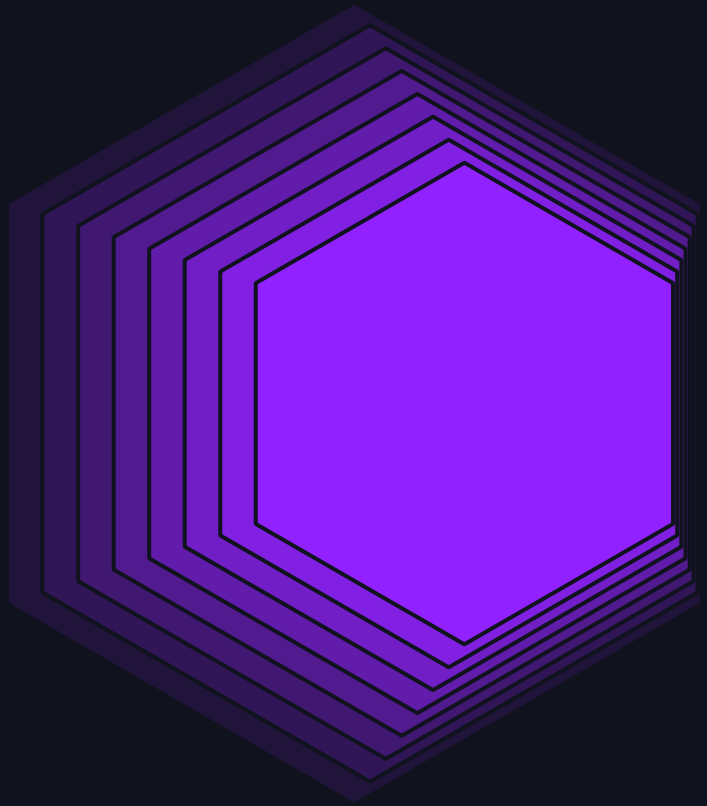
DATA+AI SUMMIT BY databricks

Production Data Apps On Databricks



Cody Austin Davis
6/10/2024

INTRO

My Journey at Databricks

Cody Austin Davis

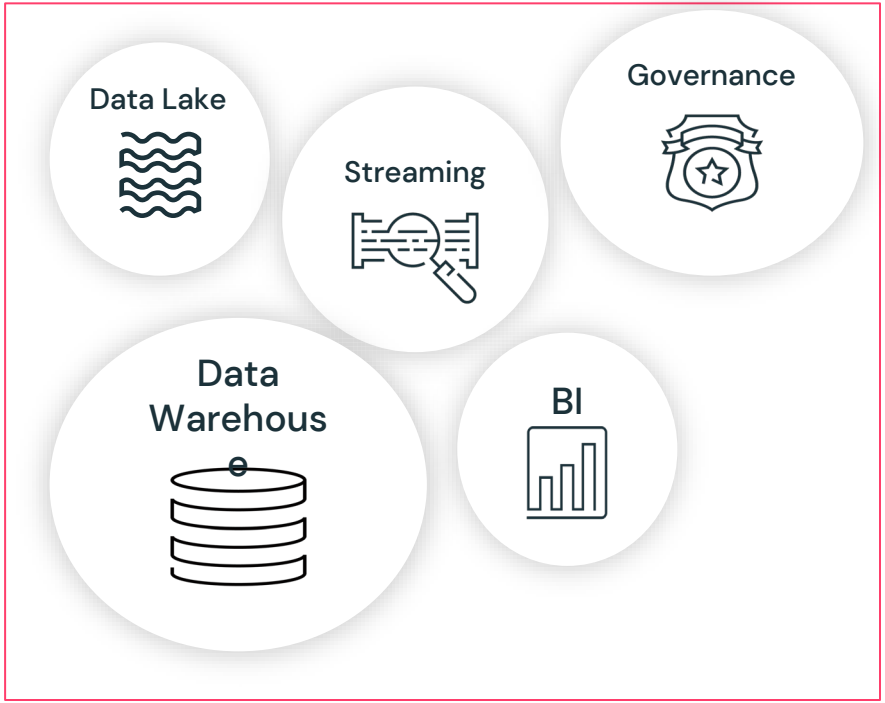
- Senior Solutions Architect @ Databricks
  - ~3 years @ Databricks
  - > 6 years using Databricks

- Solutions Architect for Digital Native Strategic Customer Segment
- Lead the Data Warehousing SME Group @ Databricks
  - Run a Warehousing Blog @ [medium/dbsql-sme-engineering]
- Focus Primarily on building out our Data Warehousing ecosystem and GTM in the field
- Developed Data Products for Healthcare Systems throughout the nation @ Loopback Analytics
- Worked in finance in portfolio optimization & investment analysis

