DATA+AI
SUMMIT 2022

Cloud and Data Science Modernization

of Veterans Affairs Financial Service Center with Azure Databricks

ORGANIZED BY databricks

FSC FINANCIAL SERVICES CENTER

2021 Winner
FedHealthIT
Innovation Awards

2022 Winner
FedHealthIT
Innovation Awards

Presenters Name

Scott Meier, Dept. of Veterans Affairs, Financial Services Center

Cary Moore, Databricks

1

Agenda

- About the Department of Veterans Affairs (VA)
- About the Financial Services Center (FSC)
- Overview of some key accomplishments at the FSC
  - Office of Community Care (OCC) Financial Time Series Forecast
  - CFO Dashboard
  - Migrating Datasets to the Cloud
  - Getting the word out with the VA Hackathon
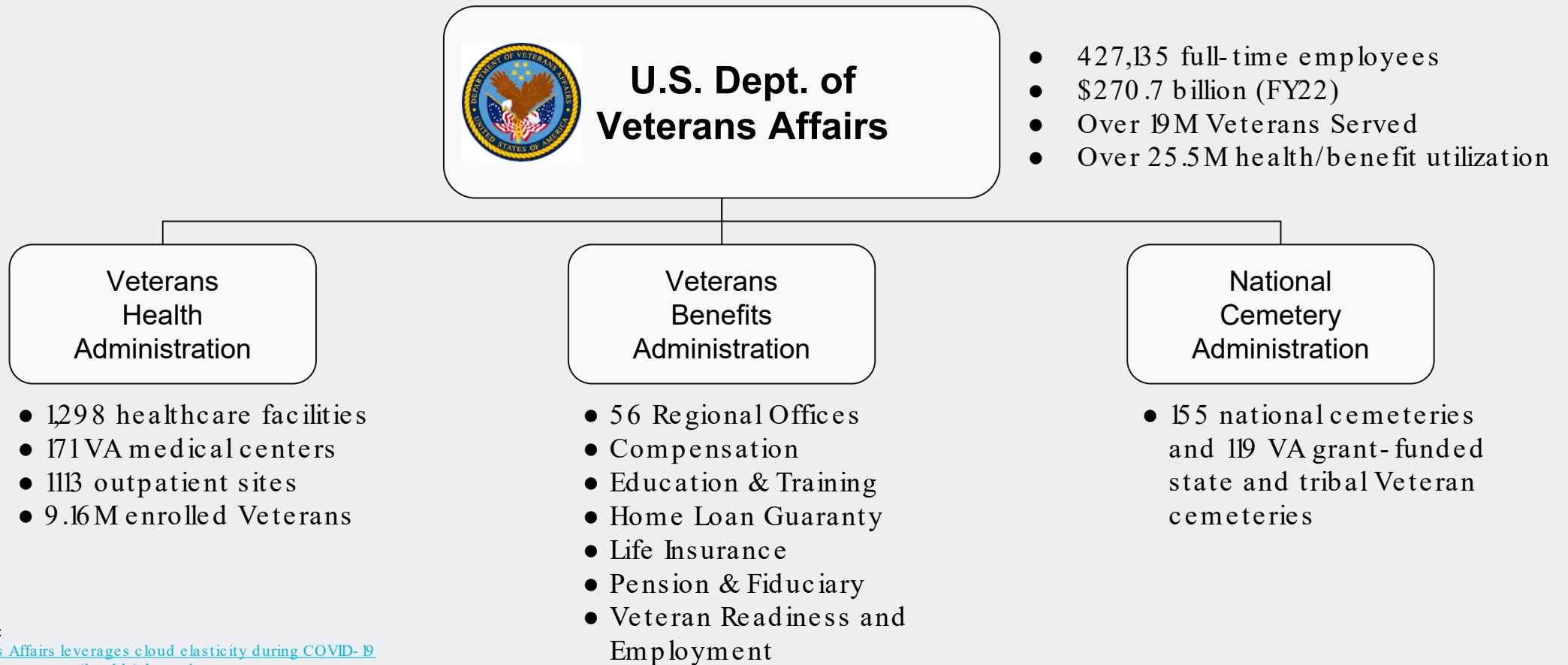
US Department of Veterans Affairs

Our mission, as the Department of Veterans Affairs, is **to care for those "who shall have borne the battle" and for their families, caregivers and survivors**. Our core values focus our minds on our mission of caring and thereby guide our actions toward service to others.

