



Unlocking Productivity Gains with AI

Fast insights for
everyone

Felix Mutzl

Munich | May 2025

©2025 Databricks Inc. — All rights reserved

print("hello world!")

Grew up in manufacturing. Built data platforms.



BUILD



ADOPT



History of Databricks



2013

Founded at
UC
Berkeley

2015

Pivot from PLG
to Enterprise
B2B

2018-19

Launched more
open Formats + on
Azure

2020-21

Defined Lakehouse,
and began scaling
(2 rounds in 2021)

2022-23

Expand Gen AI and
Data Intelligence
(Series I in Nov, 23)

2024-25

SAP Databricks
(still Private with
\$10b Series J)



Every company wants to be a

Data + AI

company

But...

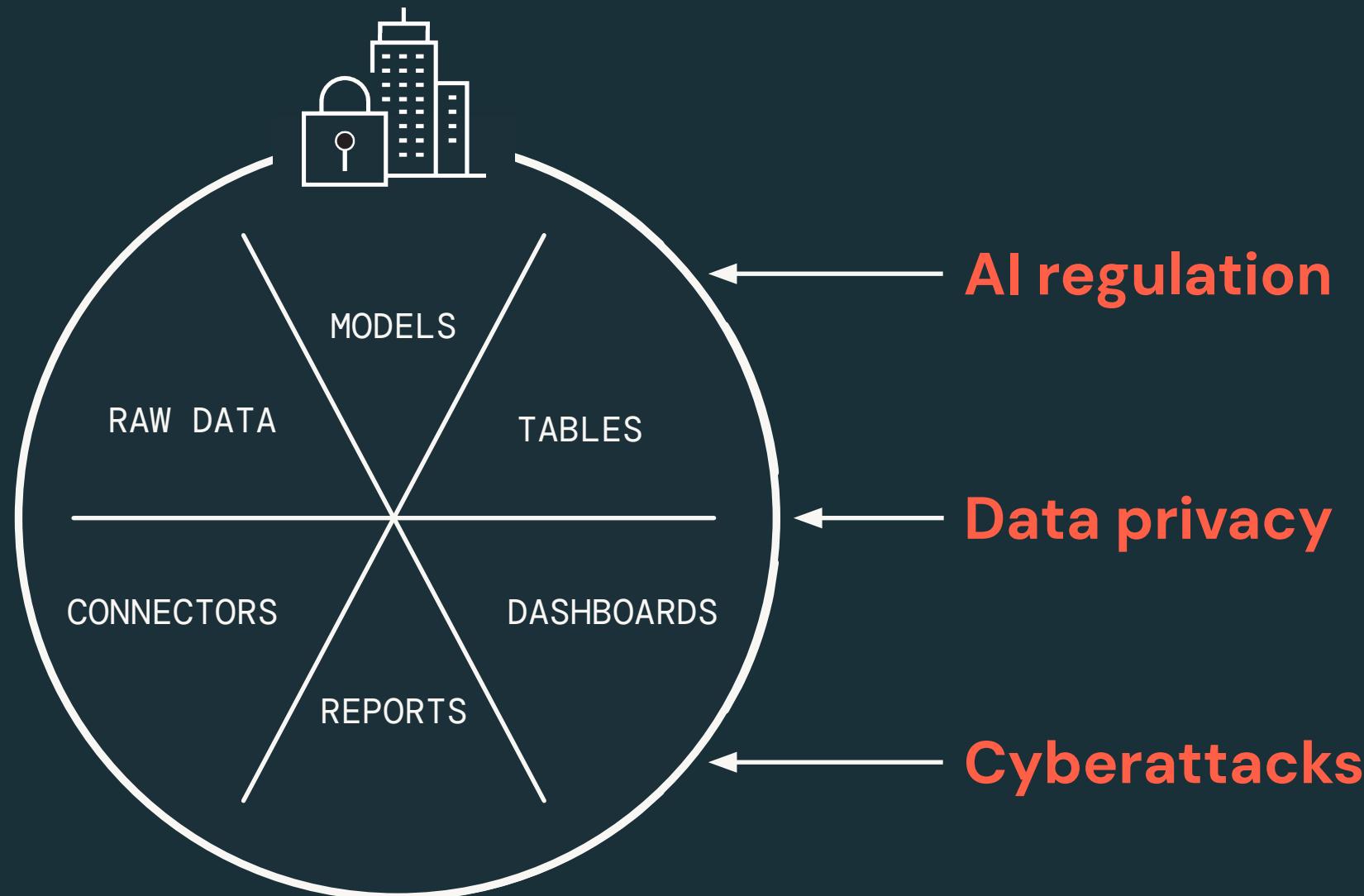
Finding success at scale is
difficult for most organizations



It's usually worse than it seems

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

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



2020

A disruptive approach

What Is a Lakehouse?



by [Ben Lorica](#), [Michael Armbrust](#), [Reynold Xin](#), [Matei Zaharia](#) and [Ali Ghodsi](#)

January 30, 2020 in [Engineering Blog](#)

Share this post



Over the past few years at Databricks, we've seen a new data management architecture that emerged independently across many customers and use cases: [the lakehouse](#). In this post we describe this new architecture and its advantages over previous approaches.

Data warehouses have a [long history](#) in decision support and business intelligence applications. Since its inception in the late 1980s, data warehouse technology continued to evolve and MPP architectures led to systems that were able to handle larger data sizes. But while warehouses were great for structured data, a lot of modern enterprises have to deal with unstructured data, semi-structured data, and data with high variety, velocity, and volume. Data warehouses are not suited for many of these use cases, and they are certainly not the most cost efficient.