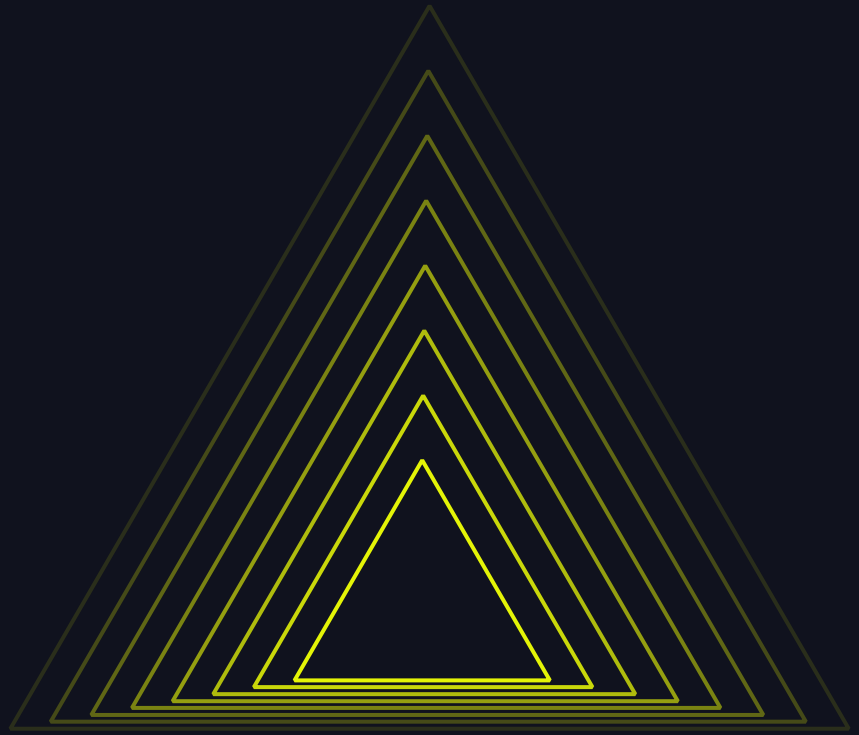
AGENDA

Our Journey to AI-embedded Data Applications on Databricks

- What is a **Data App**?
- What do **Data Apps Unlock** for companies?
- What makes **Databricks + Plotly Dash** the best combo for building Data Apps?
- **Demo** of a Data App for Observability
- What comes next



What is a Data App?



Production Data Applications

Enterprise Data + AI + BI

Connect	Embed	Share
Connect Insight To Action In a Single UX	Embed Insight Into Daily Workflow of Non(less)-Technical End Users	Share Insights & take action with end users right when they need it

DATA+AI SUMMIT
BY databricks

My favorite example of a Data App…

Bloomberg Terminal!

- Data App that helps users generate insights

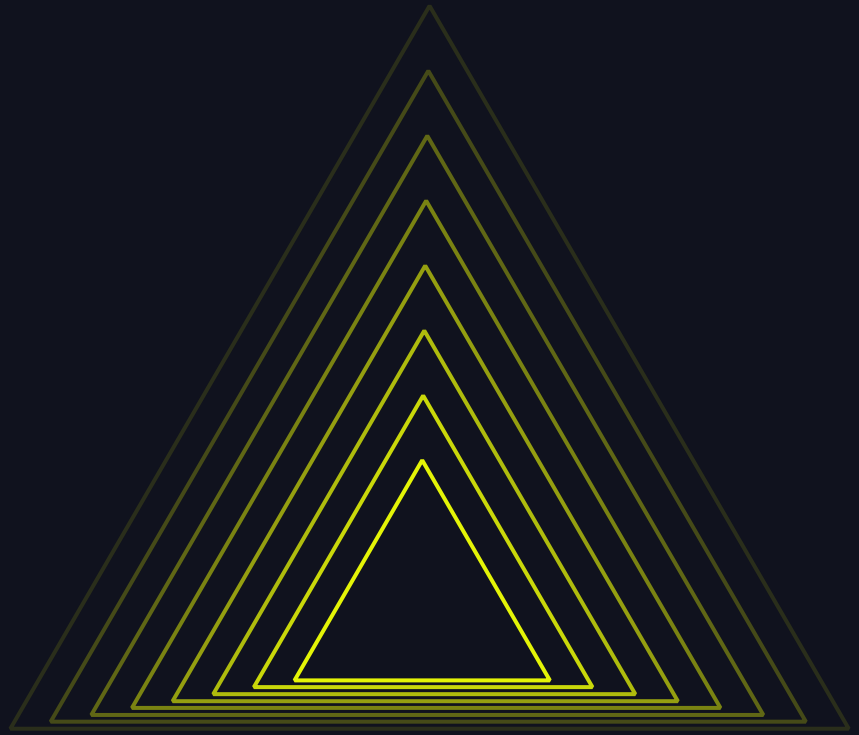
- Injects those insights into actual workflows to take action

- Empowers non-technical users

By @tech at bloomberg, Sept 2023, ref: https://www.bloomberg.com/company/stories/what-took-build-altd-bloomberg-terminal-alternative-data-function/

