changes, our team brings the leading training practices to make sure your employees have the right knowledge and skills. We also help your people get value out of technology while also assisting with cloud, advanced analytics, intelligent automation, and cybersecurity. Our passion is to create value, inspire trust, and help government clients deliver better experiences to workers, citizens, and communities.



Contact



Viral Chawda

Principal, Advisory
Head of Technology Practice—Government
KPMG LLP
832-535-8712
vchawda@kpmg.com

read.kpmg.us/modgov

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

Learn about us:



kpmg.com

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act upon such information without appropriate professional advice after a thorough examination of the particular situation.

© 2024 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organization.