

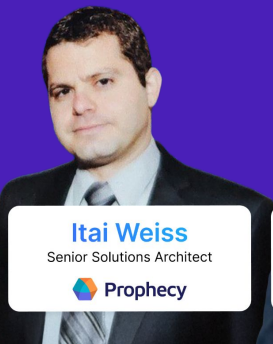


Prophecy | DATAVERSITY

## Guardrails to go faster:

# How pipeline governance unlocks data access for analysts

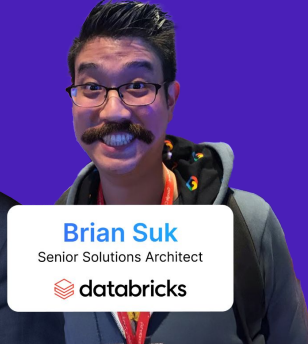
May 29, 2025



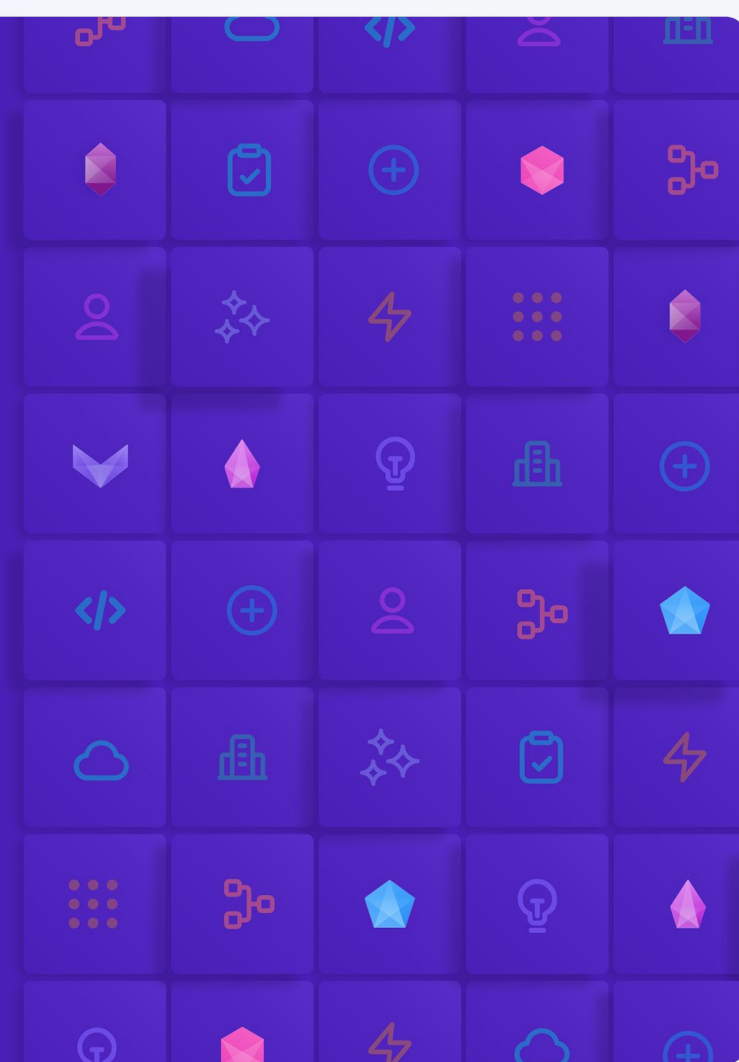
Senior Solutions Architect



Director of Product Marketing



Senior Solutions Architect



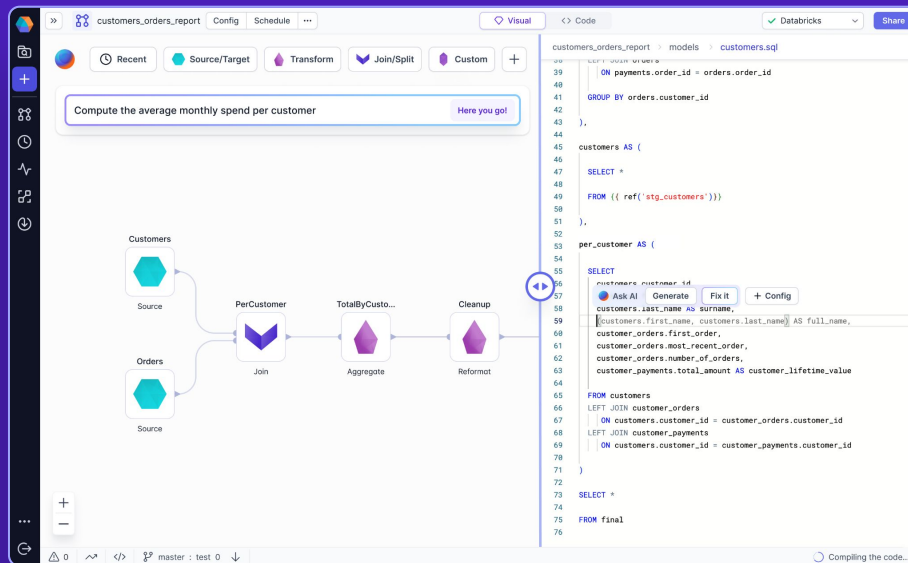


# Prophecy Data Integration Platform

## Self-service

Data analysts prepare data  
on their own

Data  
Analyst



## Governance

Working within guardrails  
set by the central data  
platform team

Data  
Engineer





# SURVEY SAYS

46% of data teams using GenAI are seeing **31-50%** productivity gains

Teams are planning to make more **data** and **AI hires**

## The Impact of GenAI on Data Teams

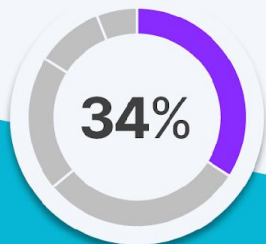
A survey of data leaders on overcoming challenges & unlocking opportunities in data and workflow transformation

February 2025

Top AI barriers:  
**data governance**  
and **faster access**  
to data