DATA+AI SUMMIT BY databricks

Why build Data Apps on Databricks with Plotly Dash

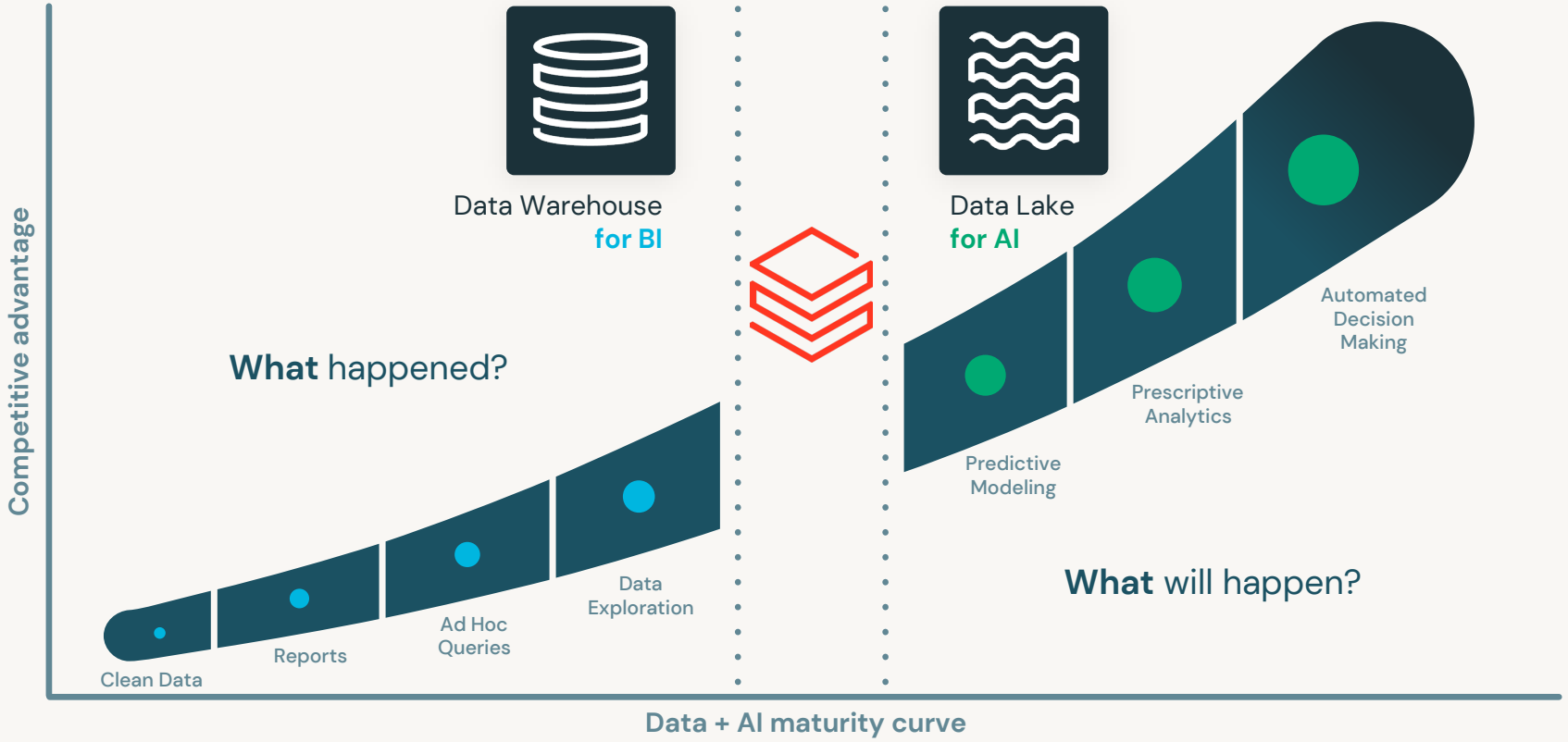


Databricks is a unified stack that business can use

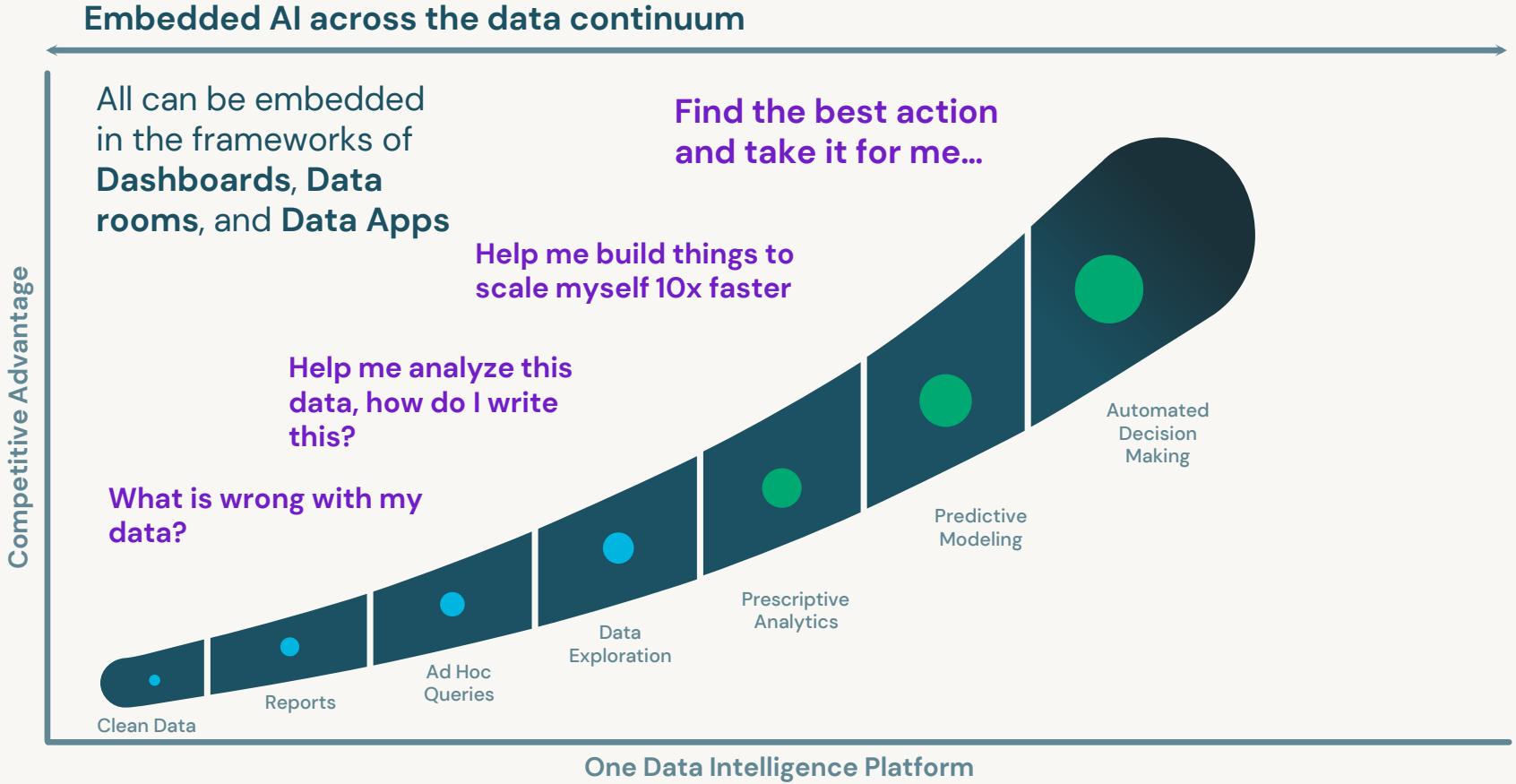
Enterprise Data + AI + BI

BI	AI	Apps
