

# 3 tips for navigating data challenges in automation

Government leaders leverage data as a key enabler for modernization, but ensuring safety and productivity requires modern infrastructure, security, and collaboration between public and private sectors.



“When we can analyze questions on a technical level and reach answers without actually having to ever exchange data, that’s when we can really see change in a secure way.”

Taka Ariga  
Chief Data Scientist, Government  
Accountability Office

## Introduction

Government organizations have embarked on a transformative journey, but navigating the landscape of automation is complex, and industry leaders are wary of the risks associated with artificial intelligence (AI) in federal programs.

Data is a fundamental enabler in the modernization process, and its role becomes indispensable for IT officials striving to create an effective and transparent government. Yet, realizing this vision requires innovative solutions, resilience, and a commitment to responsible data practices. As a result, leaders across the public sector are adopting a cautious yet forward-looking approach to digital transformation.

During the [Data In Action Summit](#) hosted by GovExec and Informatica, public sector experts discussed how they’re using data to balance productivity, security and citizen demands in the age of data-driven governance. In sessions bookended by keynote speakers Taka Ariga, chief data scientist at the Government Accountability Office, and Ken Pfeil, chief data officer for the Commonwealth of Virginia, the panelists explored how their organizations are harnessing data to deliver transformative analysis and security solutions.

### 1. Address challenges in data sharing

According to Ariga, one of the persistent challenges GAO faces is the sharing of data within the federal government. Traditional methods of data exchange often encounter statutory restrictions, making it a cumbersome process.

To address this, GAO proposed a revolutionary solution: explore data sharing without exchanging data custody. With this approach, instead of transferring the ownership or custody of the actual datasets, the process involves leveraging cloud services to facilitate the exchange of insights or analytical results without the need to share the raw, sensitive or confidential data itself. This forward-thinking approach not only overcomes legal barriers but also ensures a secure and compliant data-sharing ecosystem.

“When we can analyze questions on a technical level and reach answers without actually having to ever exchange data, that’s when we can really see change in a secure way,” Ariga said.

In Virginia, the [Commonwealth Data Trust](#), established under the Office of Data Governance and Analytics, acts as the governing authority for data across the state. Setting standards for the 63 executive branch agencies, the trust fosters a collaborative environment around data sharing that enables key initiatives to be successful.

One example involves collaborating with localities, Virginia State Police, and other stakeholders to address high crime rates in 13 localities. Pfeil explained that through data consolidation, extraction

“You need to have the right labeling, the right tagging, all the blocking and tackling done to facilitate quality of data. Because if you take an intelligent model and you put bad data into it, you’re going to continually make that model stupid and nobody is going to benefit from it.”

Ken Pfeil  
Chief Data Officer,  
Commonwealth of Virginia

and visualizations, the team improved decision-making and even developed a predictive model for Virginia State Police involvement, contributing to the abatement of violent crimes.

Overall, data sharing without exchanging data custody is an innovative solution that addresses the challenges of sharing sensitive or restricted data while promoting collaboration.

## 2. Emphasize data literacy

In the dynamic landscape of AI, Ariga and Pfeil agreed that data literacy is an integral part of successful automation. Recognizing that algorithms yield results based on probabilities is critical for employees to distinguish between accurate information and potential inaccuracies.

To Ariga, solid foundational governance practices are key for successful AI use. For example, in an experiment with a large language model, GAO found significant inaccuracies in the application’s responses. Without robust data literacy initiatives, the misinterpretation of that response could have significant consequences.

Virginia’s CDO echoed this concern, emphasizing the well-known saying, “garbage in, garbage out.” This sentiment highlights the pivotal role of foundational governance practices, data curation and quality in ensuring the success of AI projects.

“You need to have the right labeling, the right tagging, all the blocking and tackling done to facilitate quality of data,” Pfeil said. “Because if you take an intelligent model and you put bad data into it, you’re going to continually make that model stupid and nobody is going to benefit from it.”

The message is clear: Utilizing subpar data with an intelligent model degrades the model’s performance; only high-quality data will unlock meaningful, trusted benefits from AI applications.

## 3. Take a thoughtful approach to automation

As organizations navigate the complexities of AI adoption, Ariga and Pfeil recognize both the potential benefits and risks linked with AI in federal programs, sharing a vigilant yet forward-looking approach to automation.

Security concerns are paramount for both entities. GAO, an overseer of AI implementation in federal programs, opts for secure experimentation within its boundaries through its [Innovation Lab](#) to ensure the secure implementation of AI.

“What we’re really focused on is the oversight aspect of AI implementation,” said Ariga, “We want to make sure that as we are embarking on this digital transformation, we get the basics around

“We want to make sure that as we are embarking on this digital transformation, we get the basics around data, hygiene data, quality data layer and data literacy down on our end, so that we can start really exploring the downstream impacts of AI.”

Taka Ariga  
Chief Data Scientist, Government  
Accountability Office

data, hygiene data, quality data layer and data literacy down on our end, so that we can start really exploring the downstream impacts of AI.”

Pfeil advocated for a similar approach to AI adoption, acknowledging the importance of a collaborative strategy for establishing effective guardrails. For instance, collaboration with IT organizations is key to assessing the long-term complexity and costs associated with maintaining and scaling AI solutions.

The goal is to integrate controls that ensure secure AI implementation, making AI a powerful ally rather than a potential threat. With the right partners to navigate security aspects effectively, leaders and decision makers can realize the benefits of AI in a timely manner. And by focusing on foundational aspects like data quality, governance, sharing and literacy, agencies can position themselves at the forefront of technological advancements.

*Learn more about how [Informatica](#) is helping agencies maximize their data.*



**Worldwide Headquarters** 2100 Seaport Blvd., Redwood City, CA 94063, USA Phone: 650.385.5000, Toll-free in the US: 1.800.653.3871

© Copyright Informatica LLC 2021. Informatica, the Informatica logo, Intelligent Data Management Cloud, Informatica Intelligent Cloud Services, CLAIRE, and Big Data Management are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners. The information in this documentation is subject to change without notice and provided "AS IS" without warranty of any kind, express or implied.