## The top data and analytics process that will see the most impact from applying genAI

*Among those currently using or planning to adopt genAI, top ranked responses shown, N=499*



Data transformation and preparation



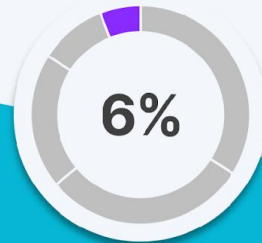
Data quality



Data ingestion



Data analysis



Data governance



# SURVEY SAYS



## GenAI use cases organizations are considering for data access and preparation

*Among those currently using or planning to adopt GenAI, N=499*



**58%**

Automatic data  
curation



**51%**

Conversational  
analytics



**51%**

Data tests and  
quality



**46%**

Speed data  
access with text  
to SQL



**36%**

Documentation



# Platform Modernization

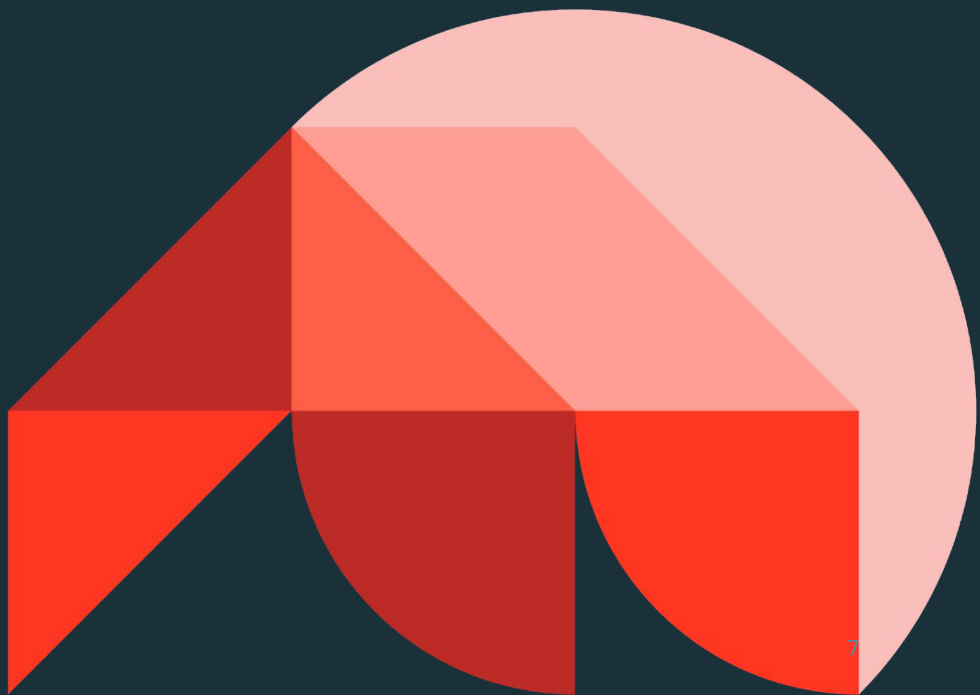


# Modernizing Data + AI for the Enterprise

---

May 2025

Brian Suk, Senior Solutions Architect, Databricks



Every company wants to be a

**Data + AI**

company

# Leaders see positive results

1

**Reduced TCO** to  
enable other  
investments

2

High **quality** data  
that is **secure** and  
**consumable**

3

**Transformational AI**  
for their organization  
and business





But...