DBSQL Lakeview + BI Ecosystem is rich and AI-embedded	Full ML Lifecycle has been on Databricks for a long time	Everything you need for rich data is is already on Databricks

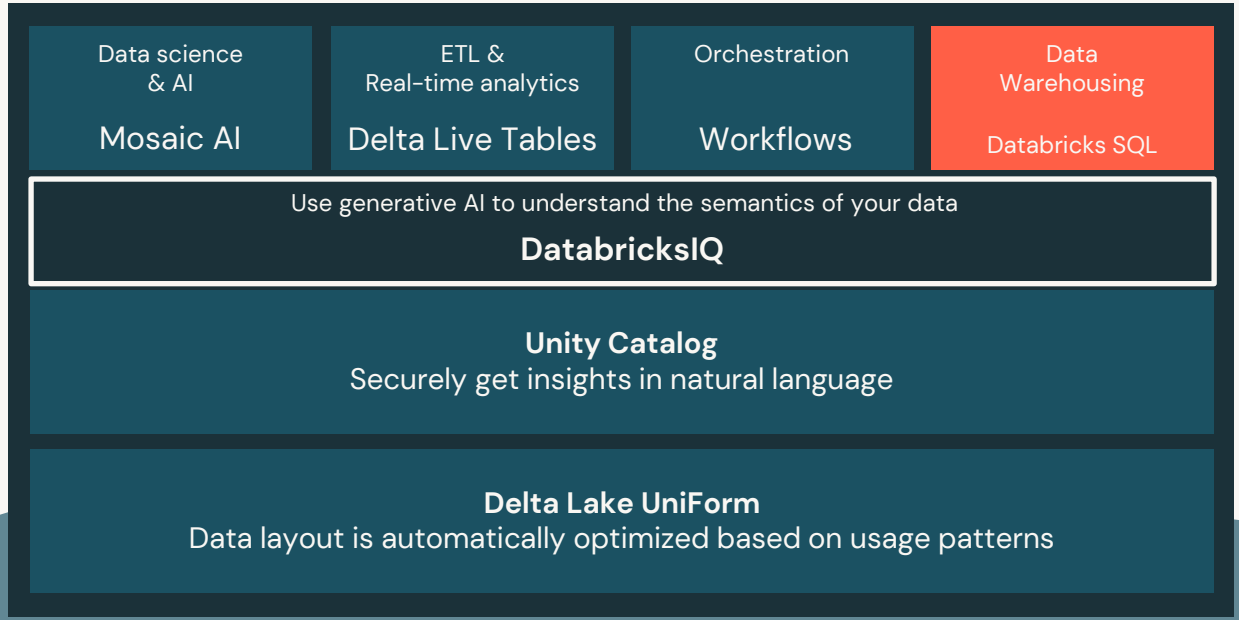
Databricks Connected the Data Maturity Chasm



