8 Ways to Put Data to Work in Your Agency

Data is a lot like exercise equipment. However much of it you collect, it won't do you any good unless you actually put it to use. In the case of data, a lot of the heavy lifting goes into getting yourself ready to use it and figuring out how. That was the focus of a recent GovLoop virtual summit, Data in Action: How Your Agency Can Put Information to Work.

This report highlights some important takeaways from the event, including:

- Building blocks for a data-ready organization

- Ways to get more value out of data

- How artificial intelligence could change things quickly

Red Hat   FORTINET FEDERAL   alteryx   govloop   elastic | aws   CHECK POINT

3

Building Blocks of the Data-Ready Organization

As organizations look to expand their use of data and shift toward data-driven decision-making, they often find they need to think about data in new ways. In particular, they must focus on making data accessible and useful to people who are not data experts. Here are three factors to consider.

1. Break Data Into Meaningful Chunks

How much data is too much data? It depends on who's looking at it.

For example, the Department of Veterans Affairs (VA) sends out more than 150,000 surveys every week to veterans who have used their services. The results provide VA leadership with a lot of information about veterans' customer experiences across the country.

But leaders at individual medical centers need a much smaller slice of that data, said Evan Albert, the VA's Director of Measurement and Data Analytics.

So Albert's team created a module that enables leaders across the VA to look at trends specific to their centers or at specific services within their centers — e.g., problem areas in neurology over the last 90 days.

"Many of our customers have Monday morning reports they provide to their senior leadership," Albert said. By providing these tools, "we help them look like superstars," he said.

2. Build Out Your Data Team

To make data useful for big questions, you need to build big teams.

For starters, you need a platform that enables your data experts to work together, said Adam Clater, Chief Architect for North American Public Sector at Red Hat, which provides enterprise open-source solutions. That is especially important when it comes to building and training AI, in part because of the scale of the work involved.

Getting something to work on a laptop or desktop computer is very different from getting it to work in a larger server environment, Clater said. "So that's one thing we're working on: How do we build a collaboration platform where data scientists can work together to solve these problems faster and more accurately?" he said.

But your data team also should include non-experts, such as program and project managers and subject-matter experts.

In particular, subject-matter experts can help data scientists get a better understanding of the problems they are trying to solve. Such "domain specificity," as it is called, can improve the quality of an AI model, Clater said.

3. Don't Let Manual Processes Hold You Back

Data involves a lot of drudge work. It's all essential, but it's also tedious and time-consuming.

In particular, employees often spend an inordinate amount of time doing the preliminary work of ETL: extracting data from various sources, transforming it into usable formats and loading it into a data lake or warehouse.

"If it takes me all week to put together a report, is that really a valuable use of my time?" said Andy MacIsaac, Solutions Marketing Director for Public Sector at Alteryx, which provides a data automation platform.

The good news is that those tedious tasks are perfect candidates for automation, being repetitive and rules-driven, MacIsaac said.

Automation also reduces the opportunity for errors. Because the tasks are repetitive, it's easy for people to lose focus and make mistakes. Automation doesn't make typos.

Alteryx provides a platform that enables organizations to automate tedious tasks across the entire analytics lifecycle. And because it is no-code/low-code, analysts can automate those tasks without manually writing code, MacIsaac said.

The result? They get to focus on the work they love doing.

"It's really about enabling analysts to unleash their creativity, their problem-solving and to really apply themselves to deeper, more valuable work," MacIsaac said.

3

Ways to Get More Value Out of Data

In the long term, agencies get more value out of data when they create policies, processes and technologies that serve as guardrails for enterprise-level initiatives. Here are three ways to drive those initiatives forward.

1. Come to Terms on Data Governance

The easiest way to ensure the privacy and security of a dataset is to keep it under lock and key. Don't share it with anyone, don't integrate it with other datasets and so on. That, of course, would defeat the purpose of an enterprise data strategy. And that is why data governance is so important.

Data governance provides rules of the road for ensuring the availability, usability, integrity and security of data, and those rules need to be applied however that data is used or shared. For example, if data is classified as personally identifiable information (PII), the appropriate policies need to be applied wherever it goes.