---

Finding success at scale  
is difficult for most  
organizations



# Data estate is fragmented

Governance



Orchestration  
and ETL



Data Lake



Data  
Science



Streaming



Data  
Warehouse



Generative  
AI



It's a complexity  
nightmare of **high costs**  
and **proprietary formats**

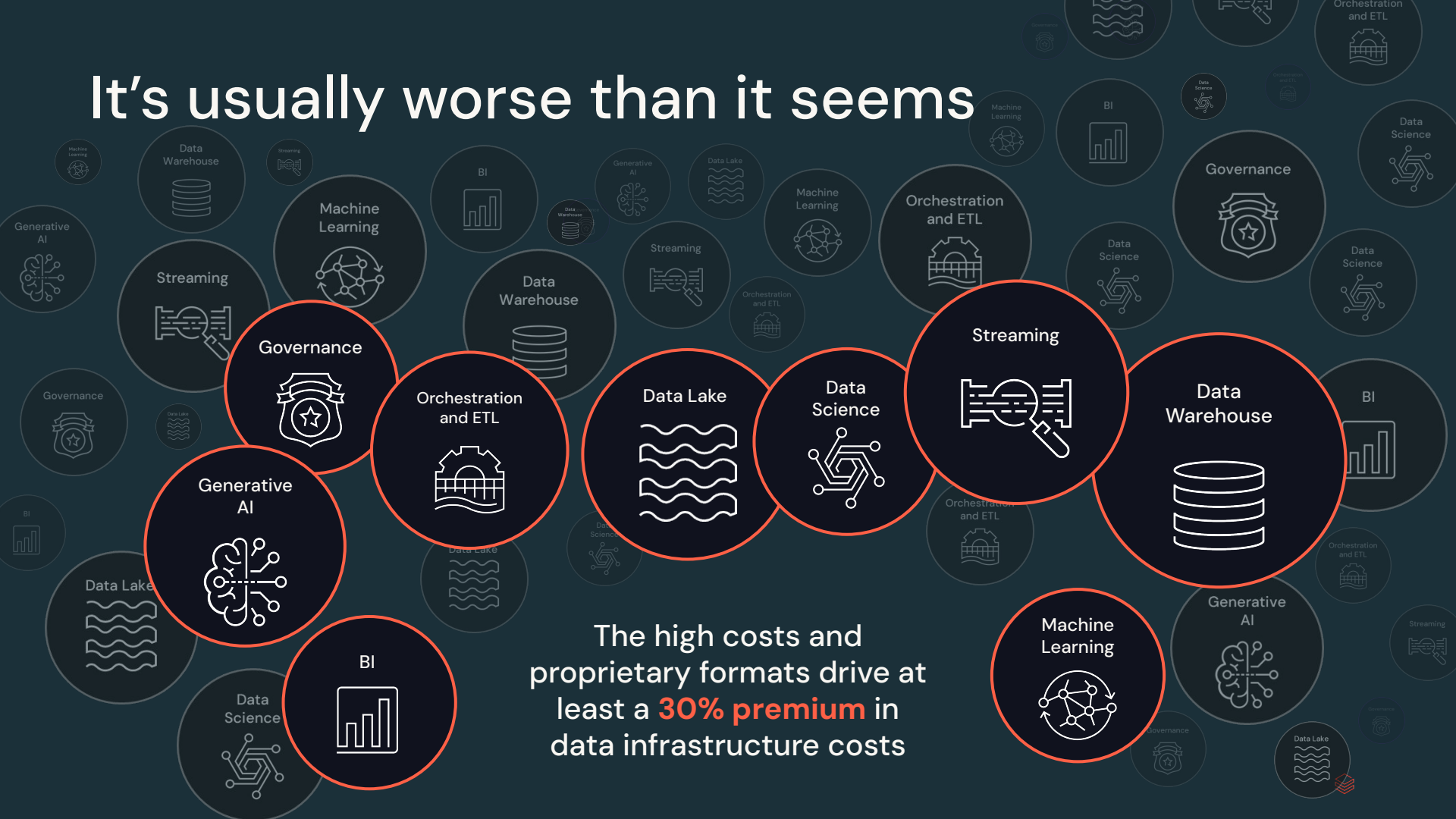
BI



Machine  
Learning



# It's usually worse than it seems



Governance



Orchestration  
and ETL



Data Lake



Data  
Science



