Portal page notes

Thursday, October 29, 2020 9:33 PM

Advanced Analytics Workspace (AAW)

The Advanced Analytics Workspace provides options to explore and analyze your data, and turn your insights into sharable solutions. It is built using Open Source and Open Standard technologies, aligning with the latest Treasury Board Secretariat (TBS) policy on information systems. We provide access to the latest data science technologies in a scalable and self-service environment.

- YouTube: https://www.youtube.com/watch?v=quYuuEAqNm0&feature=youtu.be
- Slide deck: https://govcloud.blob.core.windows.net/docs/daaas-cncf.pdf

Analysis

// Data analysis, data management, machine learning

Jupyter Notebooks

Use Jupyter Notebooks, RStudio and Linux desktops for analysis and machine learning development.

Pipelines

Run flexible and scalable data pipelines powered by Argo workflows. See our advanced MLOps pipeline for inspiration.

Object Storage (MinIO)

Powered by MinIO, you can store private and shared files in a centralized object storage environment. Browse, upload/download and share using a web interface or a command line. It is fully S3 API compatible, so it works well with pipelines and other programs. It's also automounted into your Jupyter Notebooks for easy access.

MLFlow

Track your machine learning models through your individual MlFlow central artifact repository.

Advanced Data Storage

Full integration thanks to Goofys with other data storage systems, such as Azure Blob Storage, Azure Data Lakes, and other S3 compatible API's

Platform as a Service (PaaS) Integrations

Use available PaaS services to enhance the tools provided to you by the AAW. Use cloud-provided storage, compute and analysis tools right from your Jupyter Notebooks and Kubeflow Pipelines. Examples include Azure Databricks (Spark), Azure Machine Learning (ML), Azure Data Lakes and more!

Dashboard and API's

// Visualize, publish, host

Shiny Server

Publish interactive dashboards to allow visitors to interact with your data and your analysis.

Seldon Core

Serve your machine learning models via a REST API.

Advanced

// Security and Observability

Secrets generation and management

Powered by Hashicorp Vault, we securely generate and manage credentials for use with various systems. It is used under the hood to provide MinIO credentials to Jupyter Notebooks, Boathouse and other services.

Service Mesh

Powered by Istio, in-cluster services are protected via mutual TLS authentication.

Serverless
Today we offer the Serving component of the Knative serverless platform. In the near future, we will expand this integration to the whole Knative serverless platform.

Visualize and Discover Data
We offer individual ElasticSearch and Kibana instances for your use.

Open Source Contributions

Boathouse
Simplify access to your object storage by making it available in your Jupyter Notebooks.

Jupyter API's
Rewritten backend in Golang featuring numerous user interface improvements and integration with KubeCost.

See https://github.com/StatCan for more open source tools.

Made with ❤️ by the Statistics Canada Cloud Native team and powered by Kubernetes. Learn more.