… And has embedded AI into every stage



Databricks Data Intelligence Platform

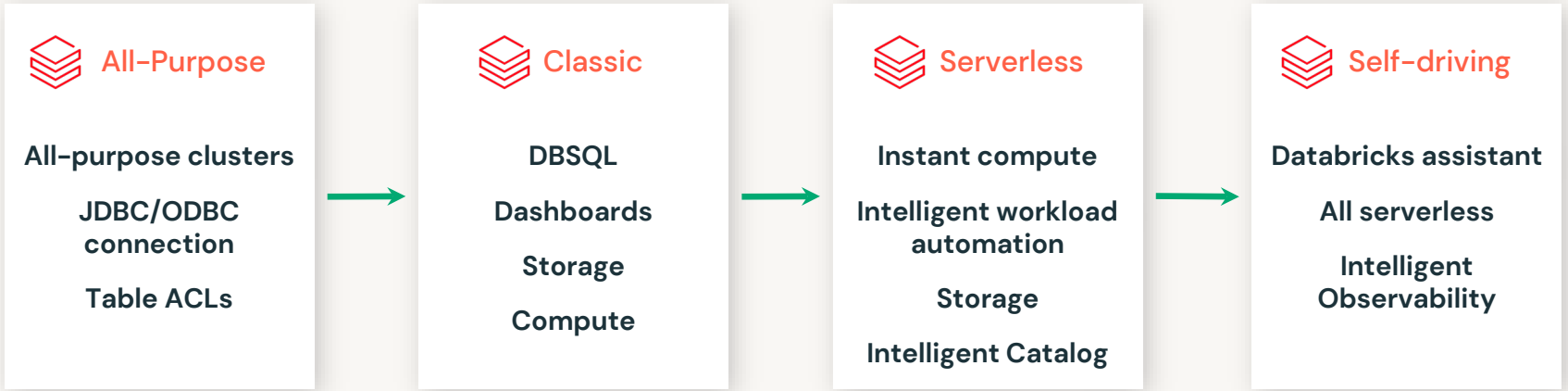


Open Data Lake

All raw data
(Logs, Texts, Audio, Video, Images)

Data Intelligence Platform Evolution

Journey to self-driving in 4 years or less



Your resources in your account
Minimal SQL-native tooling to help

Your resources in your account
Heavy lift & operational burden
Scale limited by your infra

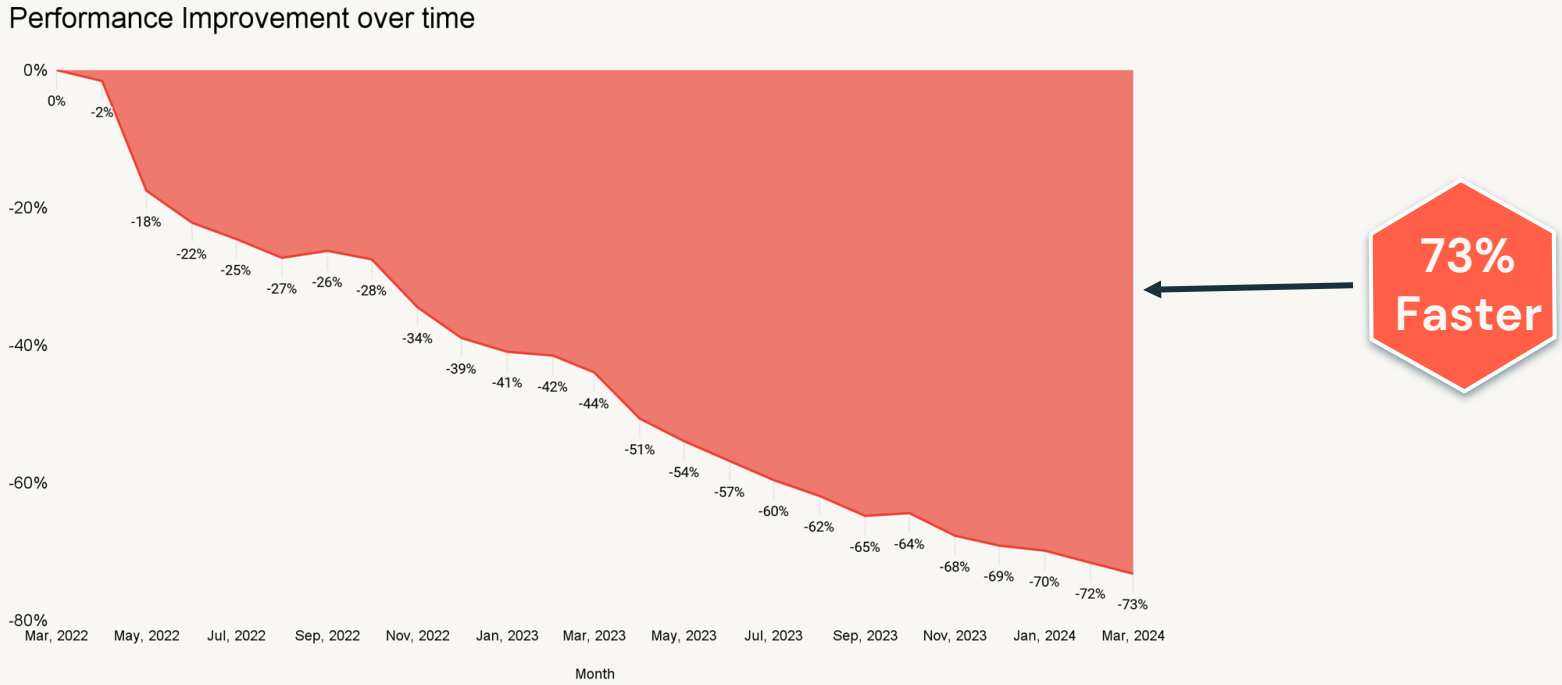
Managed compute and storage
Configure and let people work
Reduce friction for sharing & ingest

