

# UNLOCKING DATA VALUE: 84.51° JOURNEY WITH UNITY CATALOG

**Rashmi Kansakar, Director, Data Architect, 84.51°**  
**Kerry Holton, Sr Solution Architect, Databricks**  
**June 11<sup>th</sup> 2024**



# Introductions



Rashmi Kansakar  
Director, 84.51°



Kerry Holton  
Sr Solution Architect, Databricks

# AGENDA

## Unlocking Data Value with Unity Catalog

- 84.51° Journey
- Simplify to Unify
- Scale & Speed
- Data Product Mindset
- Unity: Future for **ALL** Data + AI

+ Gen AI

WHO IS 84.51°?

# About Us

84.51° is a retail data science, insights and media company. We help The Kroger Co., consumer packaged goods companies, agencies, publishers and affiliates create more personalized and valuable experiences for shoppers across the path to purchase.

Powered by cutting edge science, we utilize first-party retail data from over 62 million U.S. households sourced through the Kroger Plus loyalty card program to fuel a more customer-centric journey using 84.51° Insights, 84.51° Loyalty Marketing and our retail media advertising solution, Kroger Precision Marketing.

# PERSONALIZATION IS NOT NEW

The ability to do it **at scale for millions of households is new**

1850's



1950's



2000's



**Data is the new oil** – Clive Humby, 2006



# PERSONALIZATION IS NOT NEW

The ability to do it **at scale for millions of households is new**

1850's



1950's



2000's



2020's



... like oil, data is valuable, but if unrefined it cannot really be used



# 84.51° is a retail data science, insights & media company...

First-party transaction data from 62 MM U.S. households

VALUE EXCHANGE

SCALE

MONDAY-READY

OMNICHANNEL

UPC-LEVEL

LONGITUDINAL

...Influencing customer engagement along **the fast, nonlinear and complex** path the purchase