2020

Databricks pioneered the lakehouse architecture



Today

74% of global enterprises have adopted lakehouse

MIT Technology Review Insights, 2023



Unify and own your data

Reduce unnecessary costs
from multiple copies of data

Break free from lock-in
with open formats you control

Adopt the lakehouse data formats that are right for your organization



Iceberg



Delta Lake



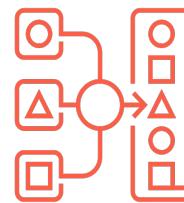
Parquet

Get the right data without the hassle



databricks Lakeflow

Connect



Easily connect key data sources



Google Analytics



salesforce
ORACLE
NETSUITE



workday



Microsoft
SQL Server



Dynamics 365



serviceNow

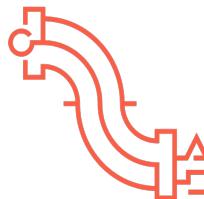


SharePoint



PostgreSQL

DLT



Reliable data pipelines made easy



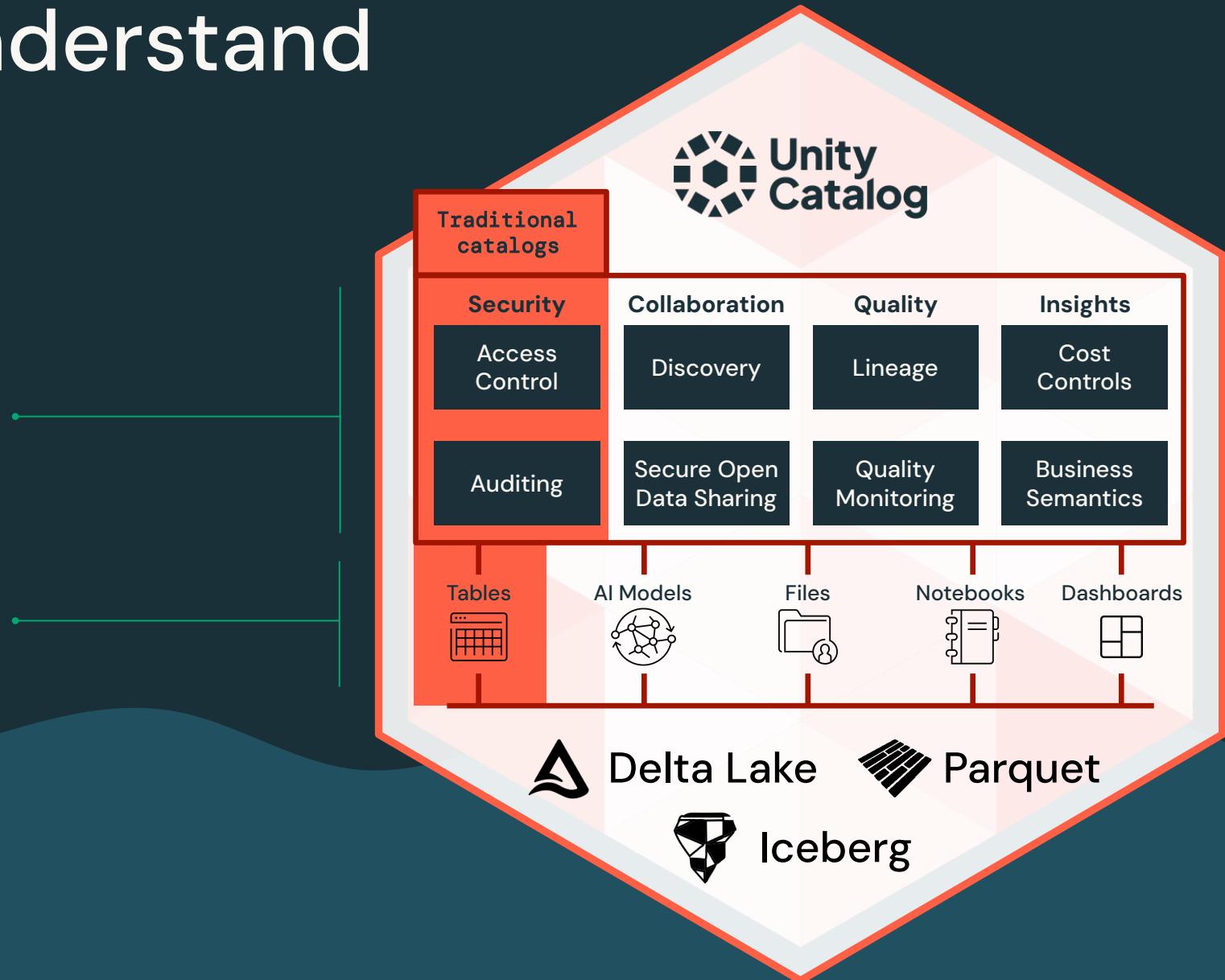
Jobs

Unified orchestration for analytics and AI

Protect and understand your data

Unified capabilities
for **every use case**

Unified governance
for **all assets**



“I don’t care about the standard
benchmarks;
I want the model to do well on my
data.”