In 2021, the Virginia Legislature created the Office of Data Governance and Analytics. This office plays a critical role in helping agencies coordinate data-sharing across the commonwealth, said Marcus Thornton, the commonwealth's Deputy Chief Data Officer.

To institutionalize that process, Thornton's office created the Commonwealth Data Trust, an information-sharing environment. Agencies abide by a standard set of requirements around security, privacy and confidentiality to become part of the trust.

"Our agency is here to help facilitate the management and sharing of data in a very safe and secure way," he said.

2. Collect Data With a Purpose in Mind

Data hoarding doesn't help anyone.

If an agency stockpiles data without a clear idea of how to use it, it will spend a lot of money to ingest and store it and will see little return on its investment, said Ian Lee, High-Performance Computing Security Architect at Lawrence Livermore National Laboratory.

For example, a lot of agencies collect and retain event logs generated by their networks and systems. In theory, such logs can help agencies detect, investigate and respond to cyber threats. But in many cases, cyber teams did not use that data. An August 2021 executive order on event logging, though, has changed the dynamics.

The executive order and follow-on guidance "is pushing agencies to level up" by focusing on how to use data to address specific challenges, such as incident response, Lee said.

To address those challenges, however, agencies need to ensure that all data is readily available when needed, said George Teas, Vice President of Solution Architecture at Elastic, whose Elasticsearch Platform is used for enterprise search, security and observability.

"We see a lot of our customers who have different datasets in different silos, which really limits their ability to take use of that data," Teas said.

The cloud can play a critical role in helping agencies capture that data and make it accessible, said Adam Cerini, Principal Solutions Architect at Amazon Web Services.

That can make a big difference when using a platform like Elastic for performance monitoring. By running it in the cloud, you can get alerts in real time "and shorten an outage from hours to minutes," Cerini said.

3. Develop a Data Framework

Aaron Rose of Check Point said he knows a data framework sounds like a dry topic but it's also essential.

A basic framework should identify:

- What data you have
- Who owns it
- How that data should be classified (sensitive but unclassified, classified, PII, etc.)

As simple as it sounds, this kind of framework goes a long way toward helping an agency map out a data protection strategy, said Rose, who is a Cybersecurity Evangelist and Security Architect at Check Point, which provides hardware and software security solutions.

And it's not necessarily as simple as saying classified data is much more important than anything at a lower level. For example, agencies store vast amounts of PII about their employees. Such data has become especially vulnerable in recent years because of ransomware, with cyber criminals using the data they capture to extort employees or even agencies. That risk should be reflected in the framework, Rose said.

The framework also might include information on who typically accesses that data, and how. Such information provides a baseline for behavioral analytics, Rose said. For example, let's say that HR personnel typically access employee data through a given HR application or through a web server. If someone outside HR is trying to access that app through an unknown server, the system can flag it for investigation.
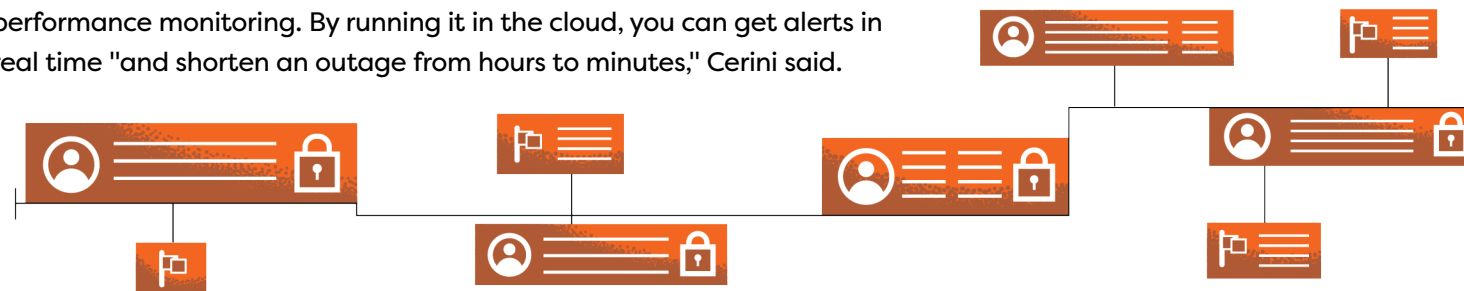
2 Ways AI Could Change Things Quickly

