Cutting through the hype of GenAI in HHS

It's an understatement to say that state Health and Human Services (HHS) agencies are facing pressure to implement generative artificial intelligence (GenAI). In some cases, it's coming from the top. In September 2023, for example, California governor Gavin Newsom issued an executive order requiring state agencies to explore ways to use it.[1] Even if there's no explicit mandate, the lure of GenAI's potential to dramatically increase worker productivity and job satisfaction or improve constituent service is a powerful motivator spurring many agencies to at least begin looking at its adoption. For anyone involved in HHS, it can't be pushed off or ignored.

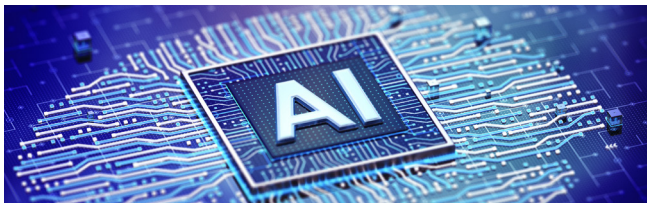
But in the rush to jump on the GenAI bandwagon, many HHS agencies may find themselves putting the cart out in front of the horse. GenAI isn't a snap-in or a "set-it-and-forget-it" solution, even if some solution providers are now touting their new "integrated" GenAI offerings that can be enabled with a click. It's an enormous collection of things agencies must account for—across people, processes, and technologies—far faster than they are used to and a vast ecosystem that must be carefully thought through, coordinated, transformed, and continuously evaluated.

At this early stage, agencies may simply be struggling to identify the ideal use cases for the technology, but the questions don't stop 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? How does it change how you work with people? How might it impact headcount and job descriptions? How could it affect contracts with labor unions? What data will it be linked to? What controls will you need surrounding it? How do you secure it? How do you ensure it delivers accurate and nonbiased responses? How are compliance and regulatory mandates covered? How will it be monitored and maintained? How will it affect the agency's budget?

Why modern government is important

Government agencies in the U.S. 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.

Introducing such a significant change into any one job will clearly require a rethinking of how that specific work gets done, but collectively across multiple jobs, it becomes an opportunity to rethink and redesign how *all* work gets done—how the agency's operating model should change. It's asking not just how the technology might affect business processes and constituent interactions but also how it *should* affect them to achieve the agency's mission more effectively. That's a big ask but an equally large opportunity.



[1] Source: Sophia Fox-Sowell, "California Gov. Newsom orders study of generative AI use in state government," StateScoop, September 6, 2023

Is GenAI even the right tool for the job?

The world has been captivated by GenAI—and for good reason. It has wowed us by how it easily understands our questions, written the way we would ask a person naturally, and the remarkably human-like responses it provides in return, mined from an incredibly vast wealth of information. Forget the stilted queries we had become accustomed to writing with search engines and the multiple clicks required to follow up on the results, or the incredibly frustrating responses we get from chatbots that clearly didn't understand what we had asked. With GenAI, it sounds like we are having a regular conversation with a very well-informed person who can create incredibly rich and relevant content in the blink of an eye.

The power of this enhanced experience is not to be understated. But it's also important to understand that GenAI isn't a panacea. You never implement technology for technology's sake. It's always to help address a "business" challenge. The first question that must be asked, therefore, is if GenAI is the right tool for the job.

Some may see GenAI as "AI 2.0"—that it has now made all other forms of AI obsolete. However, it was not designed to replace other more "traditional" AI models but rather to serve a very different purpose. Just because we now have calculus, for example, doesn't mean we no longer need algebra.

More specifically, GenAI wasn't designed to produce accurate insights into a data set; it was designed to generate novel, creative output based on it—hence the name *generative*.

As the University of Southern California warns its researchers, "Generative AI models are not databases of knowledge but rather an attempt to synthesize and reproduce the information they have been trained on." That's an important distinction, especially when it's applied to HHS applications.

GenAI cannot help you detect fraudulent activity, for example, or other patterns that may be developing within your program's data such as a shift in demographics or an unusual increase in claims from a particular healthcare provider. As a data scientist might put it, using GenAI for such applications would be trying to apply a probabilistic solution to a deterministic problem.

Other examples where traditional AI—but not GenAI—may be more appropriate include:

- Identifying actionable insights from application and case-level reporting, including denials, error-prone case compositions, and application churn to drive targeted compliance strategies
- Assessing and validating federal reporting requirements on timeliness, payment errors, application churn, etc. against the data captured in your case management and eligibility system
- Collaborating cross-functionally with appeals, claims, and other programs for insights to drive greater program coverage, stronger program integrity, and more seamless federal reporting.

There are incredible AI models built on the same or similar neural network technologies as GenAI but that have been designed for entirely different purposes. Knowing the difference—selecting the right tools for the right challenges—can save time, effort, and budget wasted traveling down the wrong path toward disappointment—or worse.

Understanding—and addressing—GenAI's limitations and risks

GenAI has its limitations. It writes poetry. It creates realistic images based on text descriptions you provide. But it can't do math reliably. It can write software, but that software isn't guaranteed to not contain bugs. The answers it provides aren't necessarily consistent. And they're also still highly dependent on the prompt or question asked.

