

**TALK
DATA
TOME**

DataOps: Unleashing the Power of data for Agency Innovation

In Hitachi Vantara Federal's (HVF) first webinar in the "Talk Data to Me" series, Dr. Pragyansmita Nayak (Dr. P), Chief Data Scientist with HVF, and Jonathan Ferguson, Head of Digital Sales for HVF, took a deep dive into the world of High-Performance DataOps and what it means for our federal government and system integrators.

As Dr. P mentioned in this webinar, "DataOps have emerged as an important topic for the federal government to improve its data management analytics and decision-making capabilities for both operational efficiency as well as the national security."

Data Ops In Agencies Today

The Federal Government is comprised of different agencies that help our country, and these agencies have different processes in operating their data. Jonathan mentioned that there is still progress that needs to be made in DataOps. Data is always relevant in the federal space and the data we collect tomorrow can be used to make influential decisions. If your agency is not willing to incorporate a DataOps strategy, then you are probably behind the curve. The conversations we have about DataOps, and data collection are endless and are always important in the federal space. Data collection is the first step of DataOps. You are building a pipeline, or a data pipeline, to integrate. An analogy Jonathan mentioned is what you put in your garbage can, "what you put in your garbage can when it's taken out, is garbage in, garbage out." If you do not have the right toolset to collect the data, it will impact the outcome and any solutions that come after.

The Pentaho Data Catalog

So, what is a data catalog? It is a tool within that pipeline that manages the data governance, and the data lineage; it can also measure and determine where that data came from. When a data analyst or scientist is solving a problem, they are "starting the process of identifying related data assets that will help them solve a given problem," said Dr. P. The Pentaho Data Integration (PDI) has been around for 20+ years now. It provides "the data, blending the data process, and then AI, ML, modeling, ML ops, and a variety of other concepts that are taking place during those pipelines." If your agency needs another data catalog to "support a variety of different complex datasets," Hitachi Data Catalog may be a better fit for you to manage both unstructured and structured data and the associated metadata management which are often non-trivial in nature.

Featured Speakers:

Dr. Pragyansmita Nayak

Chief Data Scientist

Hitachi Vantara Federal



Jonathan Ferguson

Head of Digital Sales

Hitachi Vantara Federal

