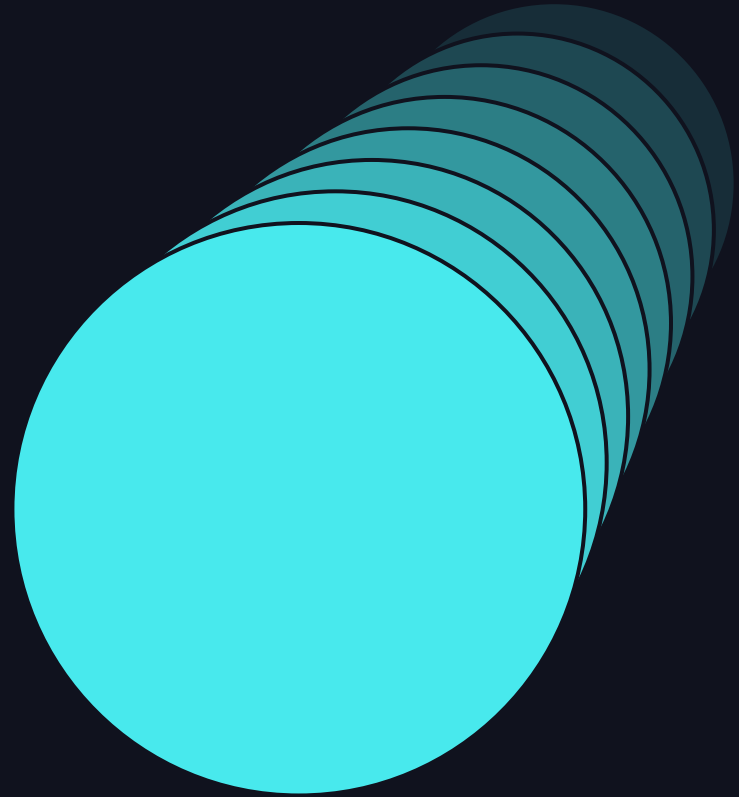


# AI GOVERNANCE WITH COLLIBRA

Extend Data Governance into your Use of AI  
with Collibra

---

Collibra  
June 2024



# AGENDA

## Extend Data Governance into your Use of AI with Collibra

- Collibra and Databricks Partnership
- How Collibra Adds Value On Databricks
- Product integration roadmap
- About Databricks Data and AI governance
- Introducing Collibra Data and AI governance
- Building our RAG with Unity Catalog, MLflow and Collibra
- Establish Trust through evidence of well-governed AI
- Q&A

# COLLIBRA & DATABRICKS PARTNERSHIP

Collibra extends the value of your Databricks investment



**225+ shared  
customers**



**2024 Data  
Governance  
Partner of the  
Year**



**2022 Data  
Governance  
Partner of the  
Year**

# HOW COLLIBRA ADDS VALUE ON DATABRICKS

Users



Data Scientist



Data Engineer



Data Analyst



Business Analyst



Privacy Manager



Data Steward

# Collibra Data Intelligence Cloud

## Enterprise

Data Catalog

Governance

Lineage

Quality & Observability

Data Marketplace

Privacy and Security



### Data Intelligence Platform

Mosaic AI

Delta Live Tables

Workflows

Databricks SQL

DatabricksIQ

Unity Catalog

Delta Lake UniForm

Metadata

Lineage

Access policies

DQ Pushdown



Collibra

Extending Databricks' value to business users & across all enterprise data and AI assets

Policies, glossaries, reference data

Enterprise data marketplace

Cross-system lineage

Privacy and security context

Find and govern data assets / products

Federate data governance

Data collaboration and helpdesk

Data curation and automation workflow

ML-based data quality

## Hyperscalers



Google Cloud

## Data Stores and Applications



## AI and BI Tools



+ a b l e a u

## Integrations

100+ pre-built integrations



# AI GOVERNANCE VALUE ADD

Collibra makes your AI and GenAI projects successful



## Collibra AI Governance

- Catalog AI use cases and validate data
- Collaborate across your organization
- Mitigate data risk and protect data