## 84.51° Loyalty Marketing

**Builds long-term engagement and loyalty** through personalized offers and messages.

## 84.51° Insights

**Understands customer needs, interests and desires** through actionable insights along the path to purchase.

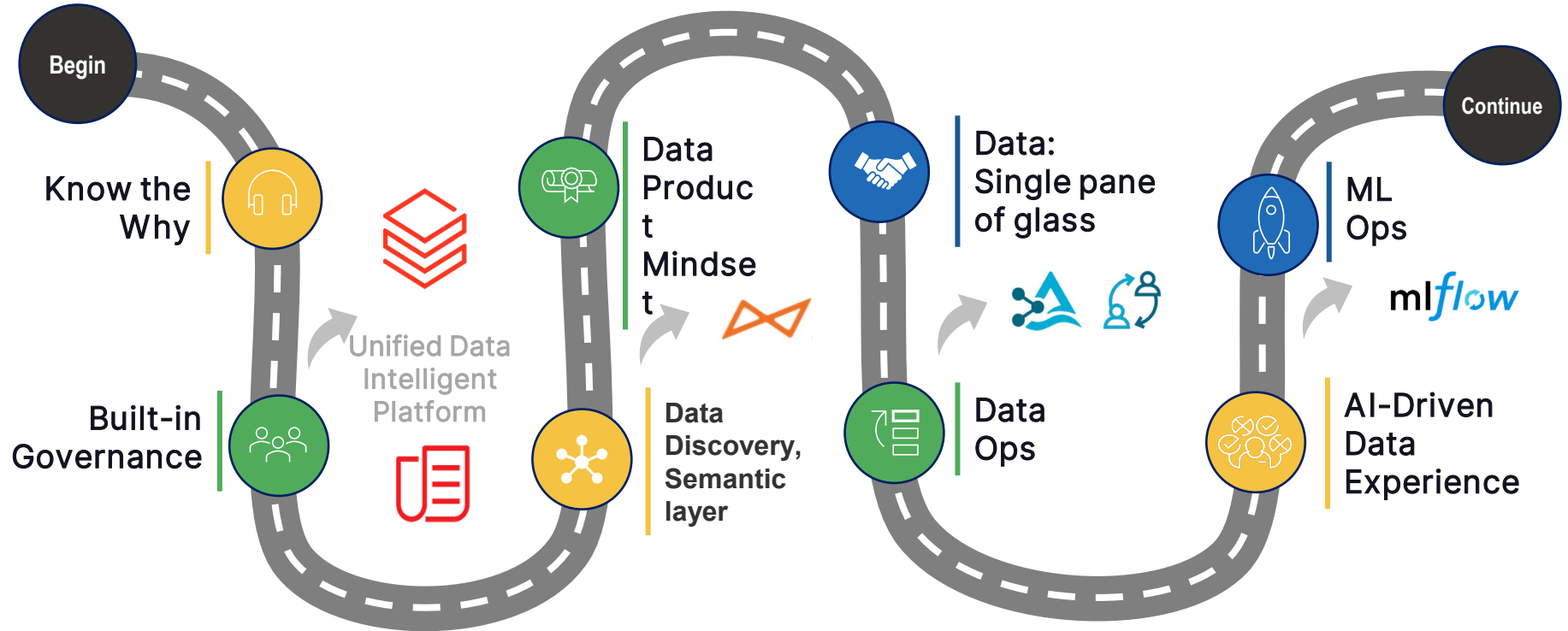


**Drives your brand's advertising effectiveness** by closing the loop between media exposure and store sales.



# THE UNITY CATALOG JOURNEY

## Unified Platform for DATA & AI



# SIMPLIFY TO UNIFY

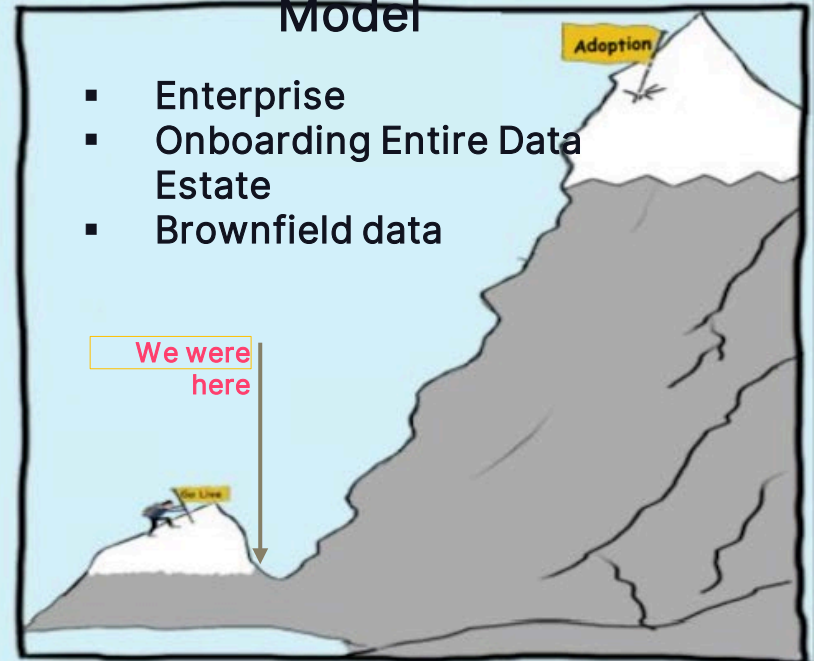
# SIMPLIFIED SECURITY DESIGN

## Bridging data trust gap

### Use Case Driven Model Design

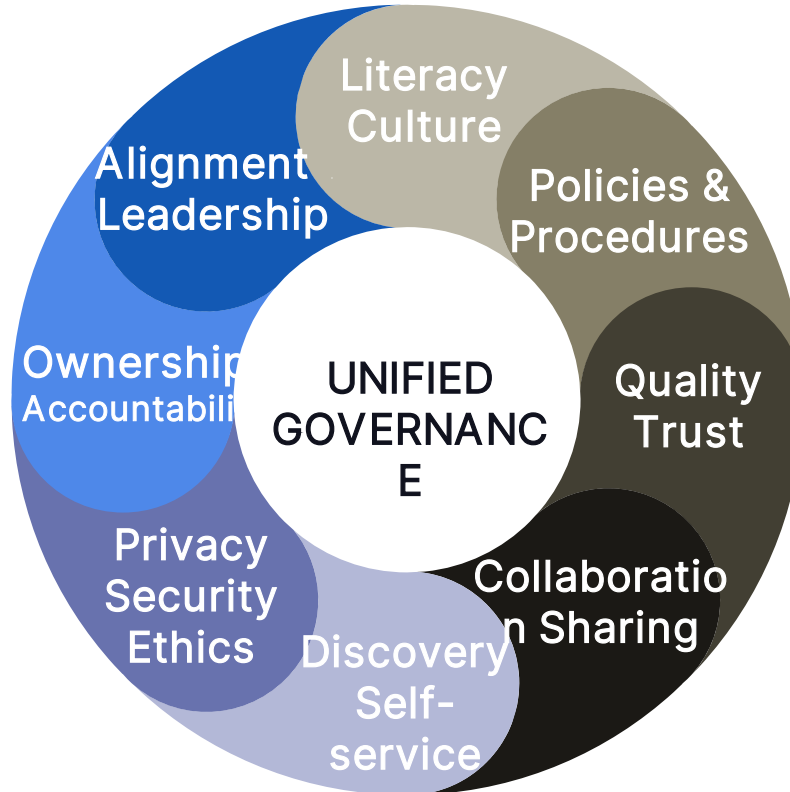
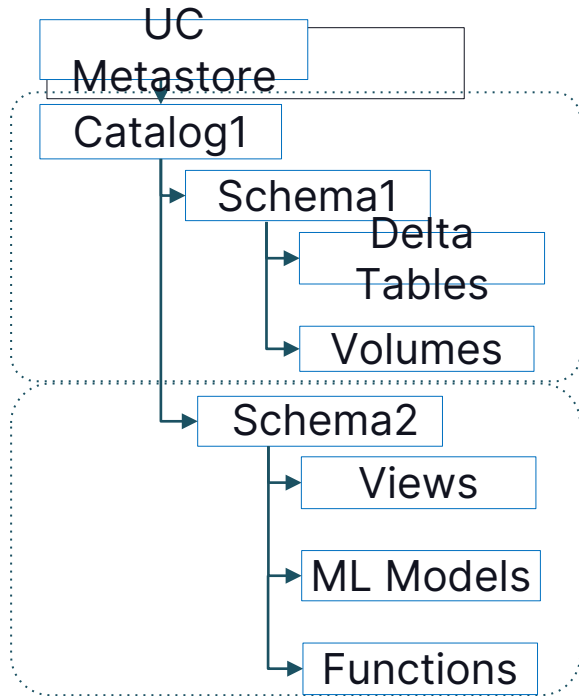