General Intelligence

Large models
trained on the entire web
leveraging scaling laws

vs

Data Intelligence

AI agents that **reason**
on your customer data
and solve domain-specific
problems



Data Intelligence Platform

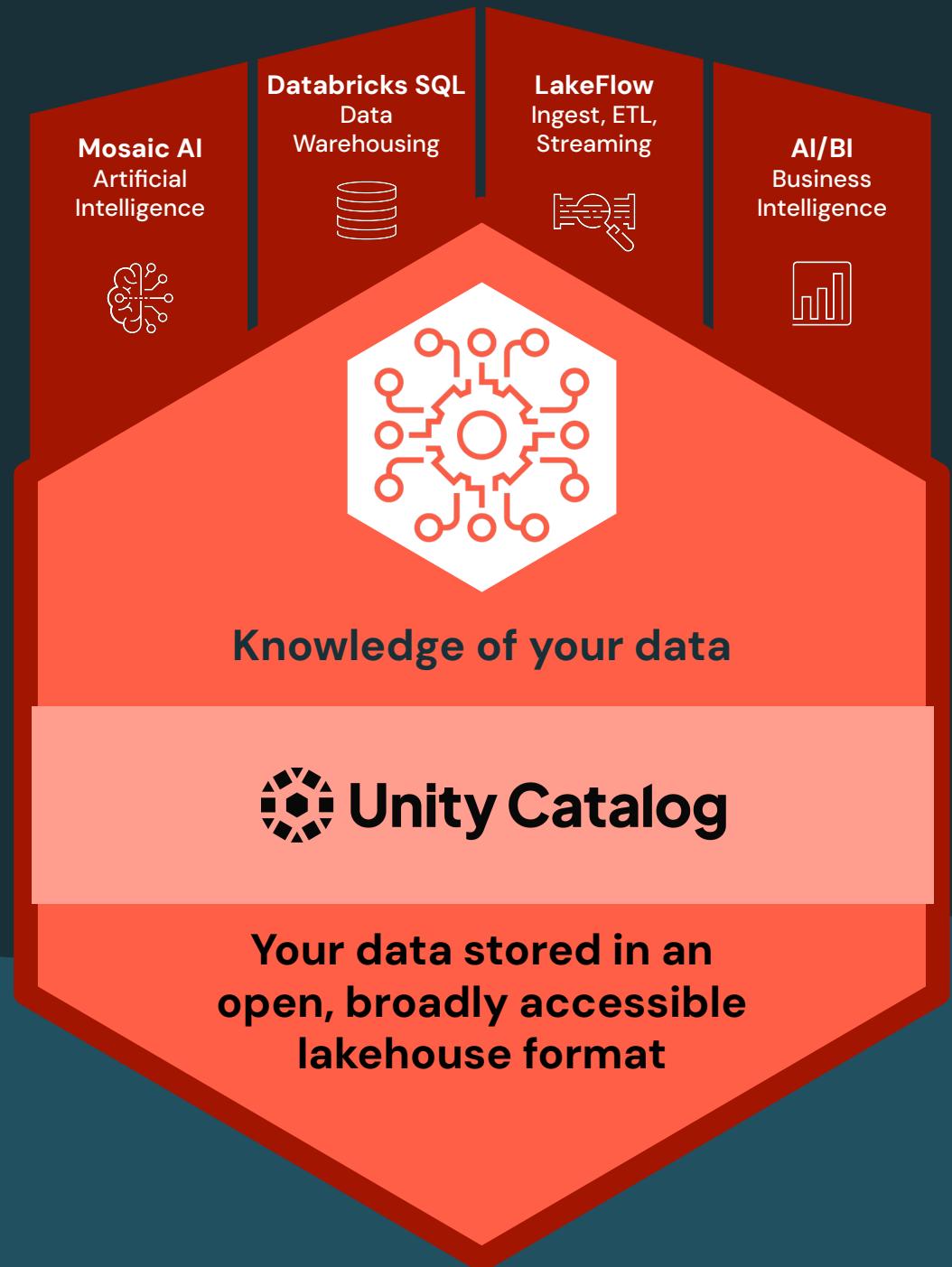
Unified data
and governance

+

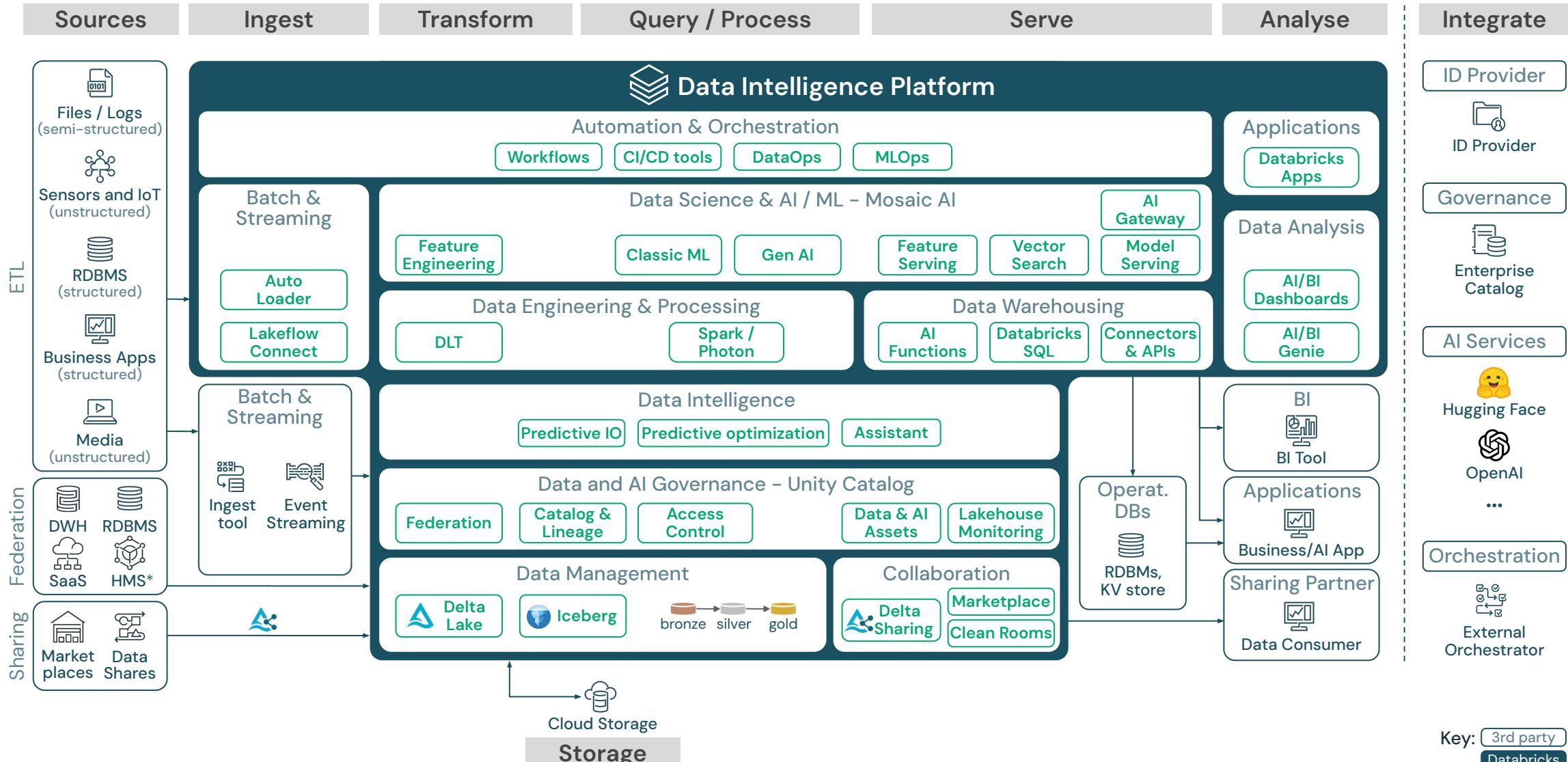
AI tuned to
your business



Data Intelligence Platform