## Databricks Mosaic AI

### Gen AI

- Custom models
- Model serving
- RAG
- DBRX

### End-to-end AI

- MLOps/MLflow
- AutoML
- Monitoring
- Governance



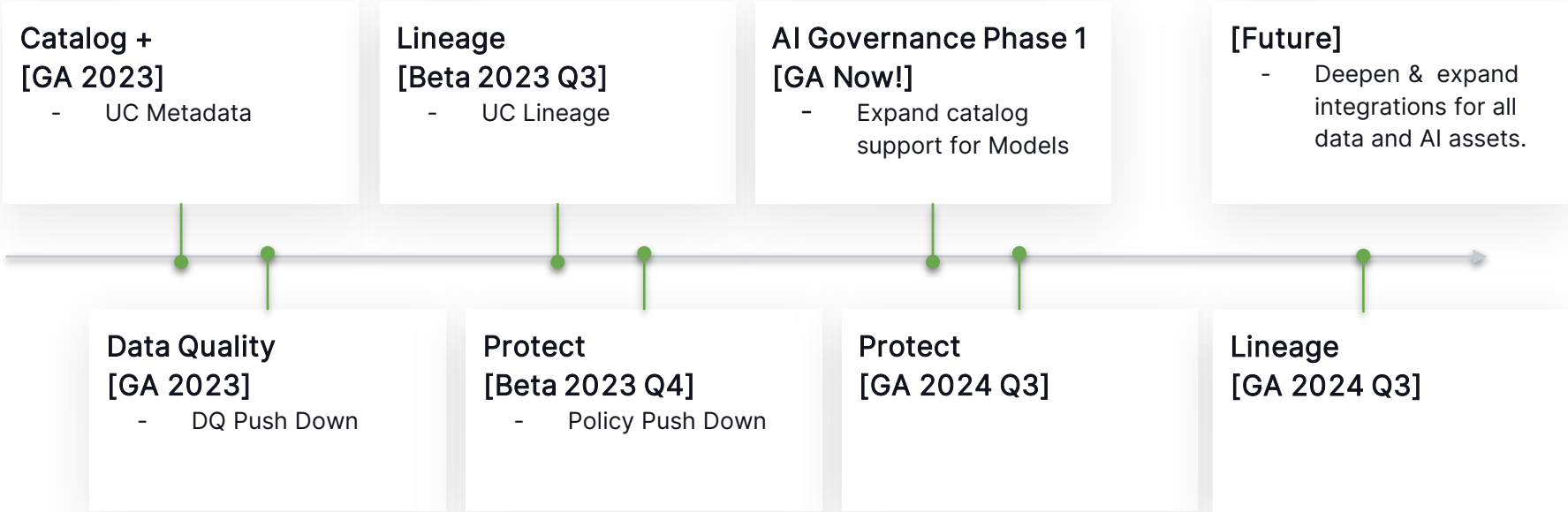
Move AI projects  
to production  
faster  
&  
Democratize AI  
products across  
your enterprise



# PRODUCT INTEGRATION ROADMAP

# COLLIBRA PRODUCT INTEGRATION ROADMAP

## Expanding integration to AI assets

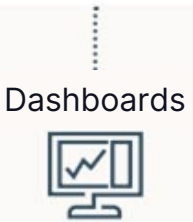
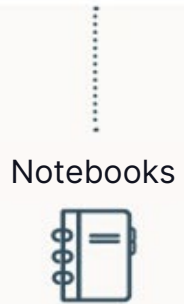
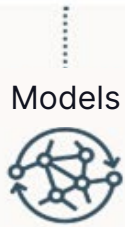
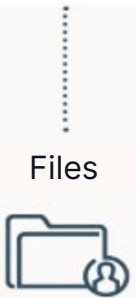
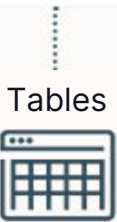