US Department of Veterans Affairs

VA Overview

U.S. Dept. of Veterans Affairs

- 427,135 full-time employees
- $270.7 billion (FY22)
- Over 19M Veterans Served
- Over 25.5M health/benefit utilization

Veterans Health Administration

- 1,298 healthcare facilities
- 171 VA medical centers
- 1,113 outpatient sites
- 9.16M enrolled Veterans

Veterans Benefits Administration

- 56 Regional Offices
- Compensation
- Education & Training
- Home Loan Guaranty
- Life Insurance
- Pension & Fiduciary
- Veteran Readiness and Employment

National Cemetery Administration

- 155 national cemeteries and 119 VA grant-funded state and tribal Veteran cemeteries

Sources:
Veterans Affairs leverages cloud elasticity during COVID-19
https://www.va.gov/health/aboutvha.asp
https://www.benefits.va.gov/BENEFITS/about.asp
https://www.cem.va.gov/cem/about/index.asp
https://www.va.gov/vetdata/docs/pocketcards/fy2021q4.pdf

Figures are as of 5/29/2022

FSC Leadership

Curt Rauhut — Executive Director

Wallace French — Executive Director

Edward Bernard — Deputy Director

Anna Mourad — Chief Experience Officer

Ed Niedzielski — Chief Accounting Officer

Rachael Petteys — Chief Management Officer

Desiree Lake — Chief Fraud, Risk & Compliance

Ron Jackson — Director, Corporate Travel & Charge Card Service

Angela Repka — Director, Financial Healthcare Service

Robert Toney — Director, Financial Operations Service

Brian Huff — Director, Financial Payroll Service

Scott Meier — Director, Data Analytics Service

Crystal Russell — Director, Financial Accounting Service

Clint Loeser — Director, iFAMS Service

Reem Mourad — Director, Strategy & Transformation

Melissa Moore — Director, Business Office

Leena Bhanot — Chief of Staff & Director, IT Business Operations Division

Jilee Rich — Director, Program Management Office

Johnny Lindow — Director, Operations & Maintenance Division

William Eytel — Director, FMBT Technology Support Division

Key — OPERATIONS — OIT

April 18, 2022

Data and Analytics Service (DAS)

Vision

To aid in fulfilling the promises to our Veterans by achieving the highest level of excellence in the practice of **data science and analytics**

Mission

To provide **accurate**, **meaningful**, and **actionable analytics** in clearly understandable formats, providing **new and deeper insights** into business successes and issues, allowing the FSC and its customers to meet their goals.

DAS Support of the VA Mission

- Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act)
- Fraud Reduction and Data Analytics Act of 2015 requires Agencies to establish financial and administrative controls
- Digital Accountability and Transparency Act of 2014 (Data Act)
- Government Charge Card Abuse Prevention Act of 2012
- VA's Enterprise Data Strategy supporting the Federal Data Strategy and Action Plan
- VA Strategic Plan 2018-2024: Data Driven Decision Making and Suicide Prevention
- Secretary of Veterans Affairs Priorities
  - "Seek to Prevent Fraud, Waste and Abuse (STOP FWA) Initiative on April 27, 2017
  - Implementing the MISSION Act
  - Transforming our Business Systems (FMBT, VALOR/DMLSS, COVID Response)

DAS Industry Participation

- 2021 & 2022 FedHealthIT Innovation Award Winner
- Member, VA Data Governance Council, Sub-Council for Analytics
- Chair of VA's Inaugural Data Analytics Work Group for FY21/22
- Founding Member and Chair of the Databricks Federal Executive Council
- Advisory Council Member for University of Texas, McCombs Master of Science in Business Analytics Program
- Partnership for Public Service's AI Federal Leadership Program
- VA's National Artificial Intelligence Institute Partner
- Databricks Strategic Customer Program & Microsoft Fast Track Program
- DoD-VA Joint Data and Analytics Summit Participant
- Partnered with Treasury in "Enhancing Program Integrity Using Data Analytics" Pilot
- Federal Advisory Committee Act (FACA) positive review of analytics program

DAS Services

Interactive Data Visualization	Data Architecture and Governance	Data Modeling	Business Intelligence
Data Analytics Consulting	Report Development	AI & Machine Learning	Data Product Training

DAS Solutions

➔ **Finance**
- ◆ Chief Financial Officers Dashboard
- ◆ Payroll
- ◆ Travel
- ◆ Congress/OMB Reporting on special appropriations