Streaming



Data  
Warehouse



Generative  
AI



BI

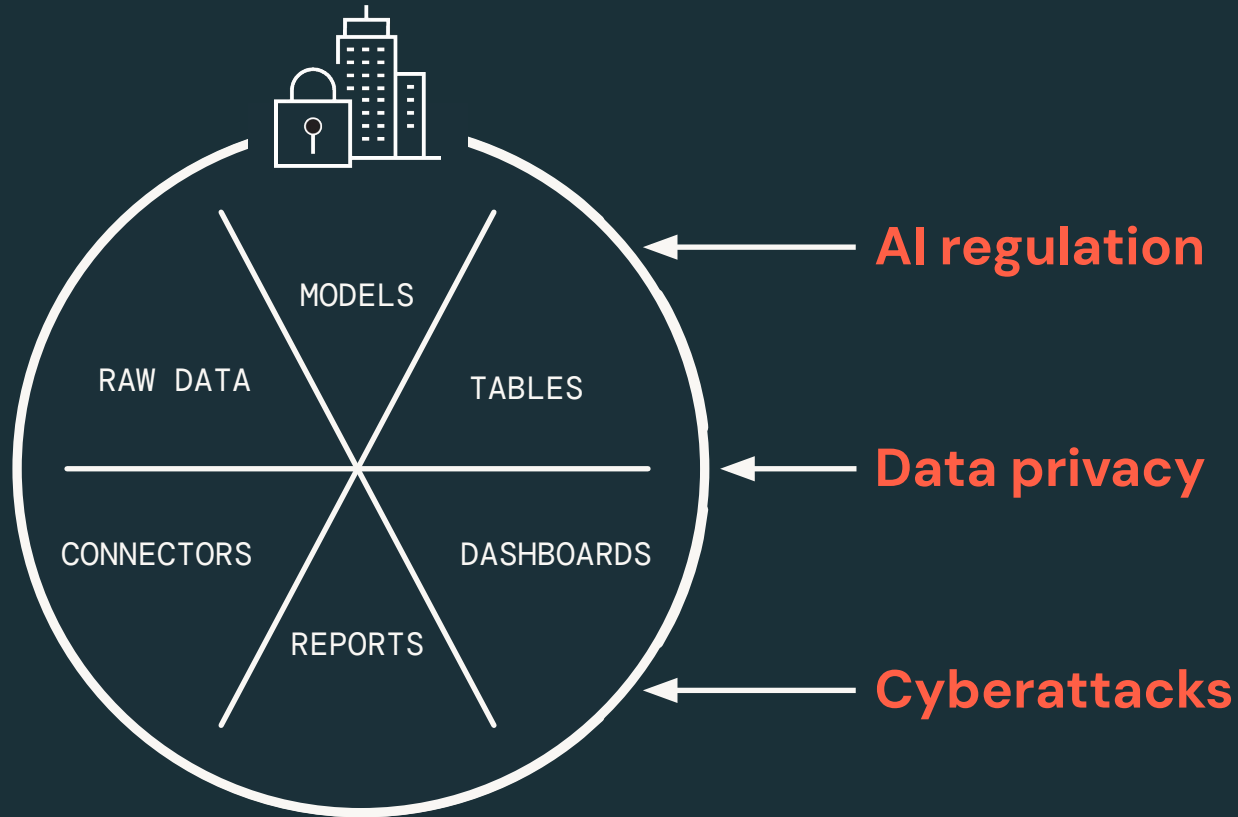


The high costs and  
proprietary formats drive at  
least a **30% premium** in  
data infrastructure costs

Machine  
Learning



# Governance of the entire data estate is hard



# Production AI that reasons on your data is elusive

**Experimentation**

Quality  
Cost  
Privacy

**Production**

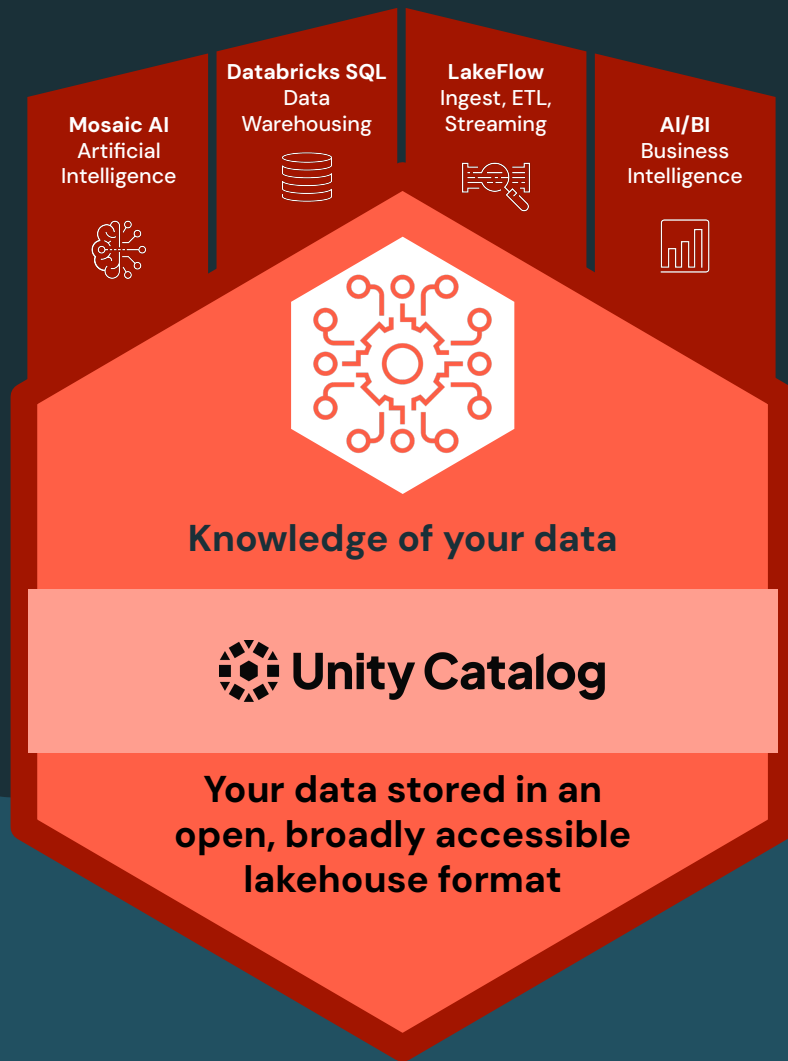
**85% do not make it  
into production**