# ENSURE RELIABLE, TRACEABLE, AND COMPLIANT AI

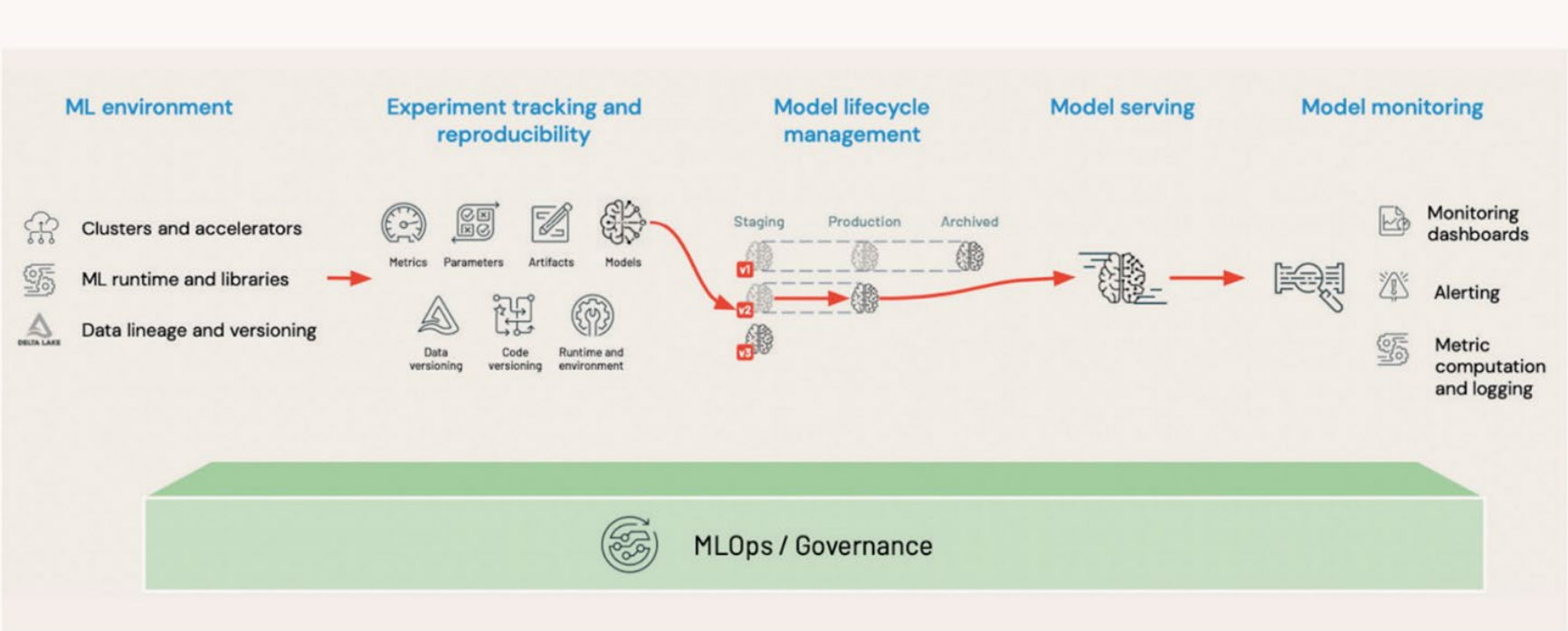
# DATABRICKS UNITY CATALOG

A unified governance solution for data+AI



# DATABRICKS MLOPS GOVERNANCE

## Managing the lifecycle of your ML models



# MODEL METRICS – ARE THEY ENOUGH?

## ML model metrics – should we trust them?

### AI risks

From sources across the web

Transparency	▼	Privacy concerns	▼	Security	▼
Discrimination	▼	Dependence on AI	▼	Economic inequality	▼
Ethics	▼	Loss of human connection	▼	Misinformation and manipulation	▼
AI arms race	▼	Autonomous weapons	▼	Biased outputs	▼
Existential risks	▼	Job displacement	▼	Law	▼
Unintended consequences	▼	Concentration of power	▼	Social manipulation	▼
Invasion of privacy and social gradi...	▼	Liability for actions	▼	Social surveillance with AI technolo...	▼