### Data Domain Driven Model



# BUILT-IN DATA GOVERNANCE

## Key to Data Enablement



## Data Innovation...

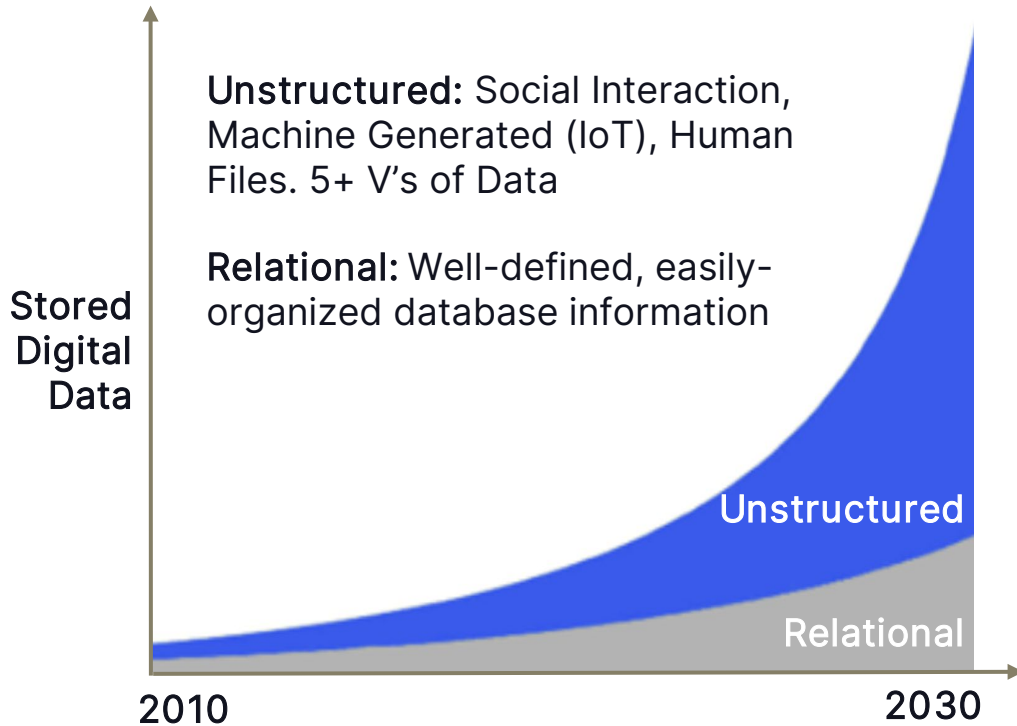
- ✓ Within our teams
- ✓ Within our organization
- ✓ Throughout our organization
- ✓ Across Suppliers
- ✓ Customers
- ✓ Data Cleanroom



# SCALE & SPEED

# HOW WE GOT HERE?

## Explosion of Big Data



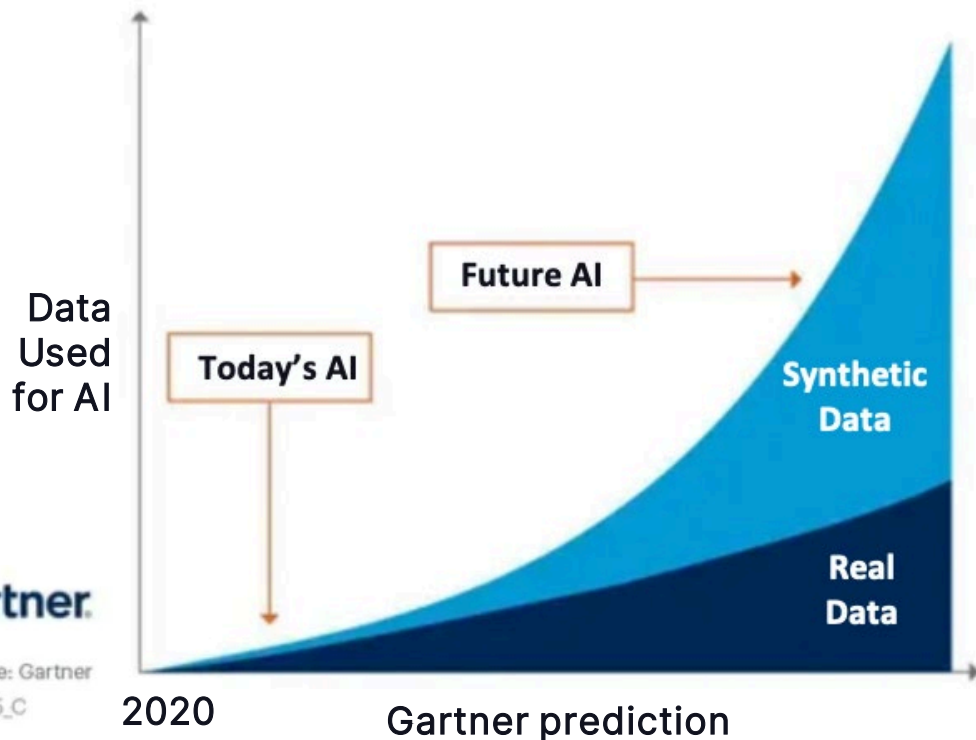
- ✓ Faster, better decision-making
- ✓ Boosted productivity
- ✓ Quicker issue resolution
- ✓ A better understanding of the market
- ✓ Regulatory compliance

