MeriTalk
Improving the Outcomes of Government IT

HITACHI
Inspire the Next

ISSUE BRIEF

Turn Data Into Accelerated Insights and Mission Results With DataOps

Data is undoubtedly the government's biggest asset today, both strategically and tactically. At the same time, data is one of government's biggest challenges.

Federal agencies have vast stores of structured and unstructured data, and those volumes are growing exponentially. Transformation of valuable data stores from static, unused assets into powerful, real-time intelligence is critical to digital transformation efforts, which seek to fundamentally overhaul operations for greater business and mission value.

Harnessed effectively, data can enable agencies to make better decisions, faster – even in real time. It can enable policy that better serves the American people. Agencies are working toward these goals, backed by the Federal Data Strategy, agency and department data strategies, data acts and regulations, and the President's Management Agenda.

DataOps can help agencies make the best use of their data. It is an agile and automated approach to data management that is designed to facilitate organizational, data-driven decision-making. Similar to the purpose of DevOps in software development, DataOps endeavors to eliminate silos and encourage collaboration across IT operations, data management, and software development, so that agency data can be used effectively to achieve positive and reliable mission outcomes.

Rather than a specific implementation, DataOps is a methodology that helps bring structure to chaotic and expanding data stores – across data formats, platforms, business units, and entire agencies.

Similar to DevOps a few decades back, DataOps is in a growth phase today, particularly as Federal agencies define data management processes and infuse them as data culture throughout their organizations. They are working to ensure that DataOps becomes an organic, day-to-day process, rather than an add-on activity. The need is urgent: The more time that teams spend looking for and assessing data, the less time they spend developing solutions to meet critical agency needs.

Desire to leverage artificial intelligence (AI) and machine learning (ML) for mission success is a big driver for DataOps. Agencies have realized that AI and ML applications are only as good as the data that feeds their models. As a result, they are refining their data management practices so their AI/ML models get the highest-quality, most appropriate data for the mission. Their end goal is the ability to extract real-time analytics and drive operational efficiency with trusted, quality data.

A Comprehensive DataOps Solution

A successful DataOps strategy relies on people, processes, and technologies working together to address data governance, data integration, data cataloging, metadata management, and data analysis. Hitachi Vantara Federal addresses these core DataOps needs with the Pentaho DataOps suite, which integrates vast stores of structured and unstructured data into a unified management framework for analysis and visualization.

The comprehensive Pentaho Intelligent DataOps Platform:

- Simplifies data discovery
- Manages data pipelines at enterprise scale
- Optimizes data movement across edge, core, and any cloud provider
- Discovers and manages metadata for accurate data lineage and data provenance, enabling reusability of datasets
- Helps protect sensitive data
- Simplifies and speeds data analytics

With Pentaho DataOps, agencies can make data-driven decisions at speed and scale.

Pentaho DataOps at a Glance

Unlike DataOps solutions that are designed to meet a specific requirement, Pentaho DataOps provides a framework that can support any data management and analysis need. Agencies can deploy the entire suite or only the components they need to address the existing gaps in their data architecture. Open architecture enabled via application program interfaces (APIs) and hundreds of prebuilt data connectors, data transformation stages, and multiple runtime engines mean Pentaho DataOps isn't a rip-and-replace proposition. It works with existing data infrastructure, so agencies can build a data fabric that provides a unified view of data, as well as enterprise scale data governance and control.

Pentaho DataOps includes:

Pentaho Data Catalog	• Automatically inventory all data. • Use AI to automate data governance. • Enable data self-service.
Pentaho Data Integration	• Ingest and distribute IT, OT, and IoT data. • Filter and protect sensitive data. • Orchestrate data flows for ML training and operationalization.
Pentaho Business Analytics	• Visualize and analyze data across systems. • Deploy interactive analytics, reports, and dashboards. • Embed insights into applications and processes.
Pentaho Data Storage Optimizer	• Cost-optimize Hadoop data lakes with smart tiering.
Pentaho Edge	• Seamlessly integrate IoT and video data. • Deploy ML and advanced analytics at the network edge.
Pentaho Master Data Management	• Ensure data quality, uniformity, accuracy, consistency, stewardship, and governance. • Provide accountability of an enterprise's official, shared master data assets.

Pentaho DataOps in Action

These are just some of the ways Pentaho DataOps can help agencies meet their missions:

- **Records management:** An executive branch organization captures every data asset in real time to not only make data-driven decisions faster, but also to comply with data retention requirements daily if requested, instead of on an annual or biennial basis.

- **Natural resources management:** The State of Arizona employed Pentaho DataOps to centralize, categorize, sort, and analyze trusted data across 330,000 water resources. Pentaho presents the data via an easy-to-use, integrated, and centralized dashboard. As a result, employees can quickly understand the data and have more time to manage groundwater conditions and adjudicate water rights.

- **Predictive maintenance:** A U.S. Navy contractor uses predictive analytics, powered by Pentaho DataOps, to manage fleets more efficiently. Pentaho's data integration and analytics capabilities harmonize data from a wide variety of source formats and enable analysts to construct complex models to detect relationships humans can't see – and identify cost-saving opportunities.

  For example, analysis discovered that ship hulls encrusted with barnacles, seaweed, and other ocean detritus cost $1.3 million per vessel in lost fuel efficiency every two years, or more than $10 million for a fleet of eight. To determine whether the cost of more frequent cleaning would be offset by fuel efficiency, analysts used Pentaho to overlay trend lines of cleaning costs with the penalty of fuel inefficiency and identified an optimal schedule of six hull cleanings every two years. The result was net cost savings of more than $800,000 per ship.

- **Geospatial intelligence:** Pentaho DataOps can consume, integrate, and process data from disparate sources, including situation reports, local news and weather, and video, sensor, and other text-based data. Pentaho DataOps can then visualize and overlay this data on maps of any size for up-to-date geospatial intelligence. The low-code, no-code Pentaho DataOps platform makes geospatial intelligence available to users on demand.

- **Self-service data:** Fannie Mae is the leading source of financing for mortgage lenders in the United States and operates under a congressional charter. It implemented the Pentaho data management and analytics solution to provide a self-service data catalog for business users. Cataloged, custom properties are attached to each dataset, and APIs enable automatic data processing for faster analytics and insights. Millions of files are cataloged each day.

Accomplish Any Mission With Pentaho DataOps

Pentaho DataOps' combination of rule-based systems, ML capabilities, and processing engines automate formerly manual tasks, such as metadata tagging and AI model testing, freeing IT staff for higher-level activities. Its collaboration features drive cross-team engagement and trust in data.

With full visibility into their data, agencies can envision the art of the possible – and make it a reality with Pentaho DataOps. No other DataOps solution can match the comprehensive data management and analysis capabilities of Pentaho.

About Hitachi Vantara Federal

Hitachi Vantara Federal is the trusted leader in mission-centric data solutions for the Federal government. We're a collaborative, full-service company with longstanding OT and IT roots. We empower data-driven insight with a deep bench of integrated partners — advancing Federal customer missions regardless of their data maturity levels. Hitachi Vantara Federal is a FOCI-mitigated subsidiary of Hitachi Vantara.

Hitachi Vantara Federal:

Data-Driven. Mission-Centric. Future-Focused.

To learn more about Hitachi Vantara Federal and Pentaho DataOps, visit:
hitachivantarafederal.com

MeriTalk
Improving the Outcomes of Government IT

HITACHI
Inspire the Next