Tackling A Foundational Challenge: Generative AI in Federal Spaces

Highlights from a Roundtable hosted by the Advanced Technology Academic Research Center (ATARC) in partnership with Google, November 2023

Generative AI is a rapidly developing field with the potential to transform the way federal agencies operate. However, there are also a number of risks associated with the use of generative AI, including the potential for bias, misinformation, and cybersecurity threats.

The Advanced Technology Academic Research Center (ATARC), in partnership with Google, brought together experts from government, industry, and academia to discuss the challenges and opportunities of generative AI use in the Federal government.

Current Status of AI Use in Federal Government

Agencies represented on the panel are working to leverage AI in a variety of use cases, typically dependent on an agency's mission. Many plan to use AI to enhance the citizen experience, better understand data, and optimize web content. Others intend to use advanced AI algorithms to provide enhanced anomaly detection and simulate fraud scenarios.

In just a year, agencies have already witnessed huge advances in AI that are helping to address complex issues. In the field of healthcare, companies are pushing forward with AI enhanced drug development tools, which are shortening the amount of time to identify drug targets and effective drug structures.

However, there are also agencies that are simply looking forward to the adoption of generative AI tools to help the workforce with research, queries, contracting, text-generation, and coding. These small adjustments can add up to improve the overall operational efficiency of the organization.

"There's so much transformational potential with generative AI, but also the potential for massive risk if not implemented properly. As we harness AI's potential, we must ensure responsible use. This is crucial for maintaining public trust and ethical governance."

The Importance of Data

Roundtable participants note that it's tempting to quickly implement a product or service without fully considering total cost, resource requirements, and data preparation needed to successfully use AI. Not doing so can result in unintended consequences. Participants concur that the first step to promoting ethical and safe AI use is to ensure data quality.

Agencies should be cognizant of the rules protecting various forms of data, especially if data is owned by other agencies or governed by the laws of other countries. Disparate data regulations must be considered when training AI models.

Panelists can easily foresee the consequences of even the slightest error in data handling. For some agencies, the inadvertent introduction of incorrect or protected data into an AI model can pollute an entire dataset, making AI outputs unreliable and potentially harmful.

> **"AI is nothing without data. With all kinds of deep fakes and advancements in AI, things are becoming complicated. We need boundaries."**

The concept of 'garbage in, garbage out' was repeated numerous times throughout the roundtable discussion. Participants reinforced the importance of introducing clean, quality data to AI models to ensure outputs are factual and to reduce the potential for bias. The importance of quality data is underscored by the fact that certain AI models simply will not work if there are holes in datasets, and other models will invent, or hallucinate, information if faulty data is introduced.

Thankfully, one participant noted that generative AI models typically do not require large datasets to function effectively. Agencies need a few quality data sets that are aligned with mission goals to experience the benefits of generative AI. However, agency requirements differ based on mission. Some agencies on the panel need AI to access data from countless datasets in order to form appropriate and useful outputs. Risk goes up exponentially the more data is introduced.

Agencies are focused on developing governance frameworks to reduce as much risk as possible. Ensuring AI is used safely, ethically, and in alignment with agency principles is of paramount importance to every panelist. Unfortunately, the challenges associated with this endeavor are widespread and complex, and differ from one AI use case to another.

The Critical Role of Humans

Panelists agree that humans are the key to attenuating some of the risks associated with AI, which means training, upskilling, and recruiting AI talent is critical to successful AI use. A particularly large problem is finding skilled talent and attracting them to work for the government. One panelist cited a recent Stanford study, which shows that approximately 60% of students graduating with a PhD in artificial intelligence go to work in the private sector, 30% stay in academia, and less than 1% choose to work in government.

Panelists suggest the government must take a different approach to recruit and retain talent skilled in AI technology. For instance, establishing labs, incubators, or centers of excellence to help attract talent interested in driving meaningful change. Others recommend investing in existing talent, since they already know the business of government and the nuanced challenges associated with their profession.

Yet, as one panelist noted, the field of generative AI is so new, there are no current textbooks written on the subject. With such a huge demand for AI talent from an industry in its infancy, there are simply not enough experts available. As such, the government must train and grow experts from their existing talent.

By upskilling existing subject matter experts on generative AI, the government can bring more value to individual roles while filling a critical staffing need. Doing so also reinforces the role of AI in the workplace as a tool to enhance the work of existing employees, not to replace them with technology.

AI Use Cases

- Help understaffed healthcare practitioners to work more efficiently and diagnose more accurately.
- Make the constituent experience better
- Empower government employees to make them better at their jobs
- Translate documents quickly with ease
- Leverage data to accomplish strategic objectives much more quickly
- Turn existing content into plain language for better comprehension of complex information

Navigating Risk Tolerance

Panelists believe they are entering into another realm of shadow IT, where employees will use public generative AI platforms, like OpenAI and Google's Bard, in their daily work regardless of agency guardrails. This is similar to employees using personal devices for government work – a practice that's prohibited, yet popular. Agencies should plan for this sort of shadow IT behavior by giving employees a safe environment to use generative AI platforms in accordance with agency terms.

Agencies can approach AI risk from several perspectives. Panelists discuss compliance risk, ethical risk, legal and privacy risk, and operational risk; the latter of which is not as widely discussed as other forms of risk. Agencies should consider the implications and impacts of when an AI model should break. Since the technology is so new, it's hard for agencies to understand the operational impacts of a broken AI model.

Since AI use cases are so variable, it's likely agencies will need to conduct separate risk assessments for each solution. To ensure this happens consistently and rigorously, agencies must begin to establish a strong risk management culture within the workforce.

> **"It's an interesting balancing act, and there's no one solution to this problem."**

Final Thoughts

> **"AI is coming. We can't turn a blind eye or act like it won't impact us. Organizations need to embrace it, and understand the risk and power of it. People have to be involved in this."**

- Agencies must approach the training of AI models as if teaching a child. Expect it to wander off and do something unexpected as a toddler would. Humans must be involved to monitor and correct undesirable behaviors and outcomes.
- Agencies should pay attention to other things that could be impacted by AI, since the technology is mainstream and becoming more powerful by the day. Security profiles are suddenly more vulnerable as attackers gain access to tools they didn't have a year ago. As attackers become faster and more agile, it's even more important for agencies to move quickly to zero-trust models.
- Agencies must be confident enough in their data management practices to ensure the AI training models are producing quality and accurate outputs.
- Agencies should start identifying the low-hanging fruit to test the risk tolerance of AI on the organization. This way agencies can get a sense of what AI functionality is in a low risk environment.