Harnessing Gen AI and Data Strategy for Federal Agencies

Highlights from a Roundtable hosted by the Advanced Technology Academic Research Center (ATARC) in partnership with AWS, October 2024

In a recent roundtable discussion, Federal experts discussed the challenges and successes with developing and implementing data strategies. Participants shared insight into crafting data strategies, the evolving role of Chief Data Officers (CDOs), the importance of strong leadership, and leveraging data and Gen AI for mission outcomes.

Importance of a Well-Defined Data Strategy

Participants discussed the importance of a well-defined data strategy that aligns with agency mission to achieve outcomes and further innovation. One agency on the panel has successfully followed a data strategy for several years and is seeing good progress. The key components of their data strategy include infrastructure, governance, and a data catalog.

The strategy aligns with the office structure and data management plan, and each initiative is tied to the strategy. Each year, a data management board reviews and approves each initiative, ensuring it's tied to the strategy. The strategy helps define where employees can share and deposit information, influences hiring decisions, and ultimately how projects are executed.
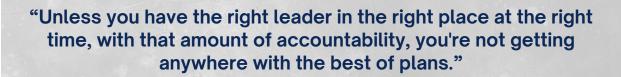
> "You have to set up the structure. It's easy to say you're going to implement this stuff, and then no one ever does. You have to have checks and balances, accountability."

However, other agencies on the panel are not as far along, and are just now starting to build an enterprise data platform, due in part to a lack of cohesive data strategy within the agency. Data is being tagged, named, and used differently across the agency, creating duplicative data sets and a lack of accountability. There is no overarching strategy to guide how data can strategically be used to further all mission areas. There is also a gap between data strategy and execution in many agencies, resulting in data being used in 'smokestacks' by separate teams.

Participants contend that ultimately, data strategies need to be built around a purposeful set of data services tied to critical use cases, whether that's a regulatory requirement, improving transparency, data sharing, or analytics. Committing to a single technology vendor or one kind of technology should not constitute a data strategy.

The Evolving Role of CDOs

Participants also discussed the evolving role of CDOs as data strategy and management become essential to mission outcomes and innovation. The CDO role has expanded beyond data governance to encompass strategic planning, security, regulation, and the adoption of new technologies.

> **"Unless you have the right leader in the right place at the right time, with that amount of accountability, you're not getting anywhere with the best of plans."**

Panelists emphasized the importance of strong leadership to implement data strategies effectively. One agency on the panel shared some of the challenges they are currently experiencing with leadership decisions being at odds with how to carry out initiatives safely and effectively.

For some agencies, there is difficulty relaying information from the top of the organization down to those implementing initiatives. Often, decisions at the top are made without including the people who are doing the work, which effectively bypasses a critical step to understanding how best to protect data. Communicating the mission down to employees implementing the work is crucial for not only how the work is ultimately conducted, but also the safety and security of data.

Others on the panel have found success by asking senior leaders to choose initiatives based on the level of risk they are willing to accept. Understanding how initiatives impact data security and access can help leaders make more informed decisions to further mission outcomes and innovation.

Data Sharing

Participants noted that successful data sharing starts with cataloging data and creating templates to ensure MOUs are written consistently. While authority issues may still arise, technologists should be focused on data protection and including data protection clauses into MOUs.

Thinking more globally, some panelists contend that agencies should work towards a true open data model where data is shared and can be leveraged by all agencies, instead of individual agencies guarding data behind MOUs.

Gen AI: Potential Benefits and Risks

> **"Rogue agents are using those technologies to get ahead of us in some ways, so it's extremely important to embrace new technologies in a very governed and secure manner."**

Participants acknowledged the potential for AI to improve efficiency and decision-making, but also emphasized the importance of responsible AI use and data security. Ultimately, the quality of AI depends on the quality of data. For some agencies, there is significant work to be done to cleanse data so that it can be used at the enterprise level. Currently, entities across the agency are tagging the same datasets completely differently, which causes bigger issues when the data is sent to the enterprise platform.

Inconsistencies and errors in data can cause significant issues with AI outputs, such as hallucinations and bias. Not only does this result in unreliable AI systems and outputs, but it increases the risk of releasing misinformation to the public, especially if AI outputs are not fact-checked by humans.

> **"Gaining the trust of the American public is extremely important for all Federal agencies. If we lose that, just by the tools that we put on our website, that would be a bad thing."**

Other concerns raised by participants are the potential HR risks and employee privacy issues introduced by integrated AI functions within popular workforce products. Some claim that AI can monitor employee activity, which could potentially infringe on employee rights.

Creating Policies for AI

Panelists agree that they do not know enough about the challenges or risks of AI well enough to write sound, detailed policies to protect data or guide future use. Essentially, agencies don't yet know the problems they need to solve with a policy.

Despite the unknowns, participants underscored the importance of ensuring policies and AI use are designed for fairness to all taxpayers. Humans must remain in the loop to ensure fairness of AI usage, and AI systems should provide mechanisms of explainability to provide context and sources of all outputs.

> **"We have got to be fair to all taxpayers in every instance, whether it's enforcement or service. And how do you know that you're being fair? You have to have that human eye on it after the fact."**

Looking Ahead

> **"At the end of the day, the data is our product. We need to protect that."**

When asked about what comes next, panelists shared their plans to start using data to tell stories. Agencies have access to a wealth of data that can be used to better inform the public on certain topics. Others are still working to establish data strategies, but have aspirations to also use data in strategic and helpful ways.

LEARN MORE AT:

HTTPS://AWS.AMAZON.COM