# WHERE WE ARE GOING?

## Gartner Prediction



Good models  
**CANNOT OVERCOME**  
bad data



### Synthetic Data:

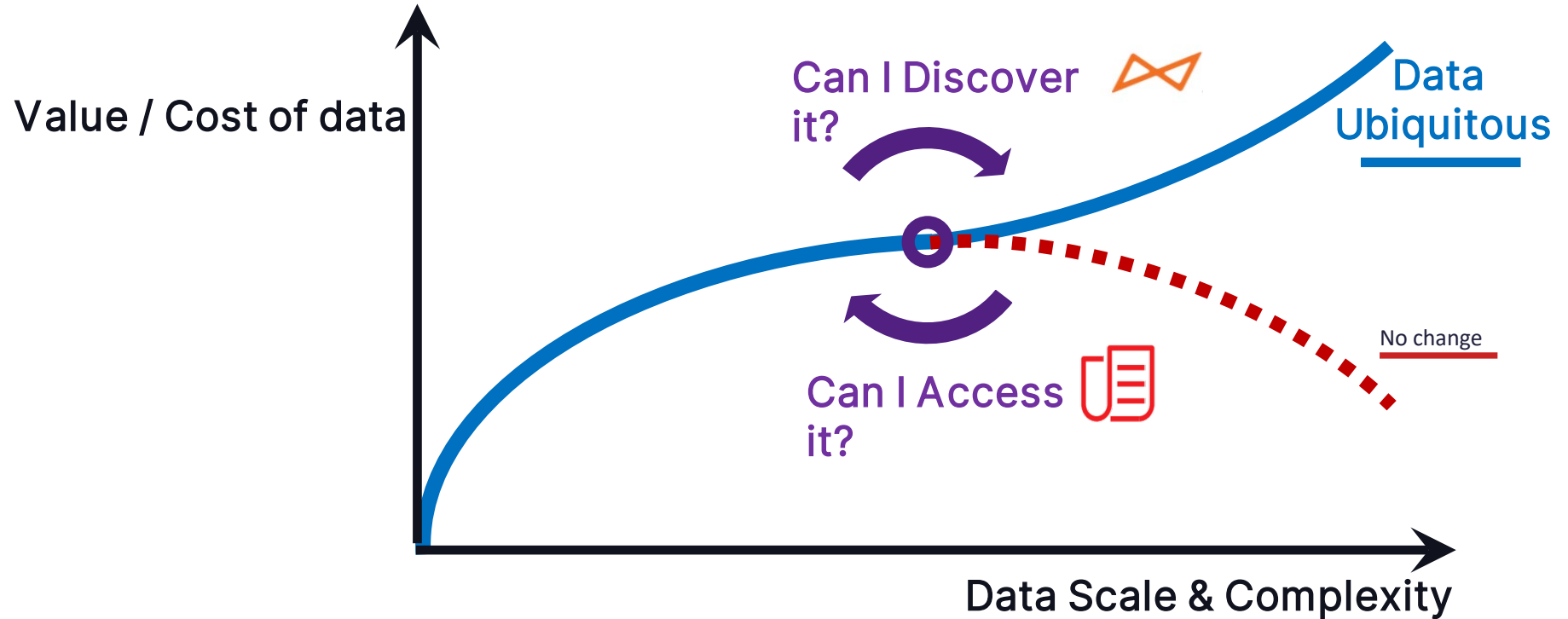
- Artificially Generated Data
- Generated From Simple Rules, Statistical Modelling, Simulation and Other techniques

### Real Data:

- Obtained From Direct Measurements
- Constrained by Cost, Logistics, Privacy Reasons