**85% dollars spent  
haven't shown a return**



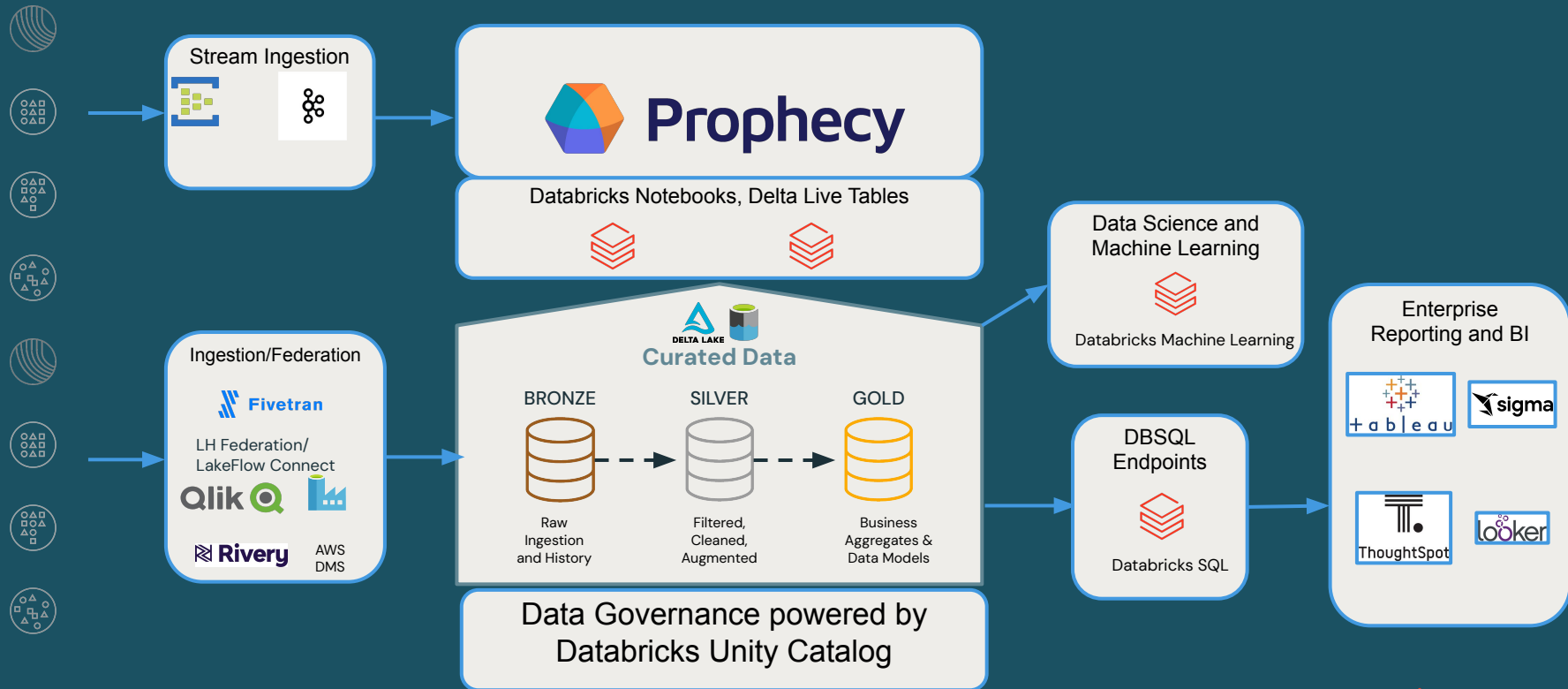


# Data Intelligence Platform



# Modern ETL Ecosystem on Databricks

Select the Ingestion, ETL, Presentation Layer and Governance Ecosystem on the Databricks Platform





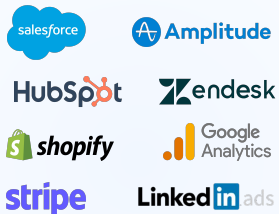
# Enable all data users

# Clean, trusted data fuels AI & analytics

Organizations that use high-quality, well-governed data are 3x more likely to see AI success (source: Gartner)