Over the last half-year, we've all been getting a crash course in the power and limitations of AI. Data is essential to AI, providing input that drives AI models; while AI is essential to data, providing new ways to analyze large volumes of it and to distill results. Here are two examples.
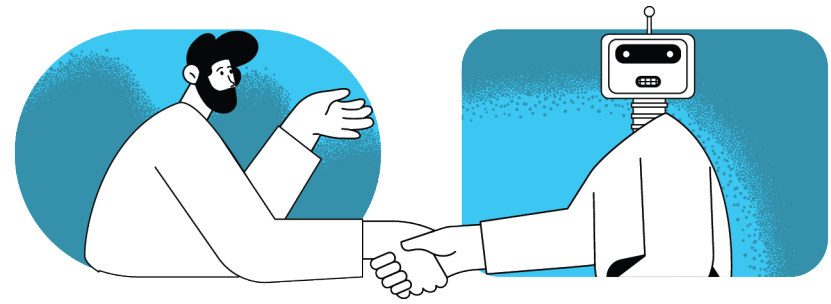
1. Make Non-Experts Feel Like Experts

"Hey, Siri, can you create a pivot table for me?"

One of the obstacles to data-driven decision-making is that often the people making decisions are not proficient at using data tools. AI could be a game-changer, said Gerald Caron, Chief Information Officer at the International Trade Administration.

Chief data officers already have been trying to provide people with tools that are easier to use, so people can do a certain amount of analysis without having to call in the experts. But AI and natural language processing are proving their ability to make it easier for non-experts to conduct analysis and understand the results, Caron said.

Here's how it might work:

- Rather than writing a structured query language (SQL) query – something most people cannot do – you could ask a question in common terms, such as, "Between 2021 and 2022, in what areas did the help desk see the biggest increases in requests?"

- AI would translate that question into an actual query and provide the results, again in plain language: "The number of trouble tickets jumped more than 50% in three areas ..."

- Based on the results, you might ask a series of follow-up questions and, in short order, have all the information you need to put together a budget request for new software.

"AI's been providing us more information in a more natural form, rather than us doing that hunting and pecking and trial and error," Caron said. "And I can talk to it naturally like I'm talking to you. ... It's just amazing."

2. Make Solutions Smarter About Your Agency

Vendors have already begun incorporating AI into their products, enabling agencies to take advantage of the technology without having any real expertise. For example, numerous cyber solutions rely on AI and machine learning to automate the process of detecting, analyzing and even responding to threats. So far, so good.

But with coming cyber solutions, AI will get even smarter.

Think about how AI works (and forgive the over-simplification). You train a program how to think (i.e., you create an AI model) by feeding it a whole lot of data, then you keep refining that model until it works as intended.

In the case of commercial products, the vendor takes care of training and refining the model. But soon, agencies could have the opportunity to further train that AI on its own data, said Felipe Fernandez, Chief Technology Officer at Fortinet Federal, which provides a wide array of enterprise-grade security solutions.

That data will give AI a more accurate baseline for how things work in your agency, and that, in turn, will improve its ability to detect an anomaly that might signal a threat.

With that added intelligence, AI should become much more proficient at stopping threats and conducting forensics before humans need to get involved, Fernandez predicted.