➔ **Healthcare**
- ◆ Data Forecasting (7 models)
- ◆ Enterprise Denials
- ◆ IPPS/eCAMS Duplicate Payments (vendor payments)
- ◆ National Artificial Intelligence Institute - Covid Dashboard

➔ **Supply Chain**
- ◆ Credit Card Systems
- ◆ VHA Supply Chain Common Oper Picture
- ◆ DMLSS / VALOR (enterprise supply chain)
- ◆ National Acquisition Center FMS Logistics Mgmt
- ◆ Electronic Data Interchange
- ◆ Undelivered Orders

➔ **Financial Management Business Transformation (replace VA's FMS)**
- ◆ FMBT Business Intelligence and Reporting
- ◆ Customer Relationship Management

➔ **Internal to FSC**
- ◆ Financial Dashboard (DAS and FPS)
- ◆ Customer Relationship Management Dashboard (FOS and FMBT)
- ◆ Nationwide Vendor Management Portal- NVMP (FOS)
- ◆ Authorized Representative Verification Tool- ARVT (FOS)
- ◆ Recruitment Dashboard (MSS)
- ◆ Procurement/Contract Tracker Dashboard (MSS and FTS)
- ◆ Payment Metrics (FHS)
- ◆ Audit Liaison Dashboard (CFRCO)
- ◆ Sprint Feedback (DAS)
- ◆ PBI Usage (DAS)

DAS Data Sources
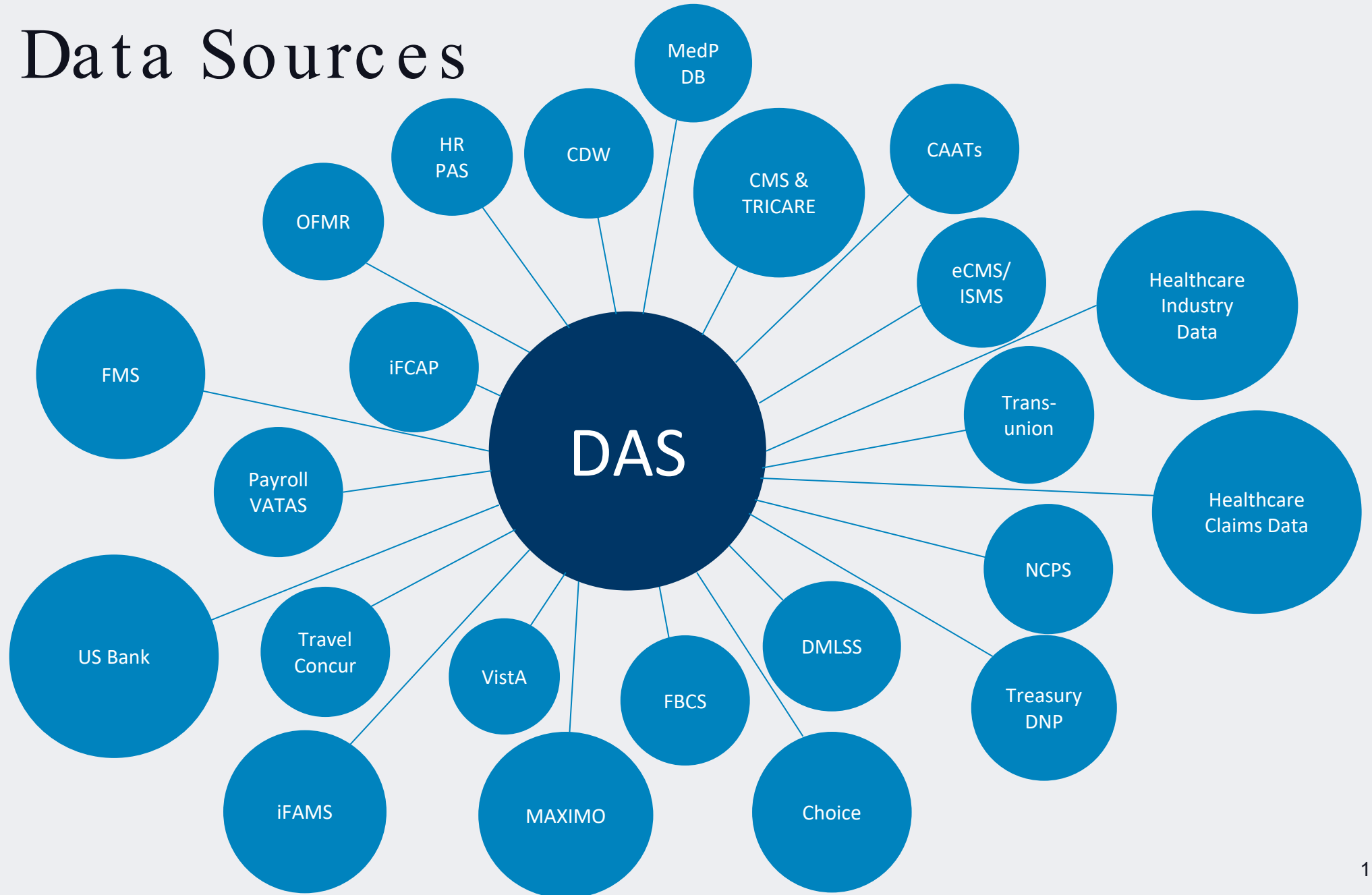
DAS Competencies and Tools