AI tuning of your workloads
Unparalleled efficiency
Minimize TCO

Moving to no knobs, no config features—serverless and self-driving Lakehouse

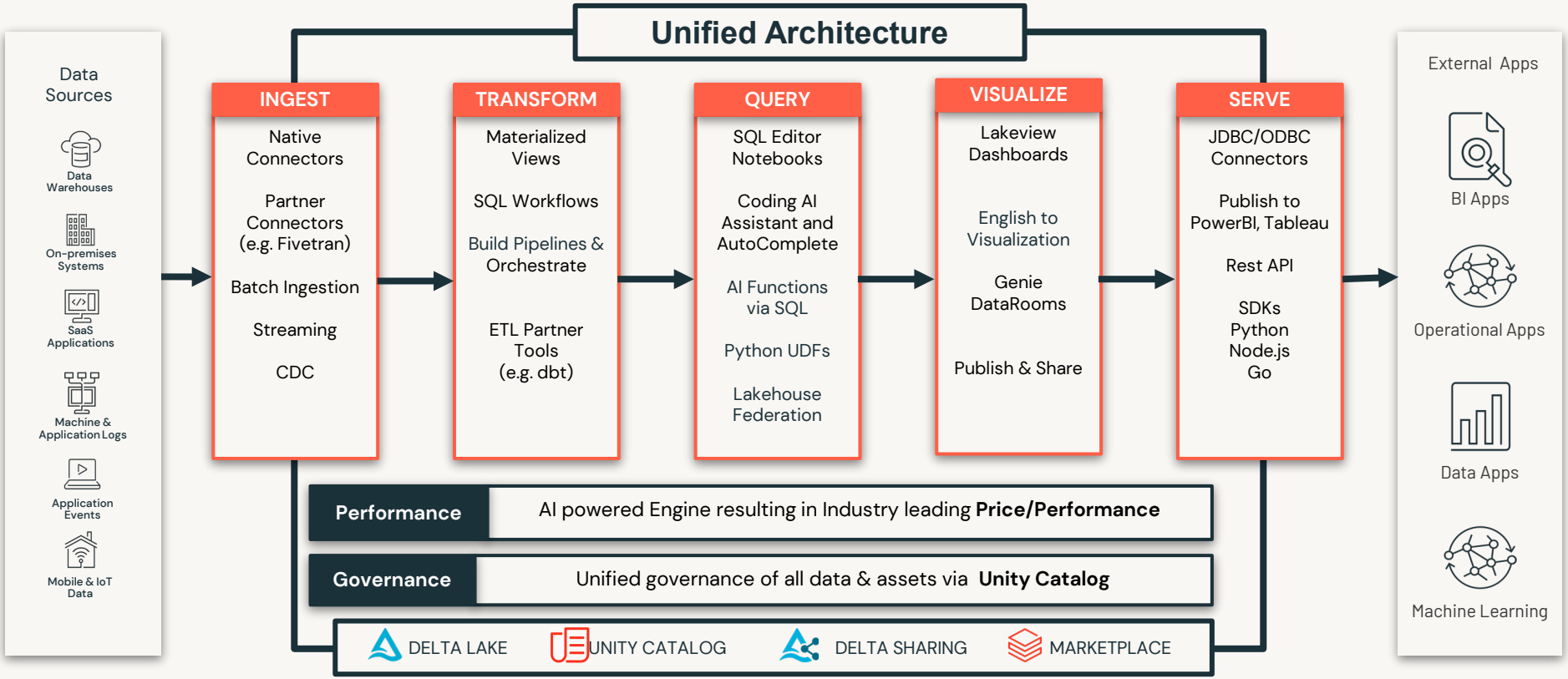
Automatic Performance improvements time

Actual customer queries improved 73% over last 2 years (4x faster)



©2024 Databricks Inc. — All rights reserved

A Complete Data Warehousing Solution



Production Data App Stack

Data Apps with Plotly Dash

- Drive every part of the Databricks Platform from a Data App all in Python

Govern + Share Insights ↓ Interact + Build Models ↓ Run Jobs ↓ SQL Querying ↓ Write to Tables / Save State

Data science & AI	ETL & Real-time analytics	Orchestration	Data Warehousing
Mosaic AI	Delta Live Tables	Workflows	Databricks SQL