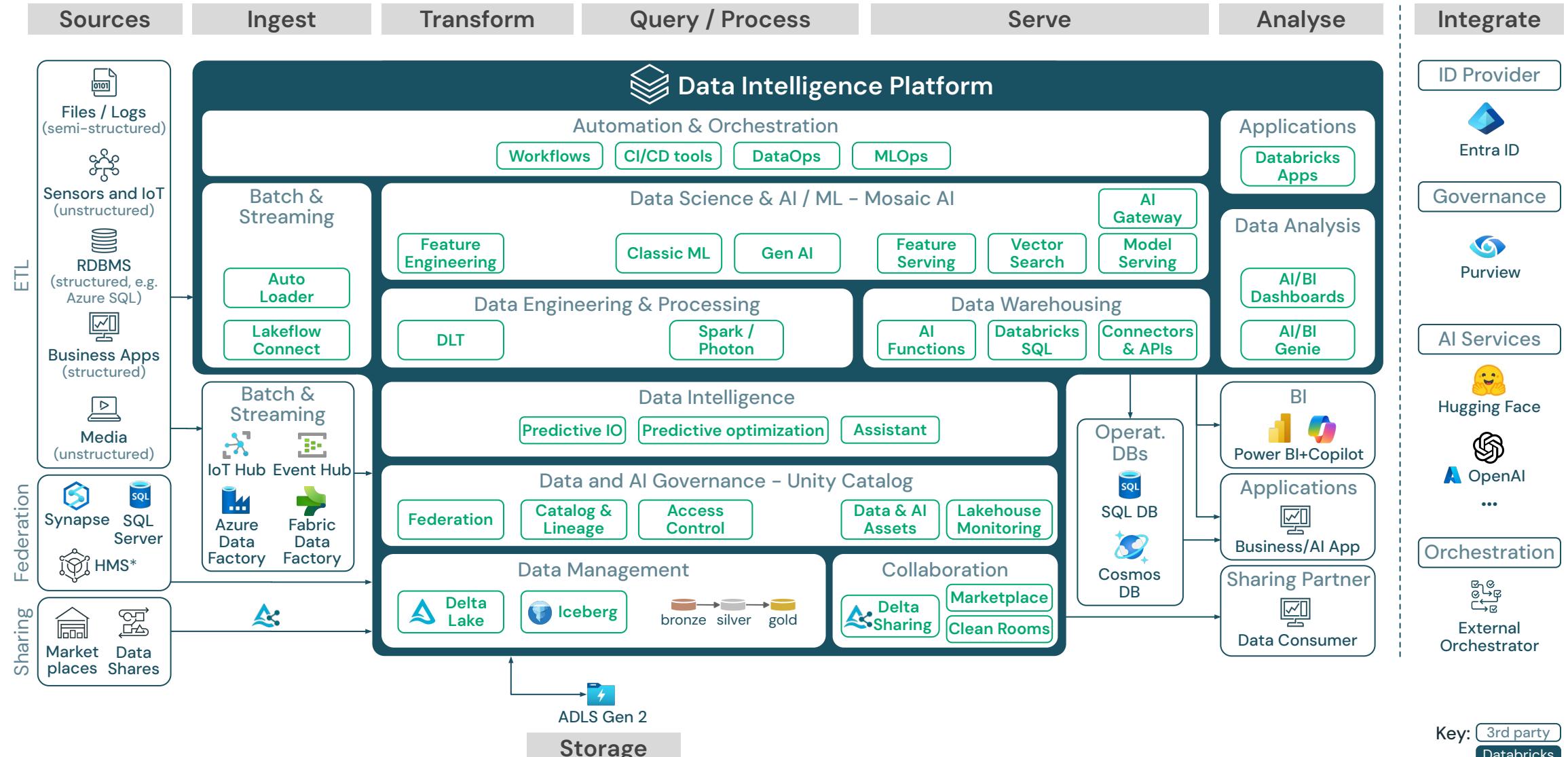
Databricks Data Intelligence Platform



* Hive Metastore



Databricks Data Intelligence Platform on Azure



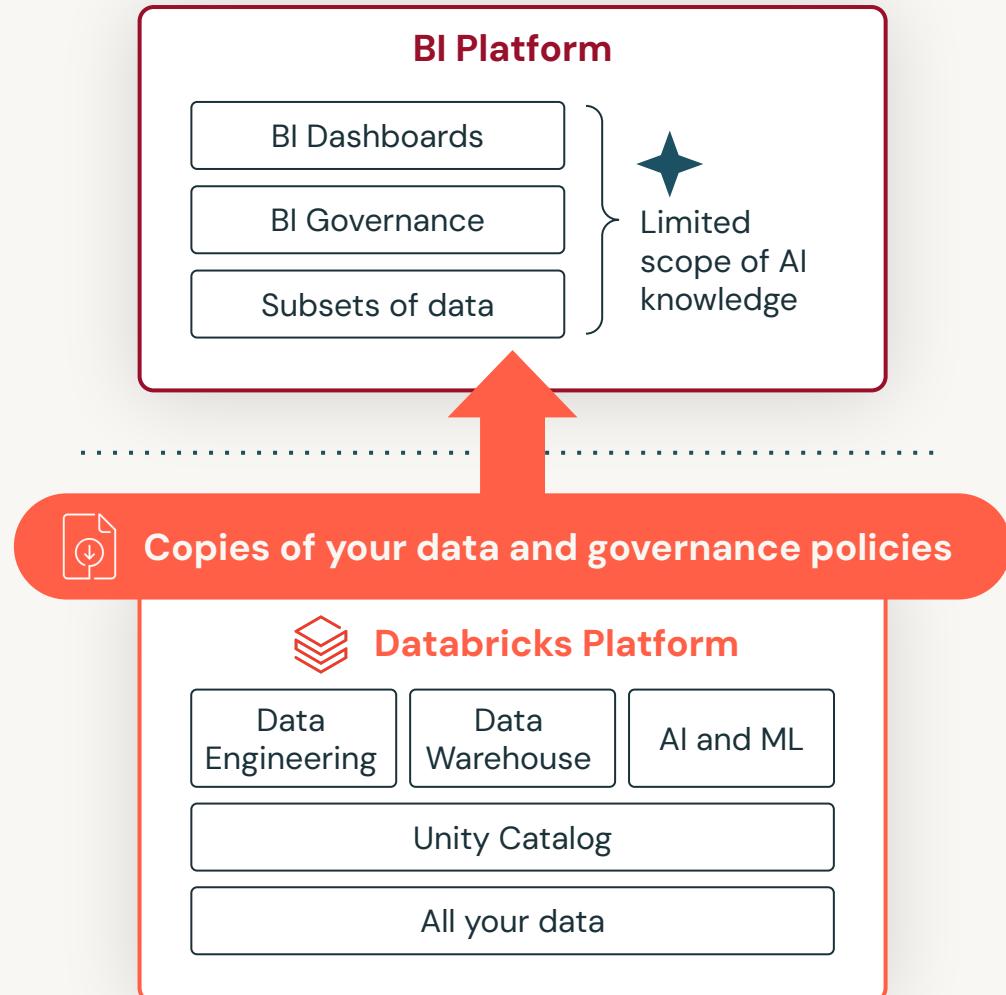
* Hive Metastore



What is
Data Intelligence

for analytics and BI?

BI silos create barriers for self-service analytics



Need to extract data to optimize performance