- R / Python / SQL / Oracle
- SAFe Agile DevOps / Lean Six Sigma / PMPs
- Jira / Confluence / Rational / GitHub / Sharepoint
- Slack / Teams
- PowerBI / MS Dynamics / Spotfire / Tableau
- MS Azure Cloud / MS SSAS / Databricks

DAS Technology

- Current State:
  - Predominately on Premises

- Interim State:
  - Cloud/Hybrid

- Future State:
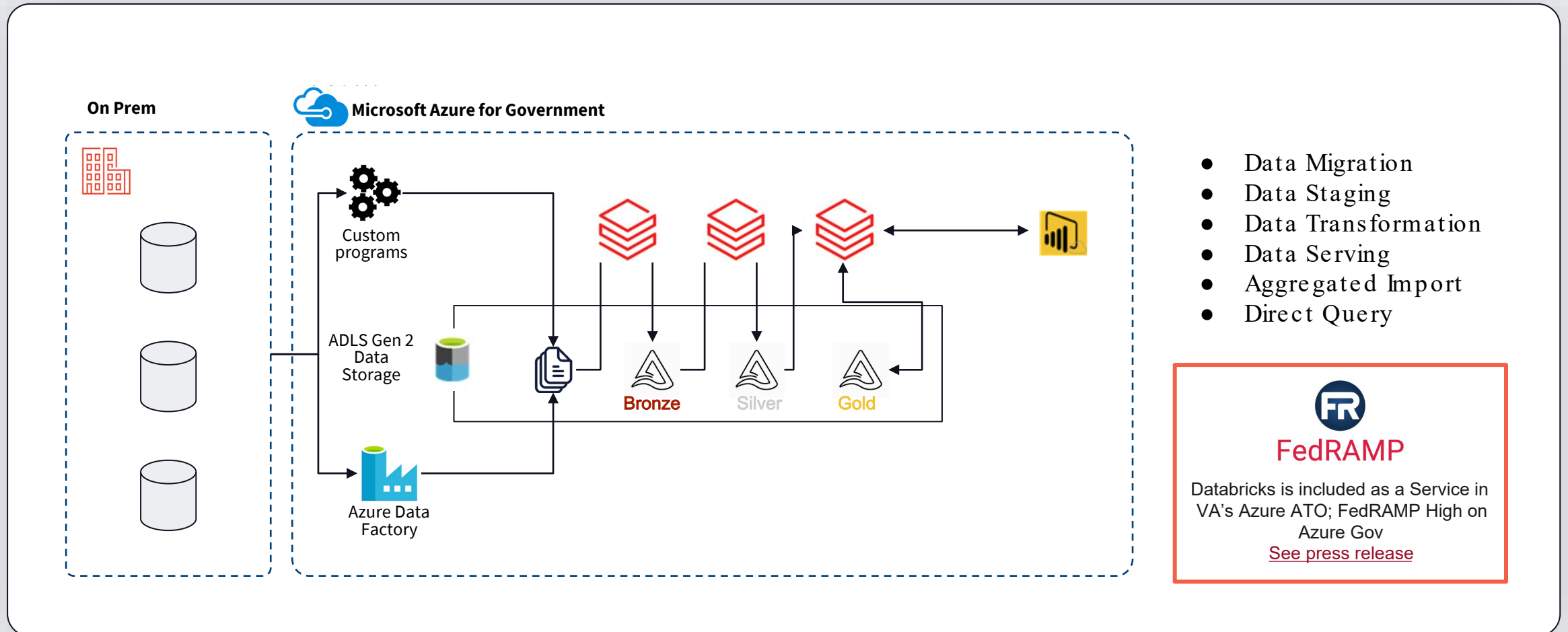 - Full Cloud Architecture