Embed AI into every part of the platform to accelerate developer productivity

DatabricksIQ

Unity Catalog
Govern, Share, and Manage Data Apps in one Platform

Delta Lake UniForm
Open and sharable format with no performance sacrifices

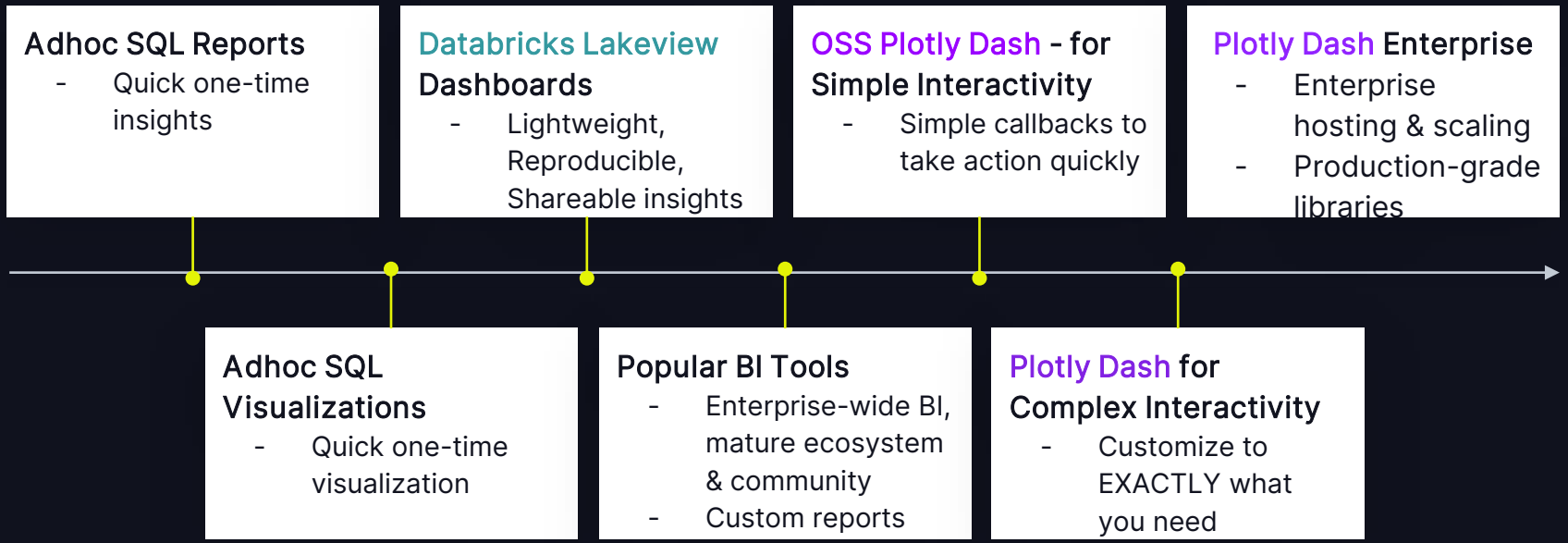
Unity Catalog

Understands all your data assets, governs and empowers your platform

- Context-aware search
- Auto describe tables and columns
- Automated lineage
- End-to-end **observability** and monitoring
- **Sharing data, functions, volumes, AI models, & more.**

Scale of Data App Severity

Databricks Supports the Entire Continuum



Building Full Stack Apps is hard - how do we make it easier?

Building full stack apps is hard, even harder when they are high analytical- much more complex state to manage

Databricks

- Databricks is the one-stop-shop to ingest, engineer, and extract insights (value) from your data.
- Fully Managed Catalog for All Data Assets
- Embedded AI for 10x productivity
- Leading Warehousing Performance

Plotly Dash

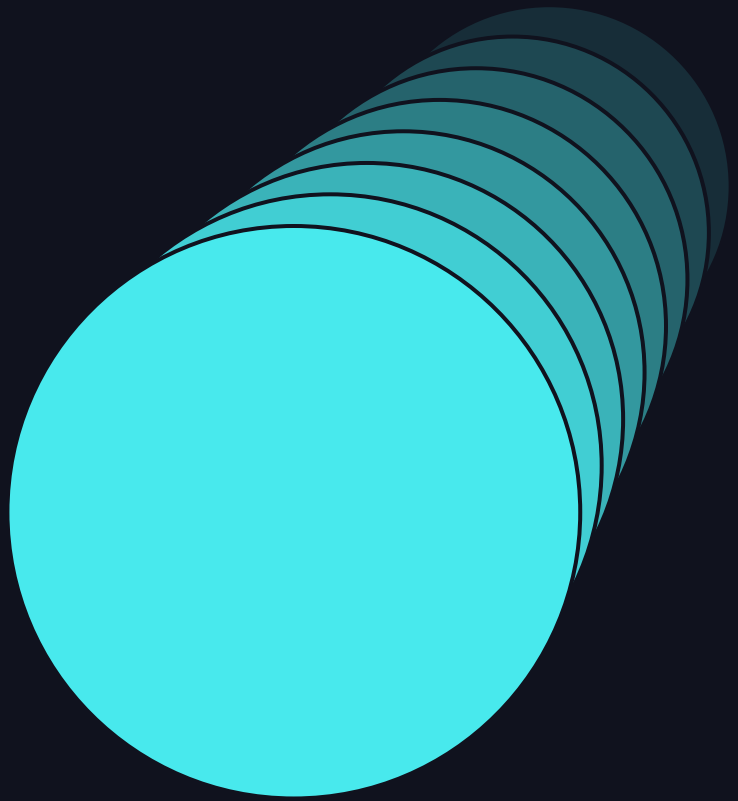
- De-facto Data Visualization leader (nearly everyone uses Plotly Charts)
- Simple full-stack app framework in Python
- Start-simple, endless customizability
- Beautiful Visualization + Simple Interactivity

Databricks + Plotly Dash for Data Apps

- *Uplevel your Data team to build more valuable & scalable data products*
- Standardize Stack on **Python + SQL**
- Start simple, customize as much as you need
- **Open Stack** - Delta/Uniform, Plotly Dash, Spark, ML FLow