## Operational Data

### Applications



### OLTP Databases



Data  
Engineer



Data  
Analyst

Ingest

Filter, clean,  
augment

Aggregate

## Analytics



Business Intelligence

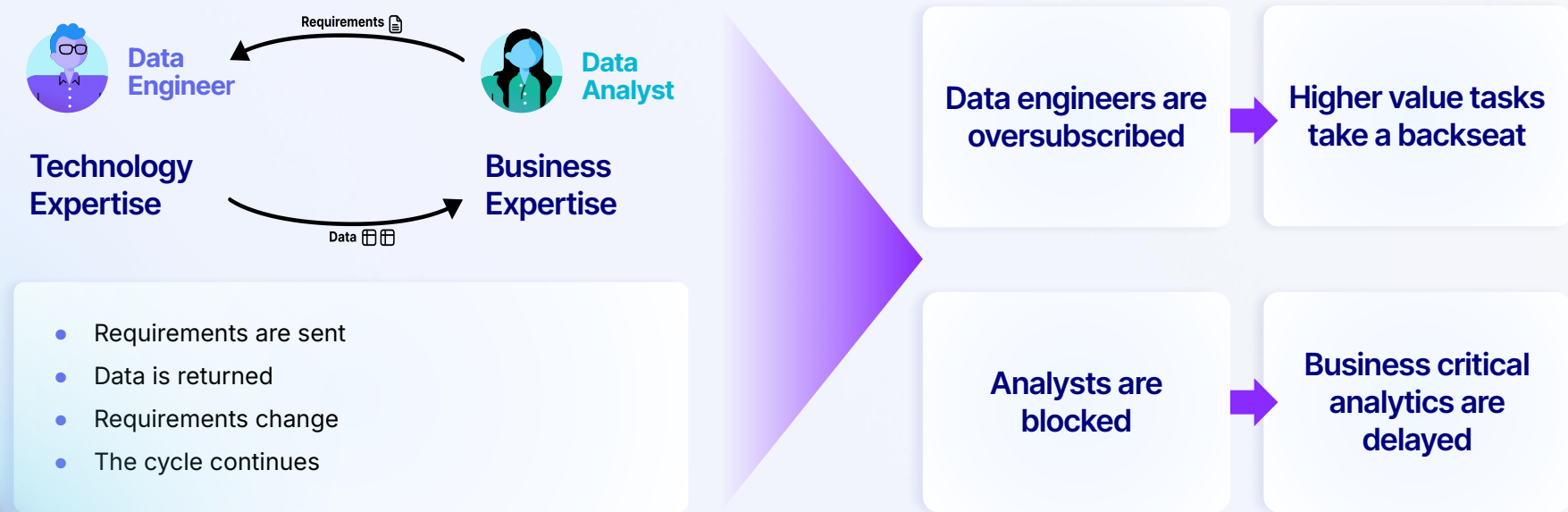


Reporting



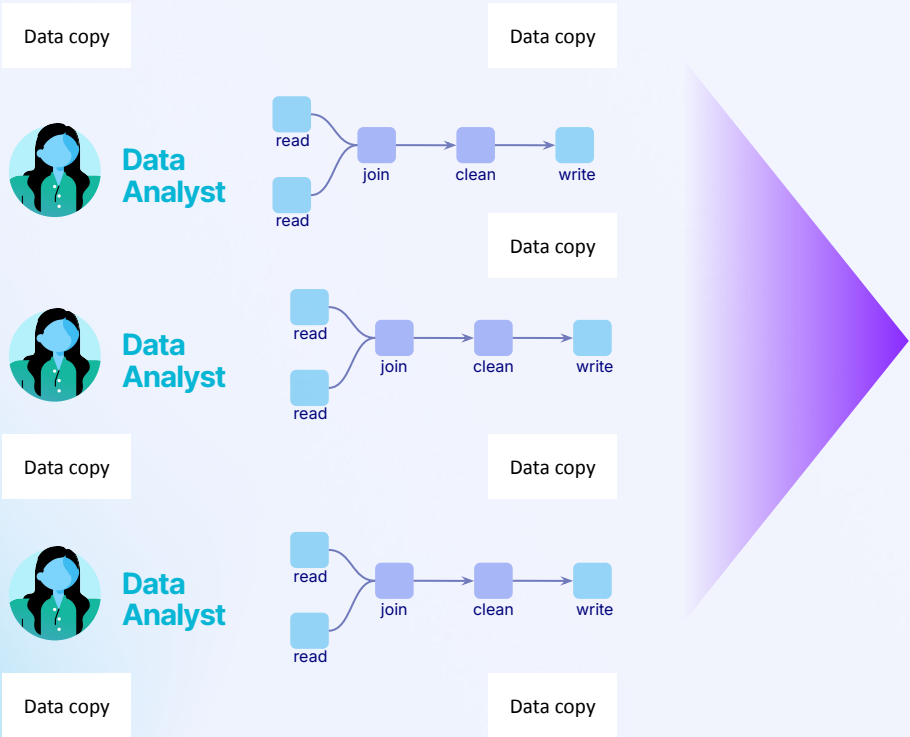
Machine Learning,  
Generative AI

# Data engineers can code transformations... But at what cost?





# Analysts can prepare data on their own... But now you've introduced more risk.



## Governance risk



### Cost

Engineering rework, as analysts use tools not built for production.



### Quality

Are analysts using the right data?  
Following company standards?



### Security

Data copies make permissions management a nightmare.



# What the business demands

## Self-service

Data analysts prepare data  
on their own

Data  
Analyst



+

## Governance

Working within guardrails set by the  
central data platform team



Data  
Engineer



# Prophecy Data Integration Platform

The data integration platform for data analysts and data engineers



Data integration platform

Business logic - code on git

Cloud Data Platform



## Self-service

Data analysts need an interface that is...

Intuitive

Familiar

End-to-end



## Governance

Central data platform teams ensure guardrails for...