Notional Future State
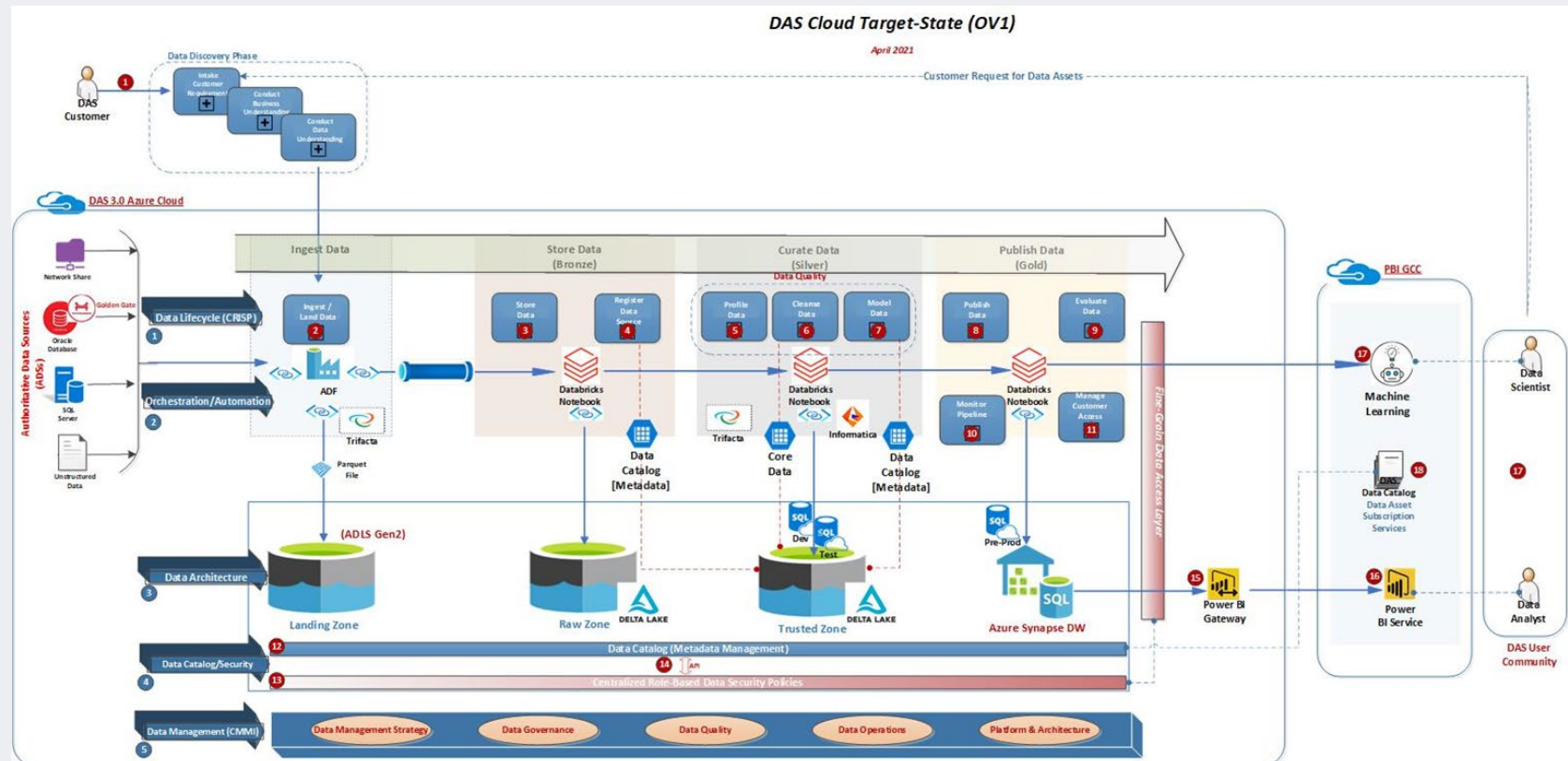
Implementing the Lakehouse at the VA