# MODEL METRICS – ARE THEY ENOUGH?

## ML model metrics – should we trust them?

In the context of AI, "**hallucination**" refers to instances where a model generates outputs or statements that are **false, fabricated, or nonsensical, despite appearing plausible...**

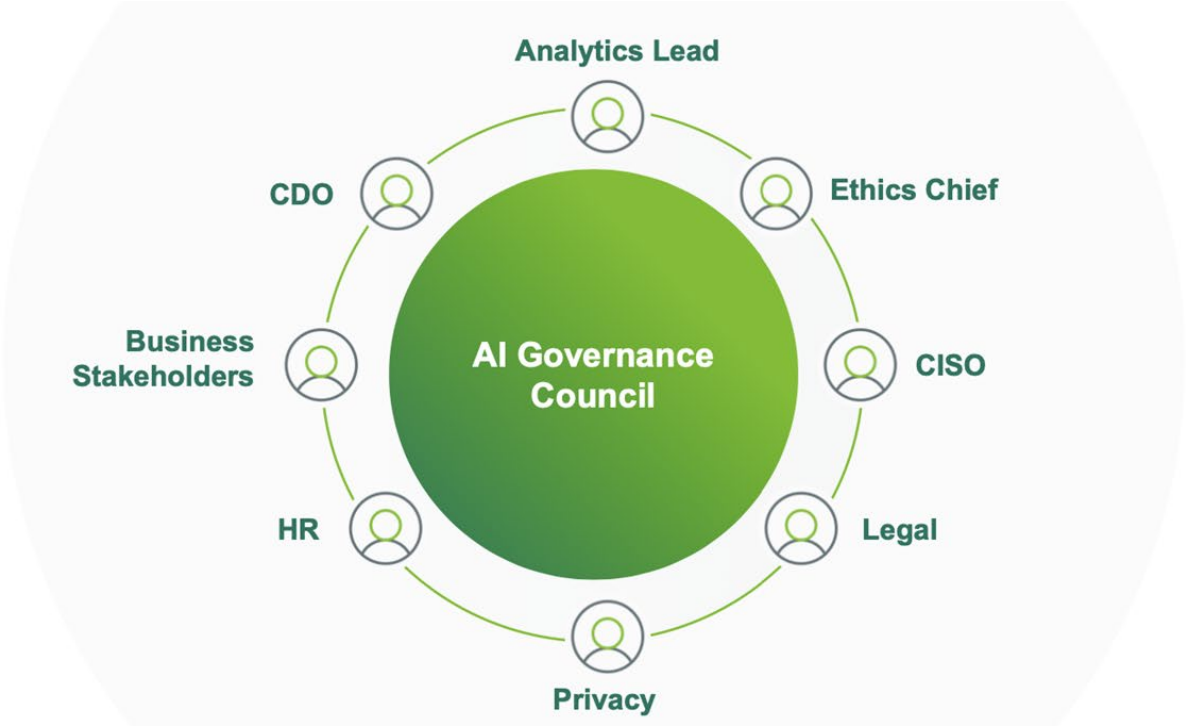
"**Bullshit**" refers to statements or information presented with **little regard for truth, accuracy, or relevance**. The term is often used to describe speech or writing that is intended to impress, persuade, or deceive without concern for **genuine substance or truthfulness...**