# VALUE FROM CONNECTED DATA

Bridging data for seamless collaboration, innovation driving connected future





# DATA PRODUCT MINDSET

# TOGETHER WE ARE STRONGER

Data Product Mindset cultivates trusted data to elevate decisions

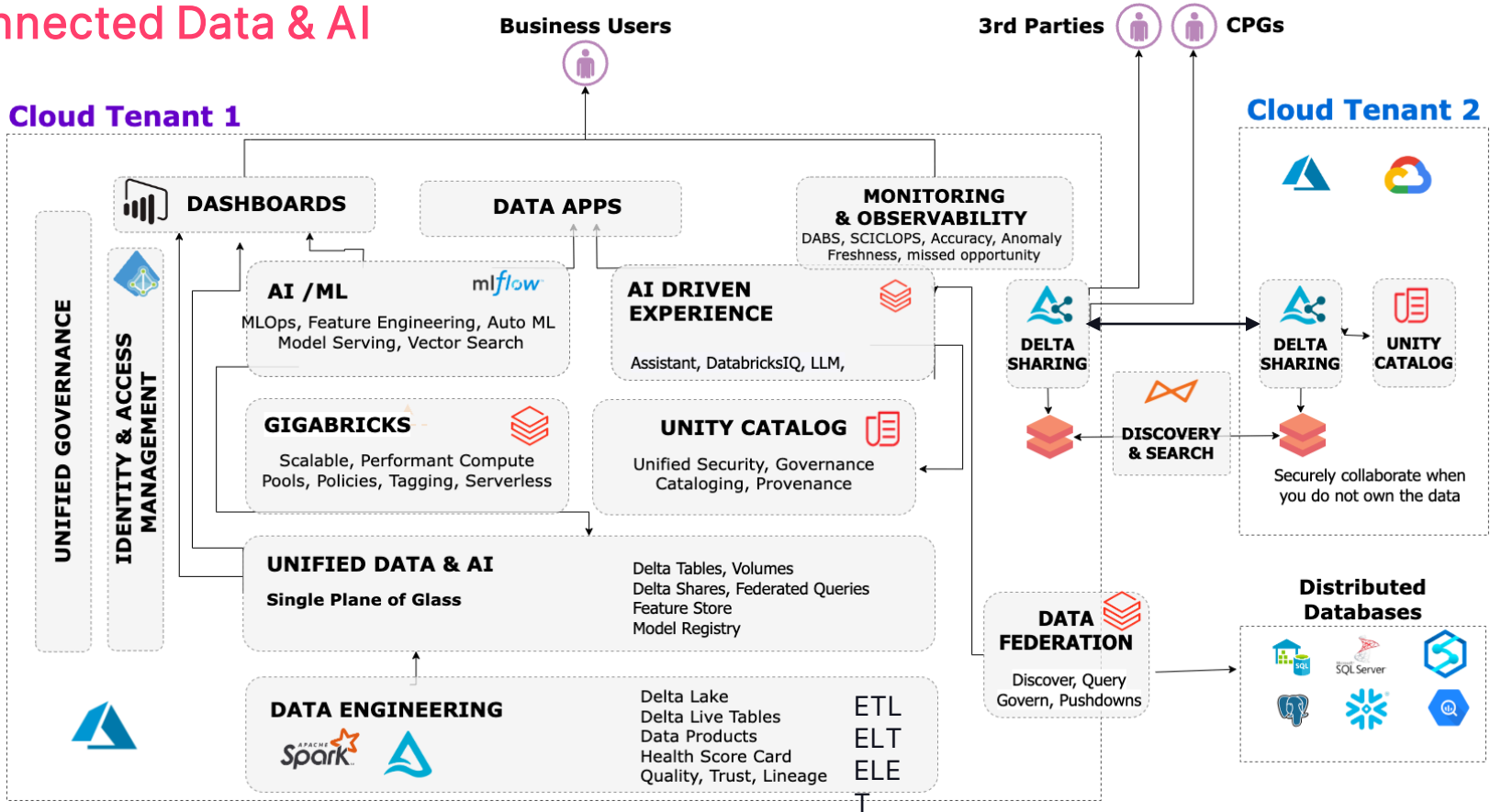


Data Product Mindset  
Pay It Forward Mindset

A society grows great when  
a person plant trees in  
whose shade they shall  
never sit!

# DATA ARCHITECTURE

## Connected Data & AI



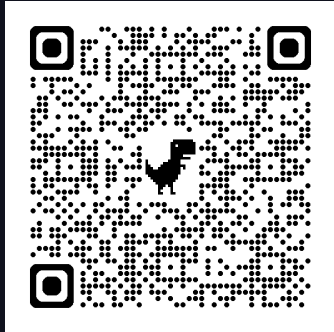
# Our Other Talks!

Check out more from 84.51°

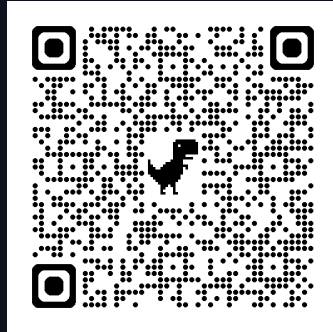
*Learn more about Unity  
Journey!*



SciCLOps: Databricks  
Quick Start for Machine  
Learning, Powered by  
DABs



Databricks Asset Bundles:  
A Unifying Tool For  
Deployment On  
Databricks



Seamless Flow: Evolving  
From Batch to Streaming  
Data Flows Using DLT



# Unity: Future for **ALL** Data + AI

# Data and AI governance drives business value

“Organizations are finally realizing the value of **data as an asset** that needs to be protected, managed and maintained to **increase asset value**”



—  
IDC

“Organizations seeing the **highest returns** from AI have a framework for **governance** to cover every step of the model development process”

—  
The State of AI in 2022, McKinsey & Co

“AI is now an enterprise essential, and as such, **AI governance** will join cybersecurity and compliance as a **board-level topic**”