- Developing extracts for BI server delays time to insight
- Creates multiple “versions of the truth”
- Extracts limit range of available analytics (e.g., real-time)

Separate governance models to keep in sync

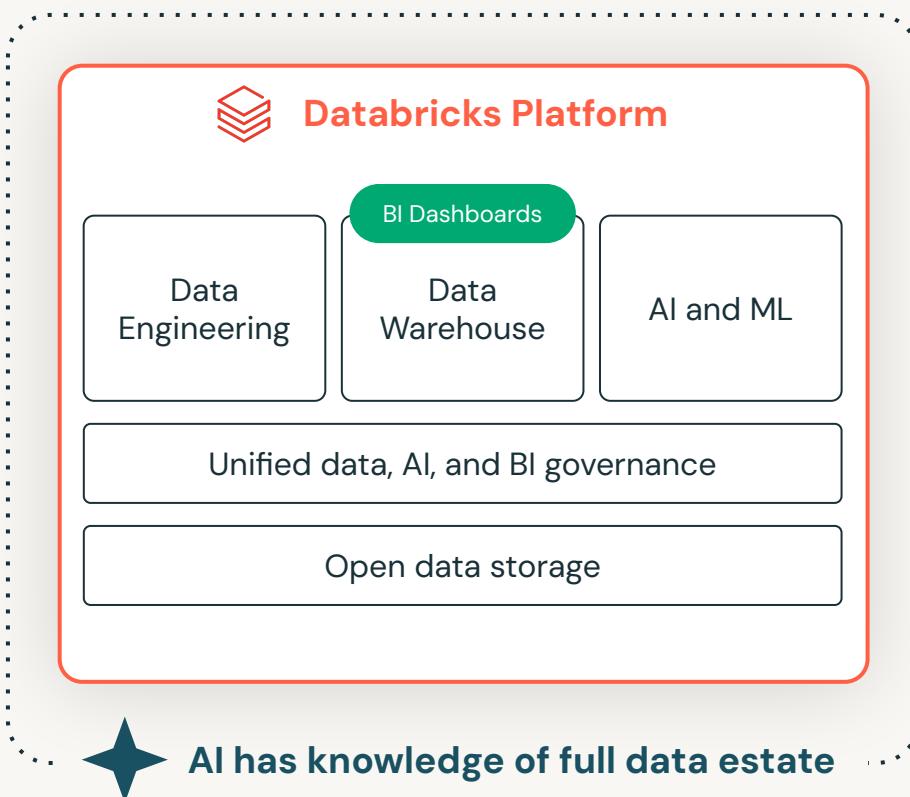
- Replicating data and access policies introduces risk
- Breaks lineage from raw data to final dashboards
- No single place to govern data and BI assets