<https://undark.org/2023/04/06/chatgpt-isnt-hallucinating-its-bullshitting/>



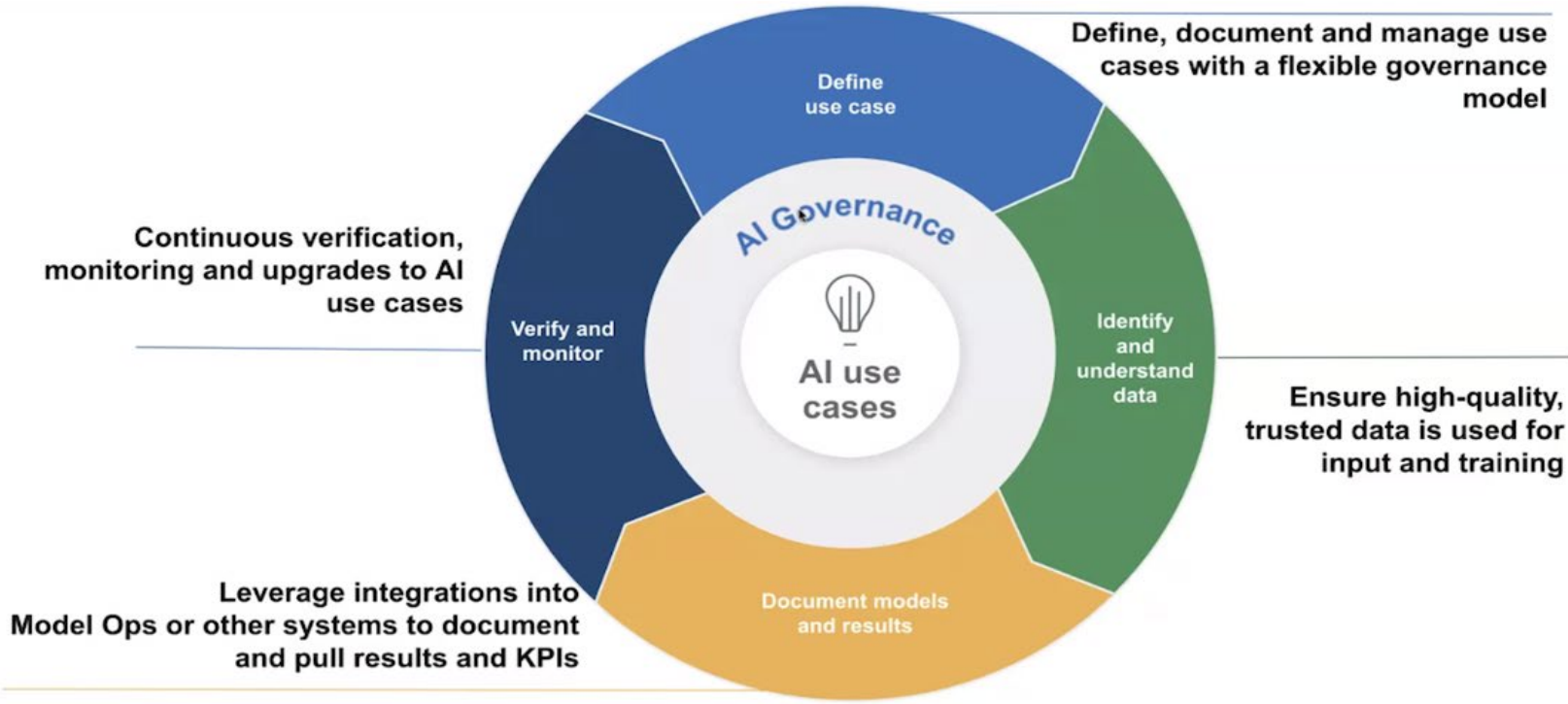
# COLLABORATION - TO GET IT RIGHT

Collaboration is crucial to delivering results



# COLLIBRA DATA AND AI GOVERNANCE

Minimize risk, ensure responsible use of AI



# DELIVER TRUSTED AI



- Prioritize your AI based on revenue potential, impact, and risk.
- Select the appropriate set of controls to implement according to your risk assessment.
- Ensure humans maintain control over AI, while providing seamless oversight to stakeholders.
- Help ensure reliable, traceable, and compliant AI