Using GenAI to help write an email or a science paper is one thing, but using it in situations where people's lives are on the line is a very different matter. Protecting personally identifiable information (PII) or personal health information (PHI) is essential for HHS agencies, as is ensuring that the people agencies serve aren't misled or denied services to which they're entitled. And that's just the tip of the iceberg. The risks can be far broader and more challenging to address than some might imagine, including:

Hallucinations: Because of the design and intended use of GenAI models—to creatively assemble information—it can be incredibly difficult to validate and properly attribute the basis of their content.[2] As a result, GenAI models can be susceptible to "hallucinations"—i.e., when they present completely false or invented information as though it were fact. New technologies are being designed to help address this risk. Retrieval-Augmented Generation (RAG) can limit the model to generating answers based on a specific set of data or documents instead of the internet as a whole, for example. While some RAG providers may tout their solutions as "hallucination free," this may not be quite accurate.[3] RAG does indeed help, but the technology still has its limitations and has been shown to generate its own "RAG hallucinations."[4]

AI overreliance: Overreliance on GenAI could potentially lead to the devaluation of human judgment and expertise, which are crucial in nuanced and complex decision-making processes. It's essential to maintain a balance between AI assistance and professional judgment.

Ethical complexities: Use of GenAI raises ethical questions about consent (especially when using PII or PHI 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, or out-of-date data could lead to flawed outcomes, such as inaccurate eligibility determinations, which could negatively impact the individuals relying on these essential services.

Bias potential: The risk of algorithmic bias is high. If the data a model is trained on contains biases, the content it creates is likely to contain biases, too. In an HHS context, this could inadvertently perpetuate socioeconomic, racial, or gender disparities in program benefits or services. To counteract this, AI training and deployment requires close monitoring to help promote fairness.

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, such as determining eligibility for government assistance programs. When benefits or services are denied, people have a right to understand why. Remedying this requires work to ensure AI is explainable or coupled with systems that offer suitable explanations. Here, too, RAG may help by providing verifiable citations for its output based on a carefully curated set of documents (such as official policies or legislation), but it still won't fully answer how the model chose which documents it selected (and which it ignored) or how it decided to generate the content it did.

Inadequate measures of performance: How do you know your GenAI solution is delivering as promised? Standard user acceptance testing (UAT) methods don't necessarily apply. The solution might reliably generate an answer to five-nines standards—but not necessarily the "right" answer or a particularly useful one. On an ongoing basis, entirely new key performance indicators (KPIs) may need to be developed to objectively assess the effectiveness of a solution.

[2] Source: "Using Generative AI in Research," University of Southern California, March 1, 2024
[3] Source: Kyle Wigger, "Why RAG won't solve generative AI's hallucination problem," TechCrunch, May 4, 2024
[4] Source: Iris Zarecki, "RAG Hallucination: What is It and How to Avoid It," K2View, May 20, 2024

Other considerations

AI in any form, including GenAI, can be incredibly disruptive to an HHS agency, and its impact can be far reaching. Other aspects agencies must consider include:

Implementation and maintenance costs: While AI can help significantly streamline processes such as eligibility determination, case management, and program administration in HHS programs, it may require significant up-front capital for its development, testing, implementation, and maintenance. Additionally, continuous investment is necessary for updating the AI models as program rules, regulations, and needs evolve.

Technological infrastructure: HHS programs are often operated on various platforms with diverse levels of technological sophistication. Any GenAI solution needs to be compatible with existing systems, ensuring it can integrate seamlessly without compromising the functioning of those systems.

Training and fatigue: While AI can assist in interpreting complex policy changes, it also requires regular updates and training based on new policy enactments, user feedback, and system performance. This could potentially lead to a form of fatigue that might undermine the enthusiasm for the system, causing it to be underutilized.

Regulatory compliance: Those who opt to embrace AI must navigate the extensive regulatory landscape specific to HHS programs such as Medicaid, Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance for Needy Families (TANF), etc. A combination of state, federal, and program-specific regulations and recommendations must be considered and incorporated appropriately into AI systems to help address compliance requirements.

Job displacement: AI technologies can automate certain job functions, potentially displacing jobs in the field. A responsible approach would involve foreseeing these impacts and planning for workers' reskilling and job transitioning.

Integrating GenAI with integrated eligibility

Many states are in the throes of multiyear integrated eligibility system (IES) projects, not all of which have been going as well as hoped. Many of these have been labeled problem-ridden and costly experiments.[5] After more than a decade of effort, just 32 states plus the District of Columbia had integrated three or more benefits programs. Most states (30) focused on the "big three" programs: Medicaid, SNAP, and TANF. The cost and complexity of delivering even a minimally integrated system for these three programs has meant that other programs have had to take a back seat. Only 4 states, for example, have integrated the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).[6]

Is GenAI an opportunity or simply another complication?

States that are locked into hard-coded, purpose-built IES solutions may also be locked into the "preintegrated" GenAI solutions their IES vendor has selected for them. But as many states begin to rethink their approach to IES, the door opens far wider to other AI and GenAI solutions. More modern cloud-native approaches to IES are designed from the ground up for simplified integration of any system, including AI and GenAI. Whatever the status of their IES efforts, states should seize the opportunity to think about AI and GenAI in the context of integrated eligibility—and perhaps rethink their current IES approach.

[5] Source: Deborah Yetter, "New benefit system a costly experiment," Louisville Courier Journal, April 15, 2016
[6] Source: "What the online enrollment experience for safety net benefits looks like across America," Code for America benefits enrollment field guide, 2023, May 26, 2023

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.

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're not just thought leaders. We're helping government organizations design and implement real use cases today. Our wide-ranging capabilities, from strategy to technology implementation, help us understand the challenges you face holistically. Our extensive network and strategic alignment with leading AI solution providers can give us ahead-of-the-curve insights and help enable us to be surgical in our approach.

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 us

Amiran Gelashvili
Managing Director, Advisory
Health and Government Solutions
KPMG LLP
816-256-1137
agelashvili@kpmg.com

Tom Briggs
Director, Advisory
Health and Government Solutions
KPMG LLP
518-728-2782
thomasbriggs@kpmg.com

read.kpmg.us/modgov