Overview of accomplishments at the FSC

- Integrated Veteran Care (IVC) Financial Time Series Forecast
- CFO Dashboard
- Migrating Datasets to the Cloud
- Getting the word out with the VA Hackathon

The Lakehouse Architecture at FSC

Implementing Azure Databricks and Delta at the VA

On Prem

Microsoft Azure for Government

Custom programs

ADLS Gen 2 Data Storage

Azure Data Factory

Bronze

Silver

Gold

- Data Migration
- Data Staging
- Data Transformation
- Data Serving
- Aggregated Import
- Direct Query

FedRAMP

Databricks is included as a Service in VA's Azure ATO; FedRAMP High on Azure Gov
See press release

OCC Financial Time Series Forecast

Financial forecasting enhancements to predict claims
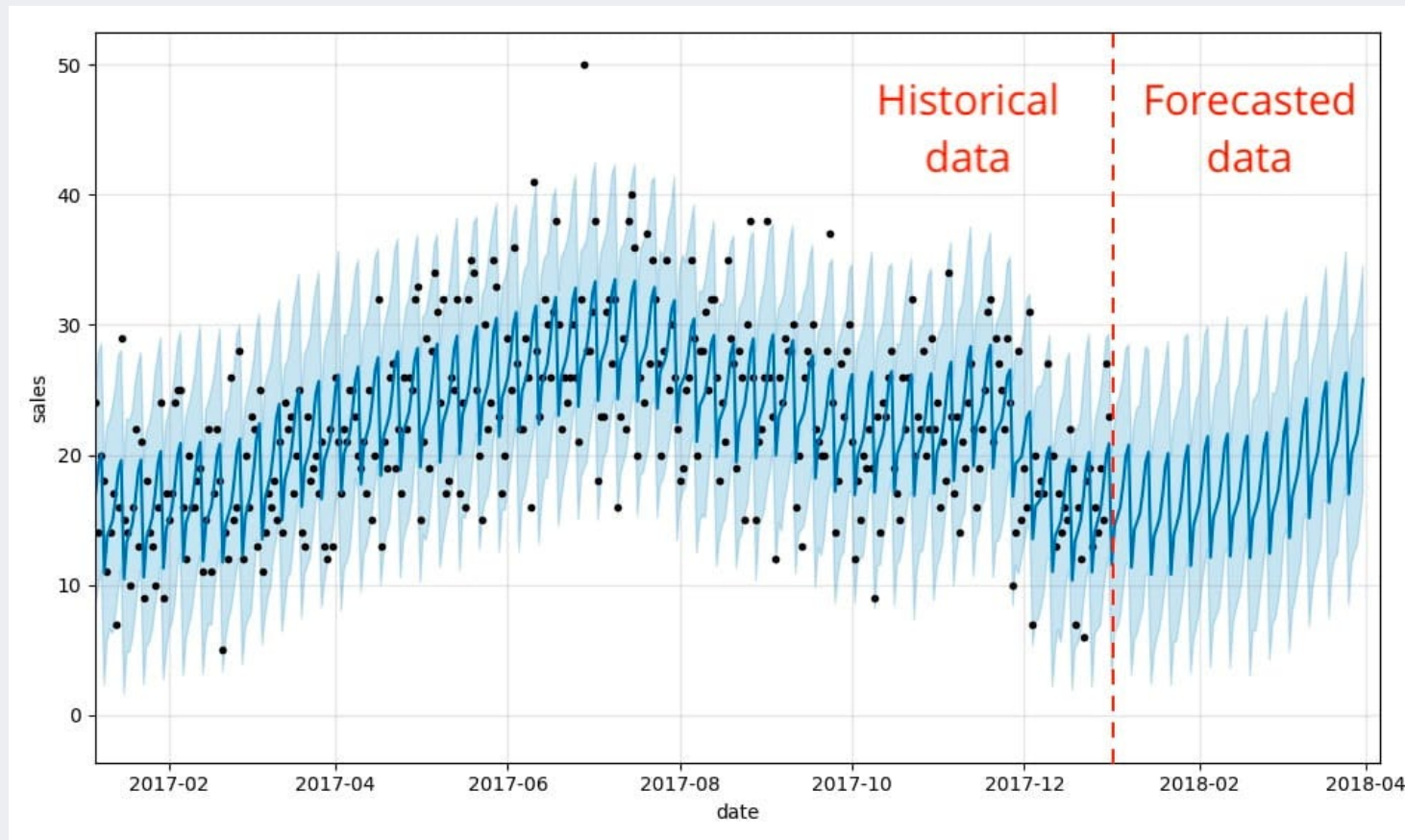
- Predict spending on claims for Congress
- Cost Drivers
  - Outpatient
  - Inpatient
  - Dialysis
  - State Home
- Time Series
  - Prophet
  - ARIMA