Given the cybersecurity skills gap across all levels of government, "we need to be able to automate some of the operational capabilities and functions that humans are handling now," he said. "I think that's where we're going to see AI take center stage in the next two to three years."
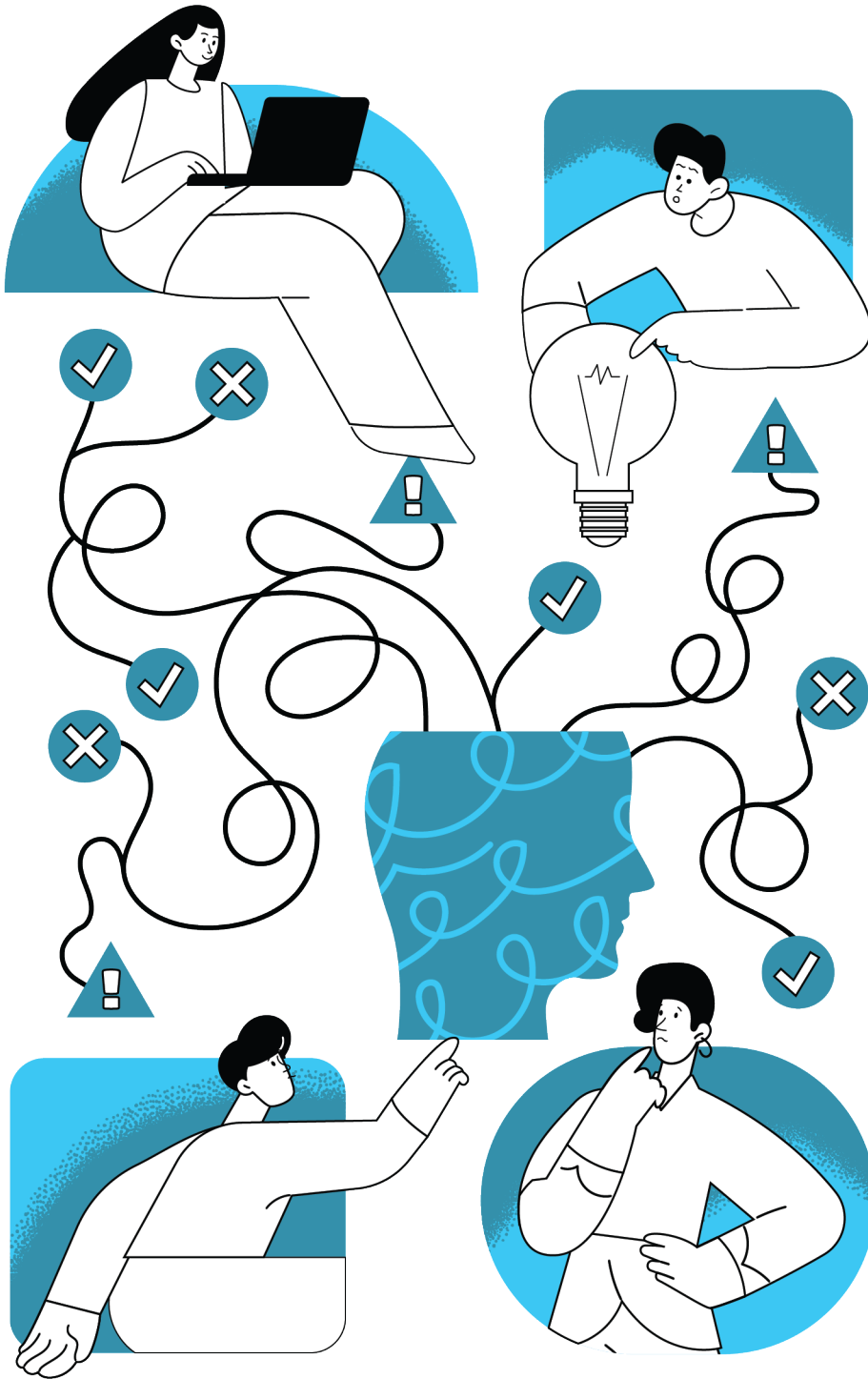
To get more insights from our subject-matter experts,
watch the entire event on demand.

Special thanks to the sponsors who made this event possible: Red Hat, Alteryx, Elastic, AWS, Check Point and Fortinet Federal.
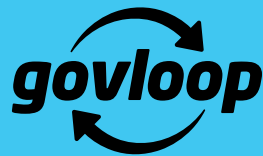
About GovLoop

GovLoop's mission is to inspire public sector professionals by serving as the knowledge network for government. GovLoop connects more than 300,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington, D.C., with a team of dedicated professionals who share a commitment to the public sector.

For more information about this report, please reach out to info@govloop.com.

www.govloop.com | @GovLoop