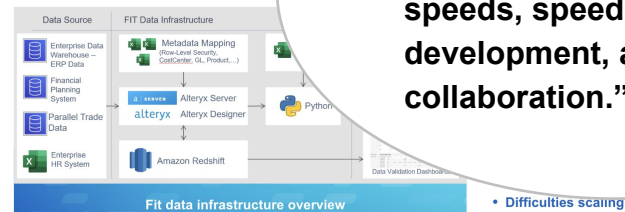
Cost

Quality

Security



## The intermediate solution

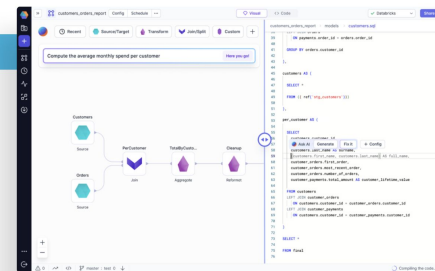


“With Prophecy we were able to fully leverage cloud data platform processing speeds, speed pipeline development, and improve collaboration.”

## The solution: Prophecy

All the benefits of Databricks with:

- + Drag & drop functionality
- + Easy to use GUI
- + AI Copilot for rapid development

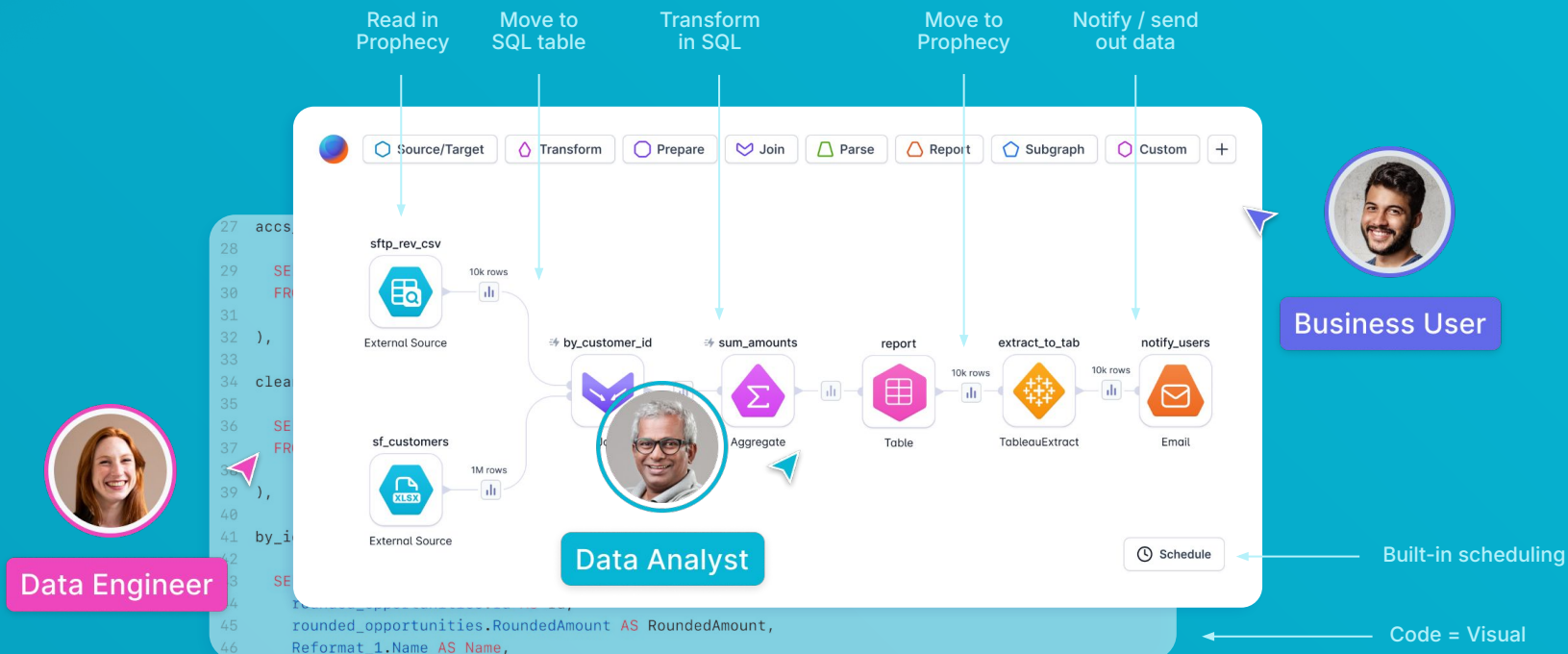


**Drew Davis**  
Principal Organizational Planning  
Insights & Analytics  
Amgen



# Prophecy Single Canvas - Unified Data Stack

Single collaborative canvas with ingest, transform, orchestrate & observe.