DATA+AI SUMMIT BY databricks

Demo - Observability with Lakeview + Dash Apps



Real-time App Demo, by Cody Davis

SCENARIO

You are a large enterprise with…

- Over 1000+ jobs
- 50+ SQL Warehouses
- And 1000+ users on the Databricks Platform
- Adoption (and cost) is growing, and you want to get a handle on where costs are going and how to control them. Right now, you dont know where to start. You just know you need to implement tagging to allocate and observe spend by use case
- The Billing Account Console, is a great starting point, but you want to implement your own frameworks to enable your data team to solve this problem themselves

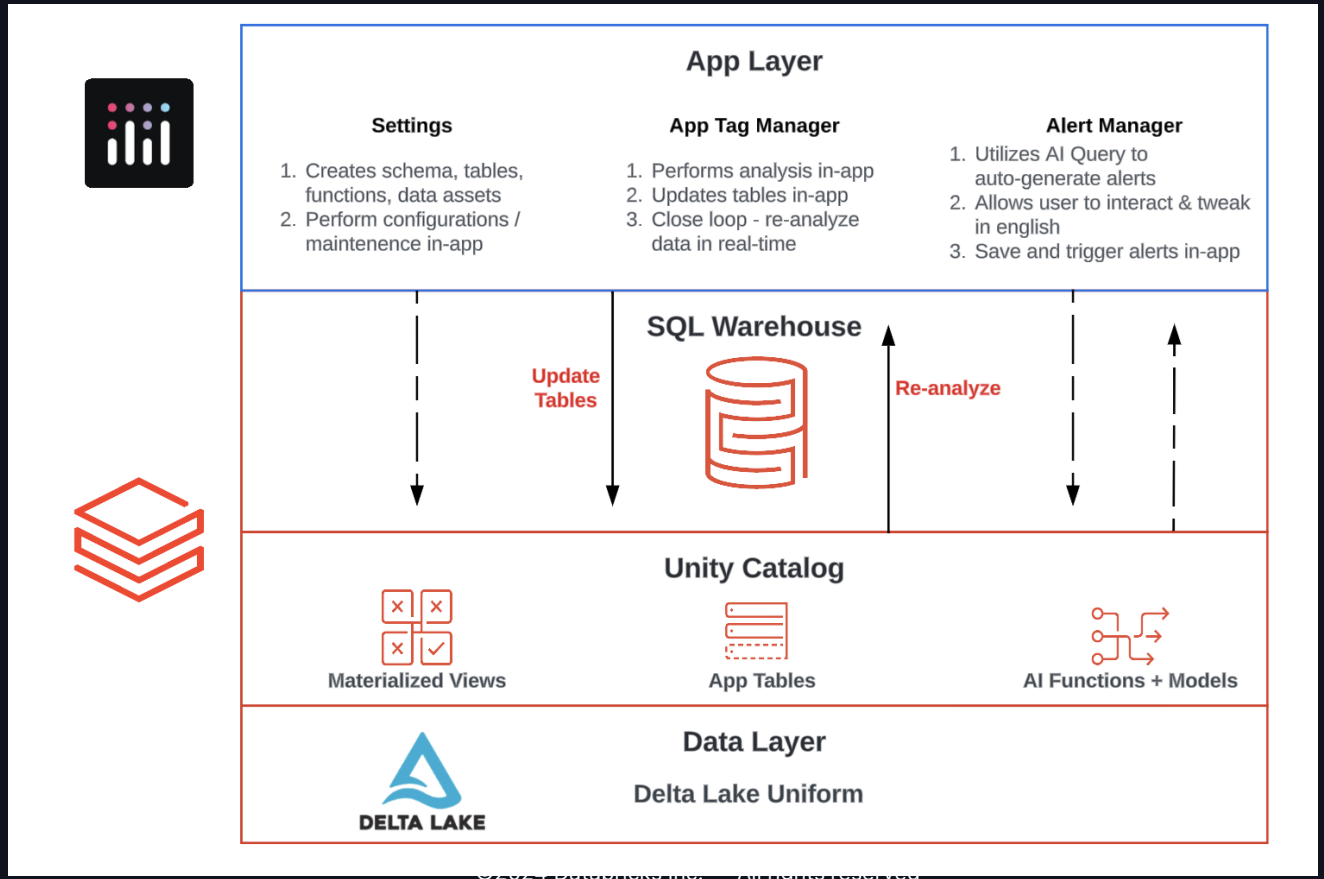
We are going to

– Review what Lakeview dashboards can do to solve this problem (and be on the lookout for more announcements for OOTB Dashboards)

– Use a Plotly Dash App on Databricks to

- Identify spend by tag (already doable in the account console, but we can dive deeper)
- Create some tagging policies
- Track adherence to those policies over time
- Identify improperly tagged assets, and fix them in the app

App Architecture



Future

What can we easily do next?

Automated Custom Alerting

AI Embedded Alerting - So non-technical users can drive

Overall Larger TAM of Users that can build and commercialize Data Apps!

Stay tuned this week on all kinds of announcements around Observability, AI, and Data Apps on Databricks!

DATA+AI SUMMIT BY databricks

Resources

- DBSQL SME Engineering Blog - For all things warehousing and apps from our SMEs
- Plotly + Databricks Page
- Lakeview Dashboard Templates
- Code to This App