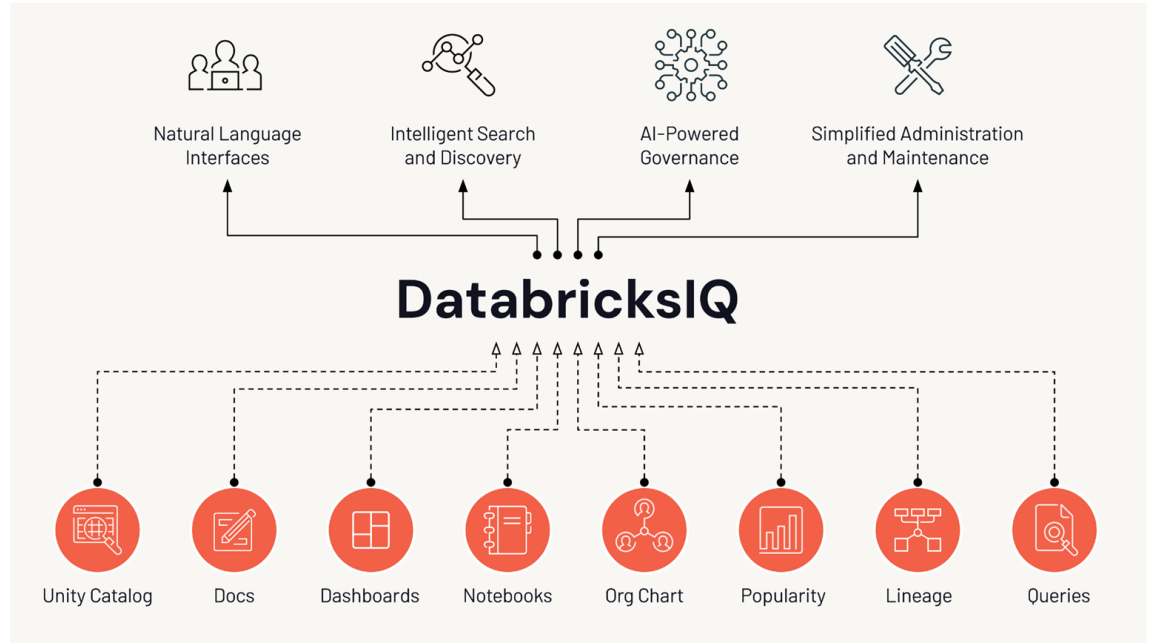
—  
Forrester, 2023 AI Predictions report

Unity Catalog Unlocks the Full Potential of your Lakehouse	w/ UC 	w/ Hive Metastore 
<b>Governance across workspaces</b> (Access controls, metastore, Identity management)	✔	✘ (Governance at workspace level)
<b>Automated table and column level lineage in real-time</b>	✔	✘
<b>Row Filters and Column Masking</b> (use standard SQL functions for fine-grained access controls)	✔	✘
<b>Volumes</b> (Govern non tabular data/arbitrary files - image, video, PDF etc.)	✔	✘
<b>Lakehouse Federation</b> (Discover, govern and query external databases, data warehouses, Hive Metastore, AWS Glue)	✔	✘
<b>Lakehouse Monitoring</b> (Monitor quality of data and ML models with auto generated alerts and dashboards)	✔	✘
<b>Delta Sharing</b> (Share files, tables, ML models, notebooks across clouds, regions and data platforms)	✔	✘
<b>Databricks Marketplace</b>	✔	✘
<b>Databricks Cleanroom</b> (Collaboration on data in a privacy-preserving manner)	✔	✘
<b>Databricks Data Room (aka Genie)</b>	✔	✘
<b>Databricks Assistant / LakehouseIQ</b> (Knowledge engine on the lakehouse to get insights in natural language)	✔	✘ (Degraded experience)
<b>Governance for AI</b> (manage and govern ML Models, Feature store)	✔	✘
<b>System Tables</b> (end to end observability for Billing, Auditing, Lineage, Marketplace analytics)	✔	✘
<b>Lakehouse Apps</b>	✔	✘
<b>Databricks Connect V2</b>	✔	✘
<b>Materialized Views</b>	✔	✘
<b>AI Powered Predictive optimization</b> (optimizes your query plans and data layout for peak performance, intelligently balancing read vs. write performance)	✔	✘
<b>Streaming Tables</b>	✔	✘
<b>HMS Interface</b> (Query data registered in UC using other data platforms such as Amazon Athena, Presto, Trino, EMR)	✔	✔
<b>Vector Search for indexing</b>	✔	✘
<b>Serverless workflows</b>	✔	✘

# DatabricksIQ Data Intelligence Engine

AI that uniquely understands your data and usage patterns

- *Data Intelligence Engine* deeply understands the semantics of your data
- Improve productivity with better search and assistant answers
- *Democratize data access* using natural language
- *Maintain governance and control* with Unity Catalog





# Databricks Assistant

Boost productivity with your context-aware AI assistant

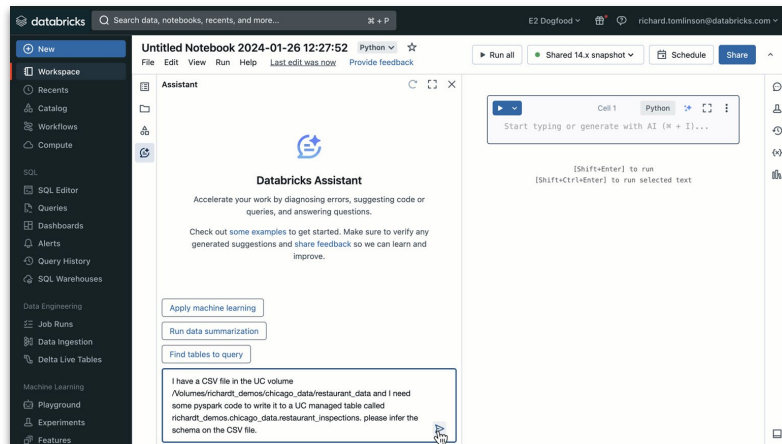
- Uses **Unity Catalog** metadata for results relevant to your data assets
- AI assistant on **ALL** Databricks product pages
- Generate, explain, and fix code with natural language
- Diagnose job failures and get help specific to the UI you are using