AI assistants have limited scope of knowledge

- Lacks knowledge of full data estate, usage and semantics
- Requires perfectly modeled data
- May hallucinate answers to unanticipated questions



Native BI simplifies the path to democratization



Fast dashboards with no data extracts

- Get new insights without waiting on the data team
- All data available to answer any ad hoc question
- Analytics built on real-time data

Unified governance for data, AI, and BI

- No need to replicate data security and access policies
- Full e2e lineage from raw data to final dashboards
- One place to govern all data, AI and BI assets

AI assistants understand your entire data estate

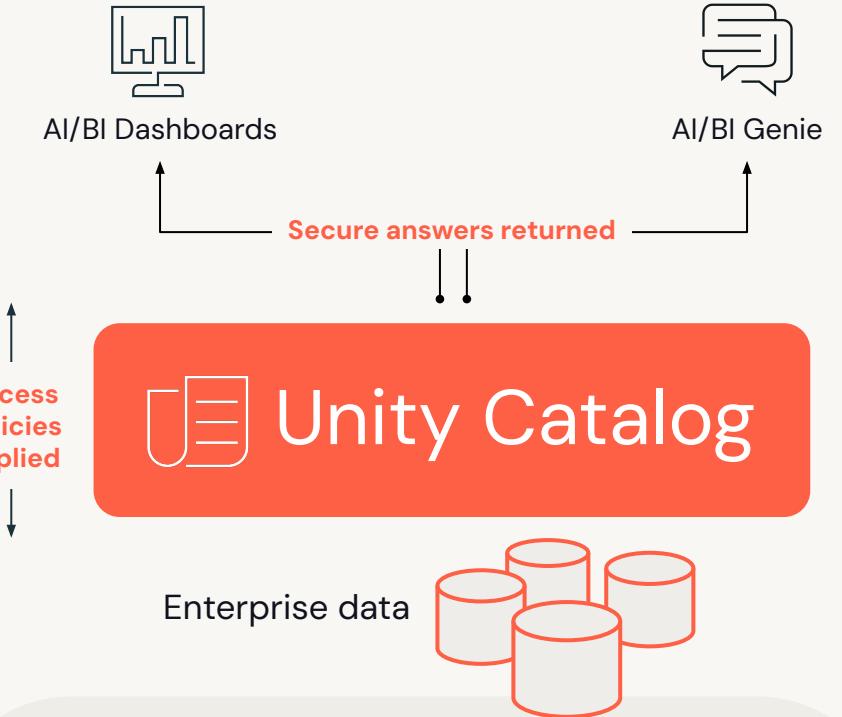
- Intelligence on all data, usage patterns and business concepts
- Works with messy data without extensive modeling
- Seeks clarification when unsure and learns from user feedback



Governed and secure AI-generated insights

The screenshot shows the Databricks Genie Spaces interface. In the top left, it says "databricks Genie Spaces". Below that is the title "Hanlin Sample Sales Room". A message states "This room is to analyze Sample CRM data". On the left, there's a sidebar with icons for Monitoring, Messages feed, and other options. The main area is titled "Monitoring" and contains a "Messages feed" table. The table has columns for Status, Question, Rating, User, and Created. The "Status" column shows all entries as "Completed". The "Rating" column includes thumbs up and thumbs down icons. The "User" column lists Hanlin Sun and chao.cai@databricks.com. The "Created" column shows dates from May 28, 2024, to May 29, 2024. At the bottom, there are navigation buttons for "Previous", "Next", and "50 / page".

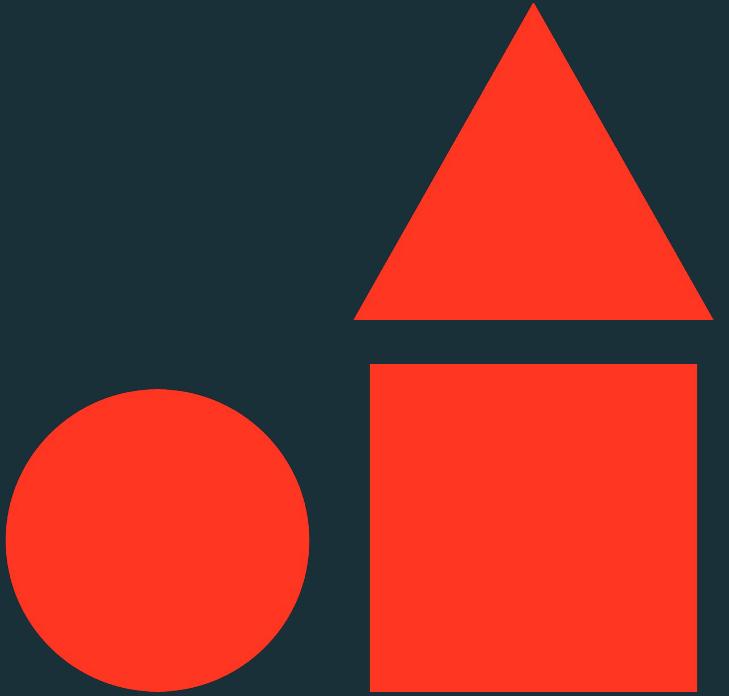
Use **Genie monitoring** to review the questions users asked. Evaluate thumbs up/down ratings to **improve accuracy for future questions.**



