

How to Use Data, Training to Maximize AI

Artificial intelligence (AI) has dominated discussions, especially as private-sector use of it becomes more popular. But in the public sector, challenges with data sourcing, sharing, training and determining the best ways to use the technology persist. As a result, government agencies continue to tread cautiously while still establishing use cases.

During a recent GovLoop roundtable, government and industry experts talked about these struggles and shared some ways they're working to overcome them. Here are highlights from their conversation.

Participants

Harry Dreany

*Intelligent Automation Division
(H60), Naval Surface Warfare
Center Dahlgren Division*

Rebecca McWhite

*Technical Lead, Cybersecurity
Supply Chain Risk Management,
Computer Security Division,
Information Technology
Laboratory, National Institute of
Standards and Technology (NIST)*

Jim Tunnessen

*Chief Information Officer
and Chief AI Officer, National
Endowment for the Arts*

James Valverde Jr.

*Director and Chief Analytics
Officer, Data & Analytics Division,
Office of Naval Research*

David Yu

*Chief Architect of Data and AI
Federal Defense Department
(DoD), Civilian, Intel Customer
Success Unit, Microsoft*

Data Is the Heart of AI

It's impossible to talk about AI without talking about data, because that's what informs the models. A crucial but challenging aspect of that, though, is identifying where the data comes from — for instance, internal vs. external data sources and authoritative vs. certified — and sharing it to work together, said James Valverde Jr. of the Office of Naval Research.

"People are protective about their data," Valverde said. "They view data in zero-sum terms."

To overcome this, the Naval Surface Warfare Center Dahlgren Division focuses on AI's tactical aspects, viewing it as a form of intelligence, said Harry Dreany of the Naval Surface Warfare Center Dahlgren Division. "Especially with the growth of generative AI (GenAI), which we want to embrace ... if you don't have the right information to feed that capability, then all of a sudden, we put ourselves at a disadvantage," Dreany said.

Still, the Navy, like many government entities, must be careful because the data it needs for AI models has different classification levels. "If our data is on the edge and it is seen, which it will be, it will be targeted," he said. "We're really at a difficult crossroads about how much we embrace the data that we have on the edge trying to use this technology."

NIST's Rebecca McWhite highlighted the importance of being aware of the data supply chain. "The data's provenance, meaning its chronological origin and its pedigree ... can often denote bias in the output of AI," McWhite said, noting news reports that a Chinese-developed large language model might not produce accurate information when queried against politically sensitive questions. "The bottom line to me here is that [data] being a fundamental component of AI means that you have to now consider data's supply chain, and how do you trace that?"



Training Not Just AI, But Users

Understanding and sharing data appropriately are just two challenges with using this burgeoning technology. Another is training the employees who implement and use AI today and preparing for future users, too, said Jim Tunnessen of the National Endowment for the Arts.

To that end, he predicts that the CIO role will include the responsibility of something like a chief human capital officer for digital personnel. That's because the CIO will need to balance AI training and workers' fears about losing their jobs to AI with excitement about potential new tasks they can tackle with newfound free time due to efficiency. For instance, if AI can do work that contractors now handle, they can be let go and the budget will shift accordingly.

"Personnel or the leadership ... concerned that in losing those contractors, they will not be as efficient, you share that you can reallocate

those funds that you were paying for those contractors to accomplish X, to accomplish Y. There's another bucket of money that has now all of a sudden become available for you to expand your portfolio," Tunnessen said.

Microsoft's David Yu emphasized the need for AI to be at the core of all training, adding that training should happen on a recurring basis because smarter AI models are becoming available at faster and faster intervals.

What's more, he said, it's important to differentiate between DoD and civilian agencies because the AI stakes are higher with DoD. "If something goes wrong with those applications, then you're talking about life or death," Yu said. "On the civilian side ... it might be a presentation that was produced but it's not accurate. That is a big impact, but it's not to the level of life and death."

AI at Work

Theorizing about how AI can help is one thing, but putting it into action is quite another.

Dreany said his organization has made strides with incorporating GenAI into preventive and predictive vehicle maintenance. It can provide data on a particular vehicle, including how many miles it's traveled and on what terrain, and predict when, say, the drive shaft will break. "I'm going to go ahead and push this part to you, and you'll have it when it does break," he said. "That's been a game-changer."

McWhite said GenAI can help analysts better understand the supply chain when conducting research. "Look at the sources to see where the information actually came from," she said. Asking GenAI, "Where can I find this information?" and validating the source of the information can save analysts' and reviewers' time.

At the Office of Naval Research, Valverde chooses use cases based on four factors:

- **Sense-making:** Determining needs
- **Preferences:** Understanding tradeoffs based on what's needed now vs. in the future and risk vs. innovation
- **The search for the optima:** This is what he calls "the quest for allocative efficiency"
- **Action:** How you execute on those decisions

"You can't be involved in everything, so you have to be highly selective, highly discriminant in your ability to parse out and figure out what those things are that are going to have maximal impact on future naval power. The use cases come out of that," he said.

Data Governance's Role

Data is key to AI, but governance is key to data. The first thing to ask, Tunnessen said, is about the expectations for using AI appropriately. “We need good data cleanliness, we need good data structure, we need to have it properly tagged and the appropriate permissions put in place in order to assure that people can access only what they need,” he said.

Still, Yu said, agencies must be realistic. “We were talking about this 25 years ago. We’re still talking about it. That means we’re not going to solve this problem completely,” he said. “You cannot be aiming for perfection.”

Instead, look at the tools and practices for improving data governance and preparing for AI, because they always evolve.

What’s more, frame data governance as a people, process and culture challenge, he said. Success comes when senior leaders champion data governance and there is a grassroots effort to embrace it. “Without you sharing data, you’re not going to have this AI advantage,” Yu said. “There needs to be a grassroots effort to have a community that encourages people to share [and] tag metadata and derive a data tier where it can be trustworthy and certified so that the enterprise can use it.”

Conclusion

AI holds great promise for the public sector today and into the future. Officials are aware of the unique challenges the technology presents and are working through them with an eye toward security, the current and future workforce, and data management. With such careful consideration, they are building a solid AI foundation with use cases that will inform more efforts down the road.

[LEARN MORE](#)