*“A groundbreaking AI-powered tool that transforms coding productivity!”*

—Saptagiri Kintali, VP/SDM Data @ Morgan Stanley

*“I was able to code 200+ lines of robust code this week in a language I’ve never coded before”*—Josue A. Bogran, Solutions Architect Manager @ Kythera Labs

*“This cutting-edge AI companion has revolutionized my data analysis journey, simplifying complex tasks and accelerating productivity.”*—Byron Exaporriton, Analytics Engineer @ ABN AMRO



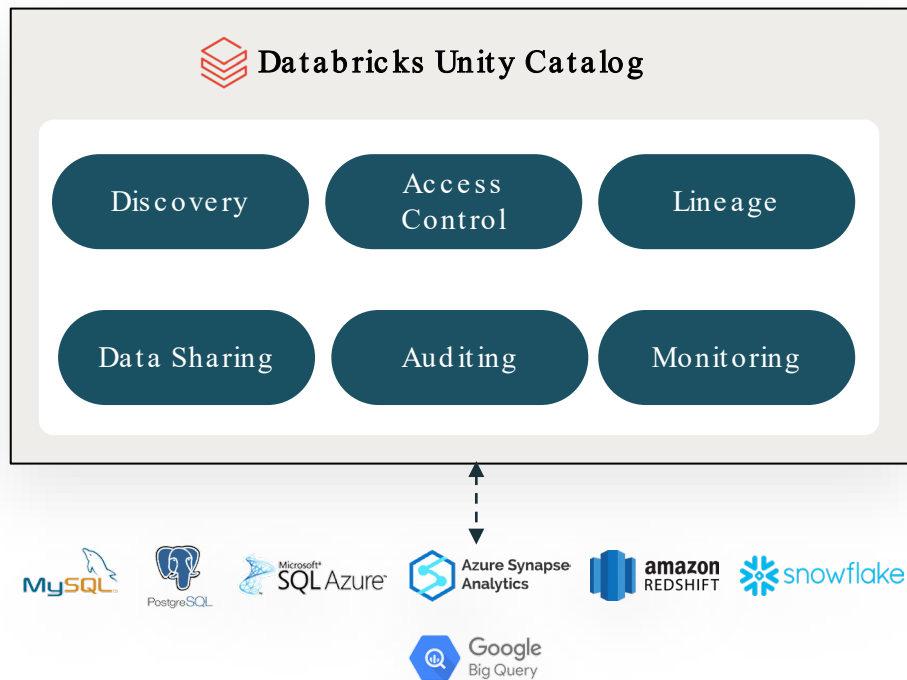
Over 100K  
users assisted  
in Jan 2024

# Lakehouse Federation

Unify your data estate with the Lakehouse

Discover, query, and govern  
**all your data** - no matter  
where it lives

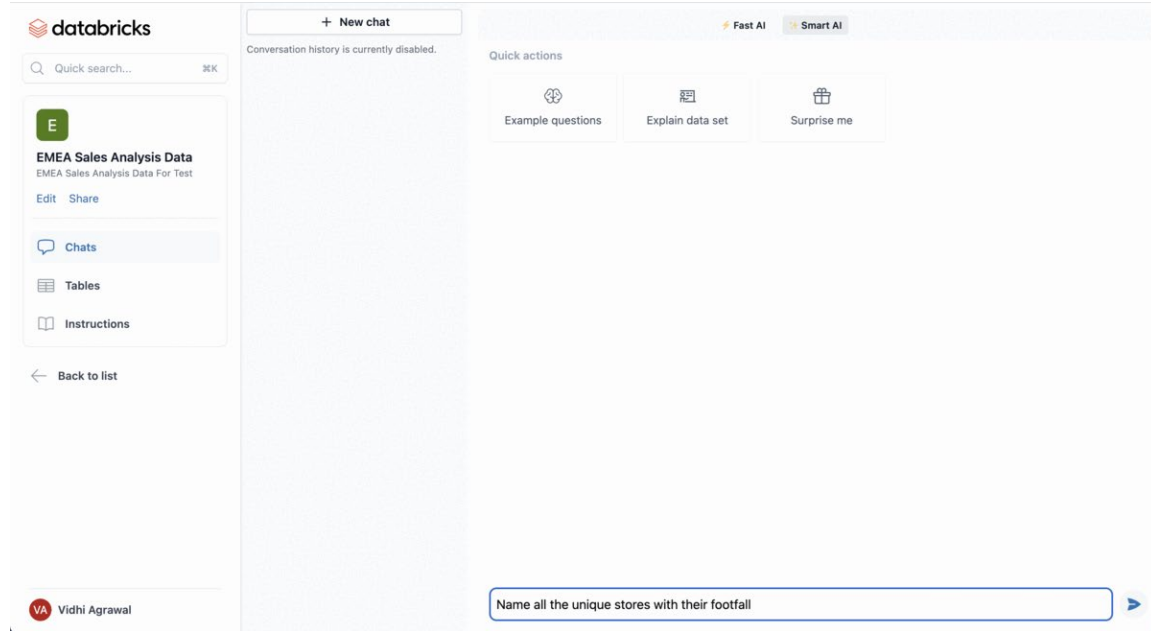
- **Unified governance** across all data sources with Unity Catalog
- Connect and query **popular databases and warehouses**
- **Single point of access** for all your data with performance optimizations