# Copilot productivity for every team member

Higher Productivity per User  
Focus on Analytics

Makes recommendations

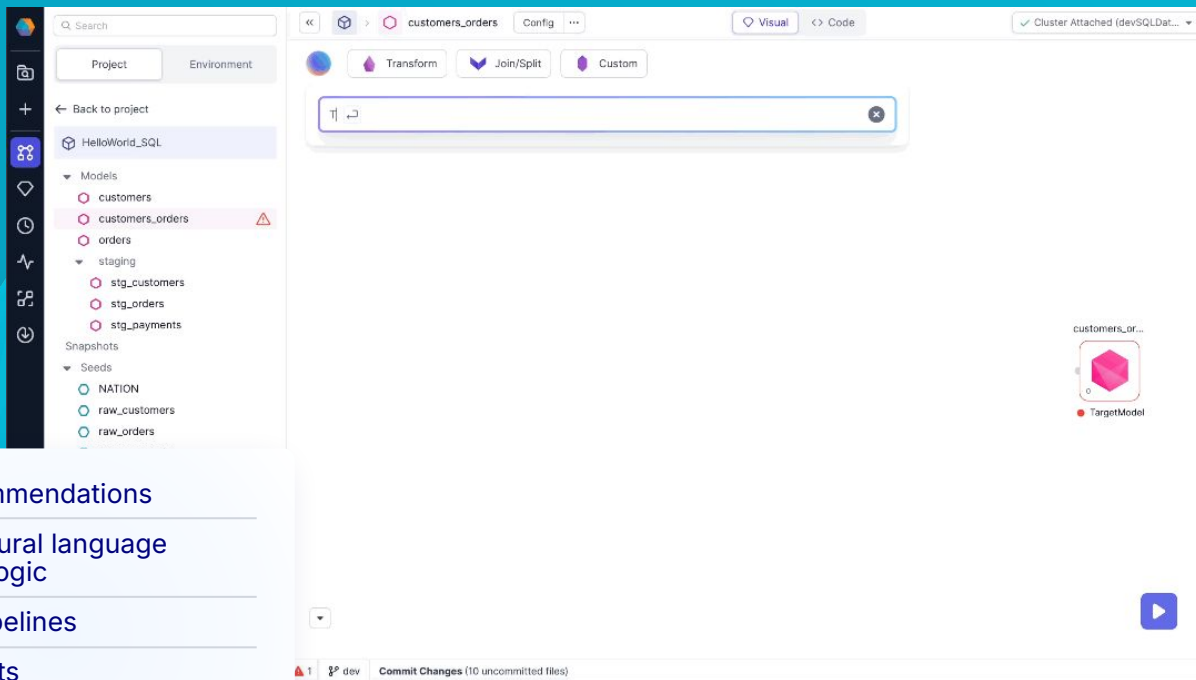
Converts natural language to business logic

Complete pipelines

Generate tests

Writes documentation

Suggests fixes for errors





# Demo



# Prophecy

## The Data Integration Platform

Visual Interface

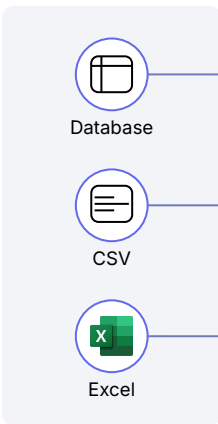
Artificial Intelligence

Compiler Technology

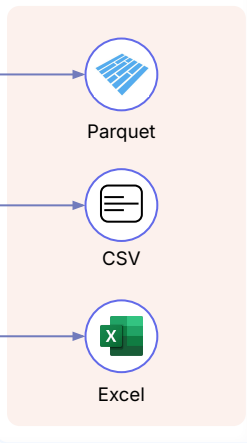


Data Engineer

Builds the unified model in Low Code Spark/SQL and lands to Unity Catalog



Cloud Storage



Bronze



Unity Catalog

Unified Model

User

Post

Comment

Silver

Databricks SQL

Data Mart 1

Data Mart 2

Data Mart 3

Custom GenAI

Vector DB lookup

Prompt template

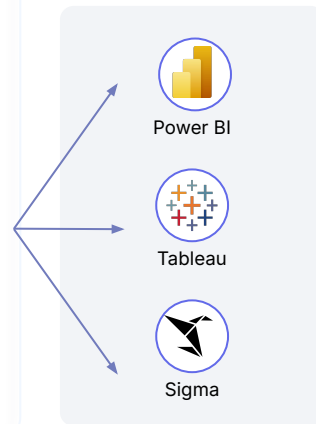
LLM

Gold



Data Analyst

Builds "last-mile" transformations in Low Code SQL to provide data for reports and analytics



# Questions



# Prophecy at Databricks Data and AI Summit 2025

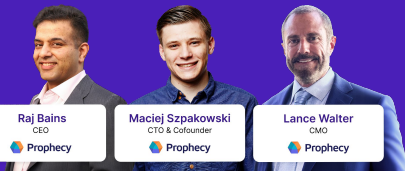
June 9-12, 2025 | Moscone Center | San Francisco, CA | Booth #E515



Learn more about our presence at DAIS

## SESSION

Reinventing data prep with AI:  
Build an agent-driven  
production pipeline in 7 minutes



Above the Clouds  
Data + AI Party.



Prophecy will be center stage with:

- AI agents, taming FHIR industry data, GenAI survey
- Booth chats with industry experts
- The View event party

# Thank You!