Genie honors access policies established in UC to ensure **only secure data returned to end users**



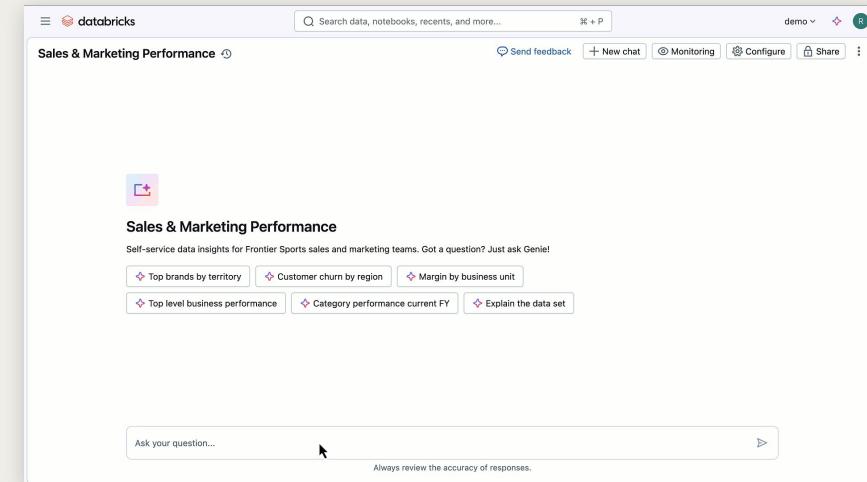
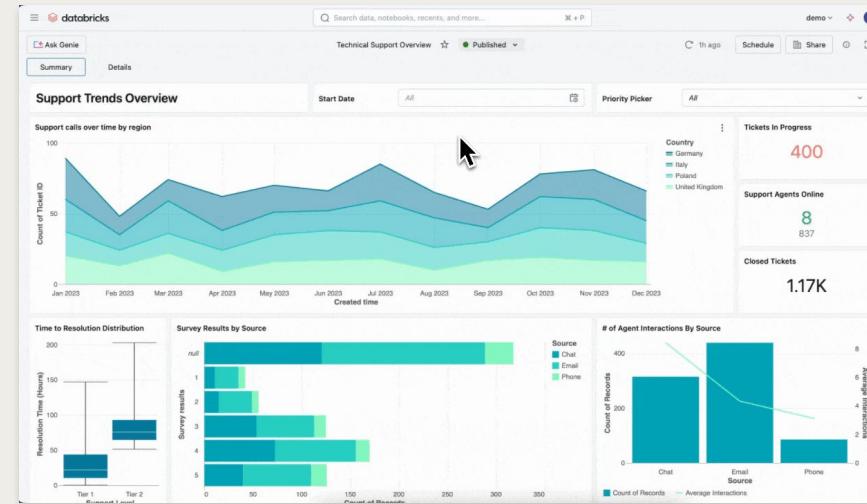
Demo



Powered by Databricks SQL

Introducing Databricks AI/BI

Intelligent analytics
for everyone



Dashboards

Fast and secure
visual insights for your
entire organization

Genie

Talk with your data
using natural
language

Governed and secured with Unity Catalog



What is
Data Intelligence

for AI applications?

2025 is the year of AI Agents

“...state-of-the-art AI results are increasingly obtained by compound AI systems with multiple components, not just monolithic models.



[read here](#)



BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH [Subscribe](#) [About](#) [Archive](#) [BAIR](#)

[The Shift from Models to Compound AI Systems](#)

Matei Zaharia, Omri Khamis, Lingjiao Chen, Jarell Quincy Davis, Heather Miller, Chris Pettis, James Zou, Michael Corbin, Jonathan Franklin, Mavrey Boo, Ali Ghodsi Feb 18, 2024

AI caught everyone's attention in 2023 with Large Language Models (LLMs) that can be instructed to perform general tasks such as translation or coding. Just by pretraining, this technology has matured to become a key component of many applications. As AI continues to evolve, with everyone wondering what capabilities new LLMs will bring. As more research is done, it is becoming clear that the future of AI lies in compound systems, where state-of-the-art AI results are increasingly obtained by compound systems with multiple components, not just monolithic models.

The example, Google's AlphaCode 2 set state-of-the-art results in programming through a carefully engineered system that uses LLMs to generate up to 1 million possible solutions for a

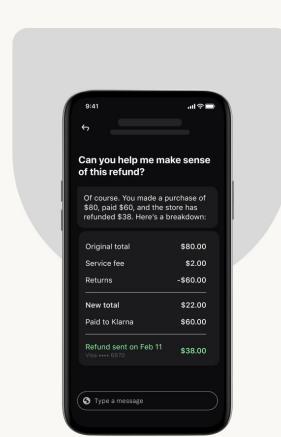
[read here](#)



Helping contact center teams improve customer satisfaction

“The GenAI agent assist tool quickly delivered accurate information to [the contact center rep]. This led them using it more often and to suggest additional use cases to make their jobs easier and spend more time with customers.”

[Source](#)



“This AI breakthrough... means superior experiences for our customers at better prices, more interesting challenges for our employees, and better returns for our investors.”

– Sebastian Siemiatkowski, co-founder and CEO of Klarna

Klarna

- **\$40M USD in profit improvement**
- Resolution time from **11 mins to 2 mins**
- **On par** with human agents in **customer satisfaction score**

[Source](#)



What is an Agent?