# Databricks Data Rooms (Genie)

## Empowering Business Users with AI-Powered Insights

- Technical analysts can use data residing in SQL warehouses in Databricks to set up Data Rooms for **business users**
- Business users can **ask questions in natural language**
- AI generated query **suggestions** make it easier to get started



# Databricks Clean Rooms

Privacy-safe collaboration for data, analytics and AI

Secure environments to run multi-party computations on joint data in the lakehouse

- Run any computation in Python, SQL, R, Java, etc.
- No data replication of existing lakehouse tables with Delta Sharing
- Scale to multiple collaborators on any data size or cloud

The screenshot displays the Databricks Clean Rooms interface. A 'Create clean room' dialog box is open, showing fields for Name, Collaborators (Sharing Identifier and Organization name), Cloud, and Region. Below the dialog, the main interface shows a 'Clean Rooms' section with a date '01-16-2024', owner 'maraki.ketema@databricks.com', and region 'us-west-2'. There are tabs for 'Assets' and 'Collaborators'. A 'Filter assets' search bar is present, along with a 'Managed by: All collaborators' dropdown. Below this, there are sections for 'Notebooks (1)' and 'Data assets (2)'. The 'Notebooks (1)' section contains a table with columns: Name, Managed by, Added on, and Approval. The 'Data assets (2)' section contains a table with columns: Status, Name, Managed by, and Added on.

Name	Managed by	Added on	Approval
Audience Overlap	wayne-enterprise	2024-01-16 10:13:37	0/2

Status	Name	Managed by	Added on
🟢	acme_catalog.default.acme_customers	acme_co	2024-01-16 10:23:15
🟢	wayne_catalog.default.identity_graph	wayne-enterprise	2024-01-16 10:21:29

# Mosaic AI

Build and scale Generative AI applications on Lakehouse

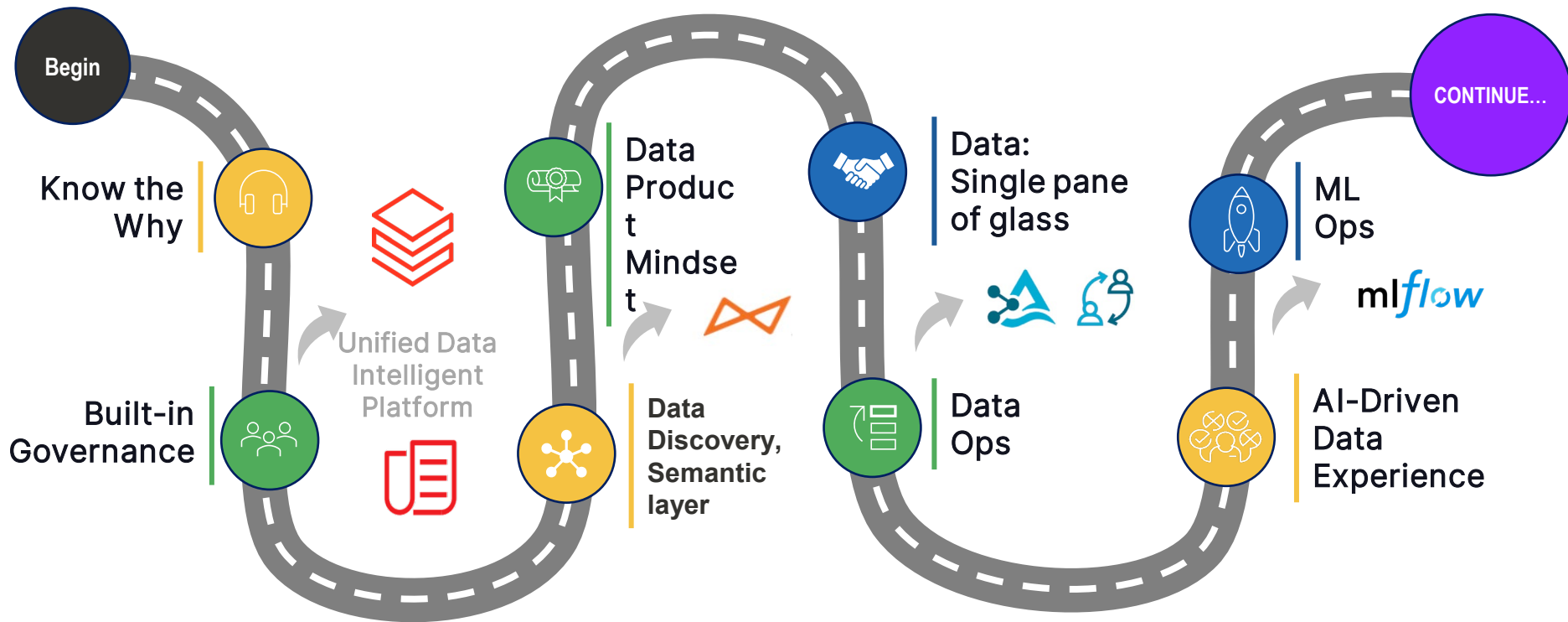


**Governance — Unity Catalog**

**Data Platform — Delta Lake**



# The Unity Catalog Journey at 84.51° Continues...



# DATA+AI SUMMIT

