Case Studies in GenAI Successes and Pain Points

About half of public-sector employees say they engage with AI applications almost every day, with generative AI (GenAI) being the most-used one, according to "EY Pulse Survey: insights into the integration of AI in government," published in August. That pushes the almost 2-year-old technology squarely out of nascency, although it still holds plenty of potential for growth.

During a recent GovLoop online event, "Beyond IT: Using GenAI the Right Way at Your Agency," thought leaders from government and industry shared use cases for GenAI, where challenges persist and best practices for using the technology.

Participants

Jonathan Alboum, Federal Chief Technology Officer (CTO), ServiceNow

Michael Berkholtz, Senior Manager for Technology Lifecycle Services, Office of Enterprise Technology Solutions, Office of Information Technology Category, Federal Acquisition Service (FAS), U.S. General Services Administration (GSA)

Brian Peretti, CTO, U.S. Department of the Treasury

R. Spencer Schaefer, PharmD, Chief AI Officer VISN 15 AI Solution Architect NAII, Department of Veterans Affairs (VA)

Jane Yang, Lead Adviser for Artificial Intelligence and Data, Administration for Children and Families, U.S. Department of Health and Human Services (HHS)

Leaning Into AI

As with any technology, GenAI works the same way for all agencies, but each uses it in specific ways to meet its individual needs. For instance, Michael Berkholtz said FAS use cases focus on how GenAI can help with contracts. One way is through chatbots that can answer questions from vendors and internal workers, and another is to translate governmentwide contracts.

So far, GenAI use at VA has centered on getting comfortable with the technology, said R. Spencer Schaefer. "Given the concern and risk we have about patient care and hallucinations [which are incorrect or misleading results], we initially dove into leveraging its capabilities on doing data mining and data summarization," he said.

Rights and privacy are also top of mind at HHS, said Jane Yang, so the agency started with lower-risk uses, such as reviewing and summarizing public comments on proposed policies and rules. "The large language models have really accelerated our ability to do some of that parsing for us and tag information by topic, so that then we can give our policy experts on lead exposure or on teacher qualification … just the comments that are relevant to them as well as a sumçmary," Yang said.

GenAI helped Treasury net a big win: saving $4 billion by preventing fraudulent payments from going out or recovering those already paid, said Brian Peretti. "By being able to use AI to identify these kind of transactions, we'll hopefully be able to reduce that fraud number, to be able to save the money for the taxpayer," he said.

GenAI's Growing Pains

For Yang, a major concern is bias in the data that feeds GenAI models. "If you're training things on biased data, you're going to get biased results," they said. "There's a lot of risk management and change management and training that goes into ensuring that our workforce can responsibly use AI in a helpful way that supports their work."

Orchestration is becoming a pain point for VA, Schaefer said. "I can't staff 100 people just to watch all the large language models that are running to see how accurately they're performing," he said. "So, it's working through these orchestration layers that help in an automated way."

Workforce skills are also a sticking point, Peretti said. When someone's job can be automated, what happens to the worker? "Those people hopefully will be able to upskill, to be able to really add the value that a human can to the process," he said. "But the challenge is also that some people may not be able to be upskilled or don't want to be upskilled. How do we make sure those folks do not fall behind?"

ServiceNow's Jonathan Alboum wants agencies to get the most out of their investments in automation. It's important for agencies to understand how work and data flow through an agency to maximize automation's effectiveness, he said. "We get all excited about new technologies and plan to either further automate or use generative AI to expedite some kind of process, but we're not operating in an ideal environment. We need to be ready to modernize the process, not just the technology. That's the hard work," Alboum said.

Click here to watch the entire session on demand.

govloop servicenow

Best Practices for GenAI

The first step to using GenAI is demystifying it, Berkholtz said. One way to do that is to let people safely test it. To that end, FAS created sandboxes outside its network where employees can experiment with models using public data. Additionally, in April, GSA published a resource guide for the acquisition workforce on using GenAI.

Collaboration has been key to HHS's successes with GenAI. "Every single generative AI project that we have done has only been successful when we've had the strong pairing of the technical folks with the policy and subject-matter experts and the processes," Yang said, adding that agencies also should "trade notes" on how they use the technology.

Ultimately a GenAI investment must prove its value, Alboum said.

"AI has the ability to really change the way service delivery happens. When you make service delivery more effective, you can get your people focused on more complex assignments, not the mundane tasks," he said. "But it's very hard, in my experience, to be able to tell that value story after the fact. You need to know what it looks like today, so when you measure the future, you can demonstrate those improvements."