Data Science Use Cases at FSC

Implementing AI/ML in Azure Databricks

Claims forecasting and budgeting using Prophet

- Migrated from a single machine used to iterate through several models.
  - Difficult to maintain and run
  - Took several hours to run
  - Could only do a higher level of granularity
- Using Databricks the team was able to improve speed and increase flexibility
  - Leveraged the Facebook Prophet model library - later Python ARIMA
  - Reduced the speed from 1.5 hours to ~20 minutes
  - Faster throughput allowed for more iterations and refinement
  - Leveraged the Spark Pandas User Defined Functions for parallel executions on spark dataframes

CFO Dashboard

Leveraging Azure Databricks to increase performance

Implementing the CFO Dashboard

Taking the Lakehouse to the extreme

Getting the data ready

- Dramatic time reduction with Databricks, Delta and Photon
  - Multi-level ragged hierarchy with up to 5 uneven levels
  - 750 Million Row dimension
  - 1.5 Billion Row Fact table
  - On Premise solution was taking 10 days with failures - lots of effort to maintain
  - Databricks with Delta / Photon reduced time to <3 hours
- Some tuning required
  - Cluster selection - Delta Cache Enabled - Photon - Larger Instance size
  - Use Delta
  - File size optimization
  - ZOrder on Dimensions and Facts
  - Analyze table to gather statistics

Migrating datasets for additional uses

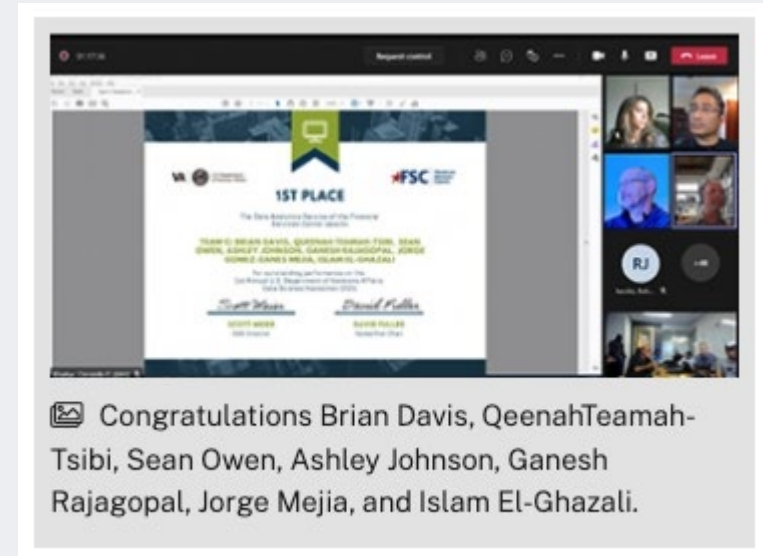Expanding the datasets available for the Lakehouse

Many datasets added and more to come

- Several Data Assets available in the Lakehouse
  - Data for Financial Services Center
  - Data for Office of Connected Care
  - Data for Purchase Card
  - Data for Office of the Inspector General (partial)
  - Data for Supply Chain (partial)
- Technologies used
  - Azure Data Factory - SQL Server On Prem and in the cloud copies
  - Custom Python routines to create parquet files in ADLS Gen 2 (phasing out primarily using ADF)

Getting the word out with the VA Hackathon

Demonstrating the art of the possible with AutoML

- Helped to promote the use of Azure Databricks
- Data from the VA's Office of Integrity and Compliance (OIC) for identify inaccurate coding
- Three teams took two days to transform data, create model and present the results
- 3 Judges from the VA scored the outcome
- Winning team achieved an 87% accuracy
- Featured in a blog article published in the VAntage Point Blog

🖼 Congratulations Brian Davis, QeenahTeamah-Tsibi, Sean Owen, Ashley Johnson, Ganesh Rajagopal, Jorge Mejia, and Islam El-Ghazali.

Thank you!