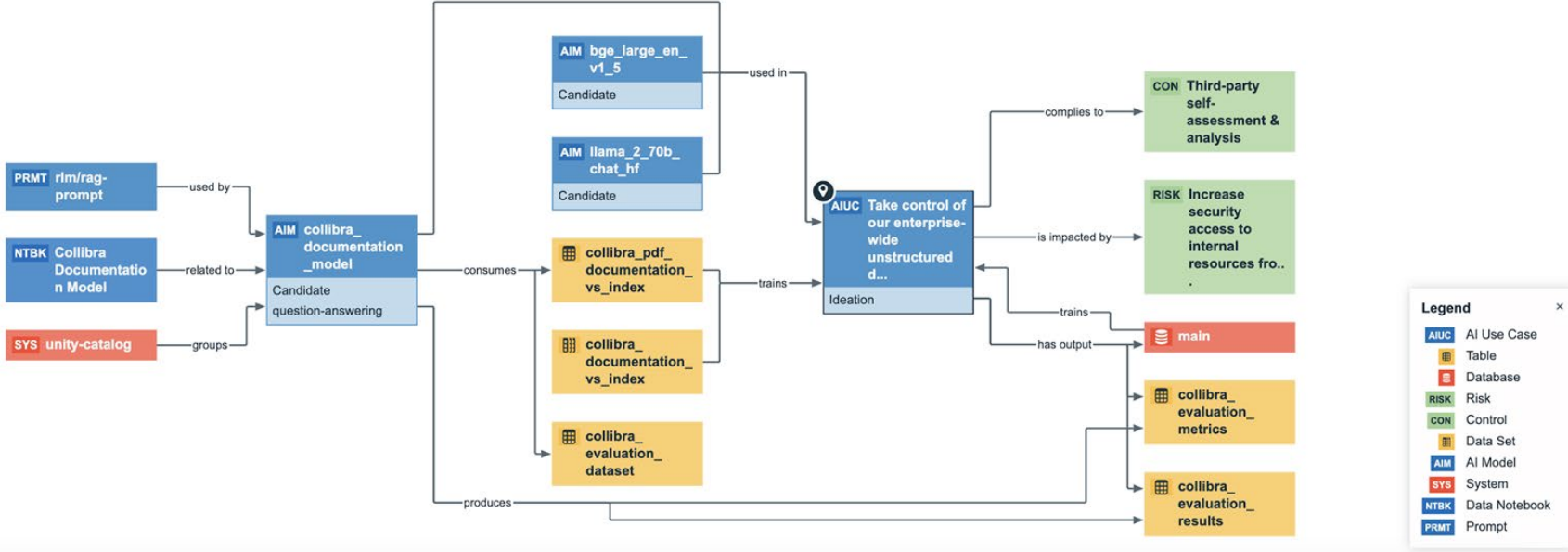
- Centralize model registry to manage the entire lifecycle of your models.
- Simplify development, testing and monitoring of your Data and AI products
- Track experiments and compare each individual model run parameters and resulting metrics





# BUILDING OUR RAG WITH UC + MLFLOW

## Take control of our unstructured resources



# BUILDING OUR RAG WITH UNITY CATALOG, MLFLOW AND COLLIBRA

# ESTABLISH TRUST THROUGH EVIDENCE OF WELL-GOVERNED AI

# DELIVER TRUSTED AI

With streamlined workflows guiding stakeholders through the design, review and approval process, organizations can **confidently deploy reliable, traceable and compliant AI products** that meet their business requirements and accelerate positive business outcomes.

# DATA+AI SUMMIT

DRIVING TRACEABLE AND TRUSTED AI  
THROUGH COLLABORATION

SCAN QR CODE TO LEARN MORE