How would a human do this?

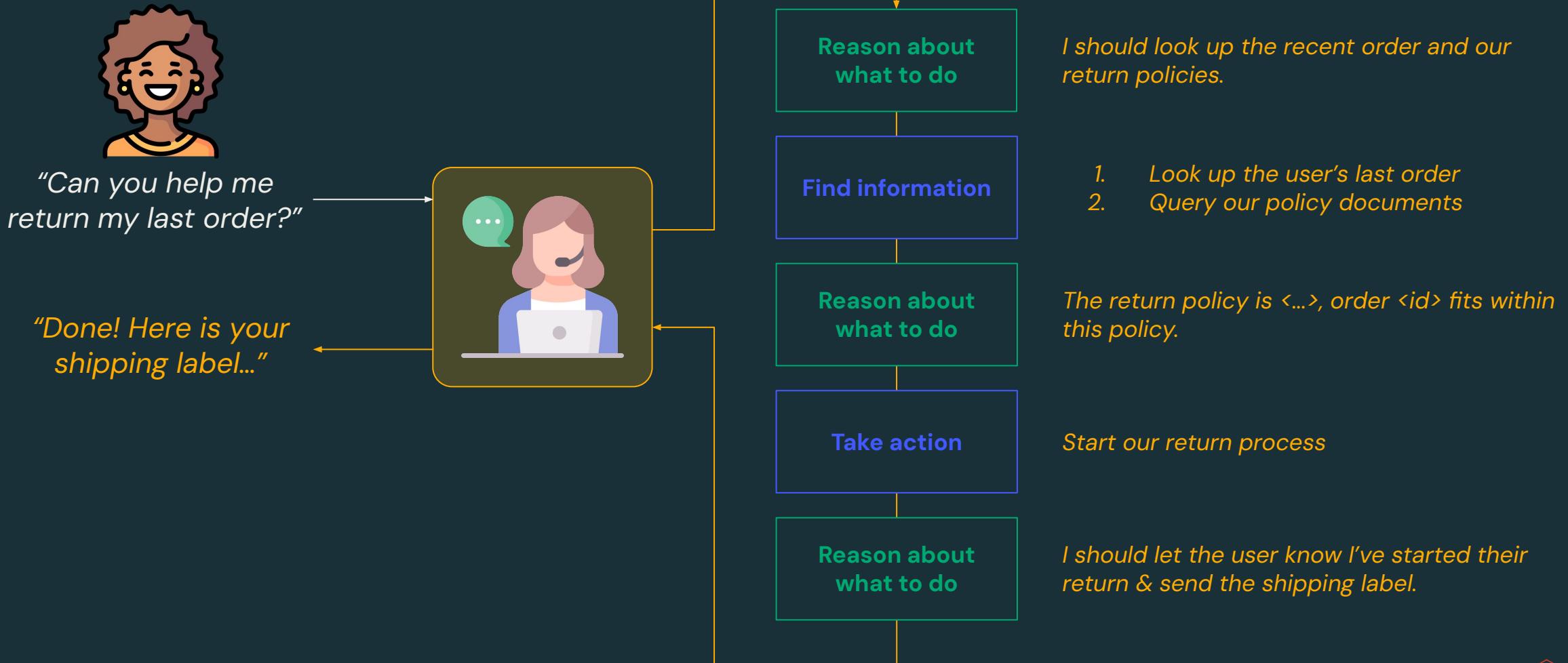


*"Can you help me
return my last order?"*



What is an Agent?

How would a human do this?



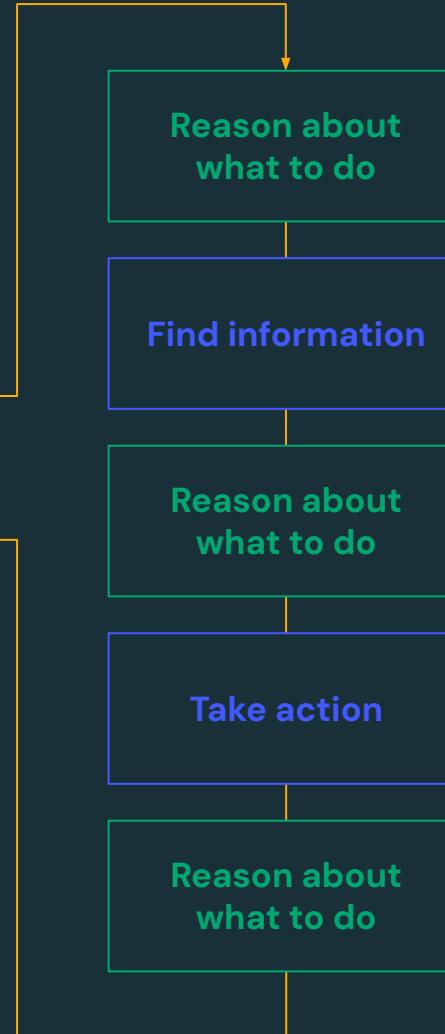
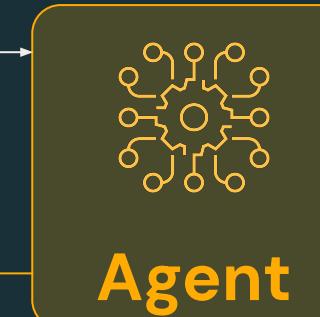
What is an Agent?

An Agent does the same thing ... an Agent is a LLM acting if it has a brain



"Can you help me return my last order?"

"Done! Here is your shipping label..."



I should look up the recent order and our return policies.

1. *Look up the user's last order*
2. *Query our policy documents*

The return policy is <...>, order <id> fits within this policy.

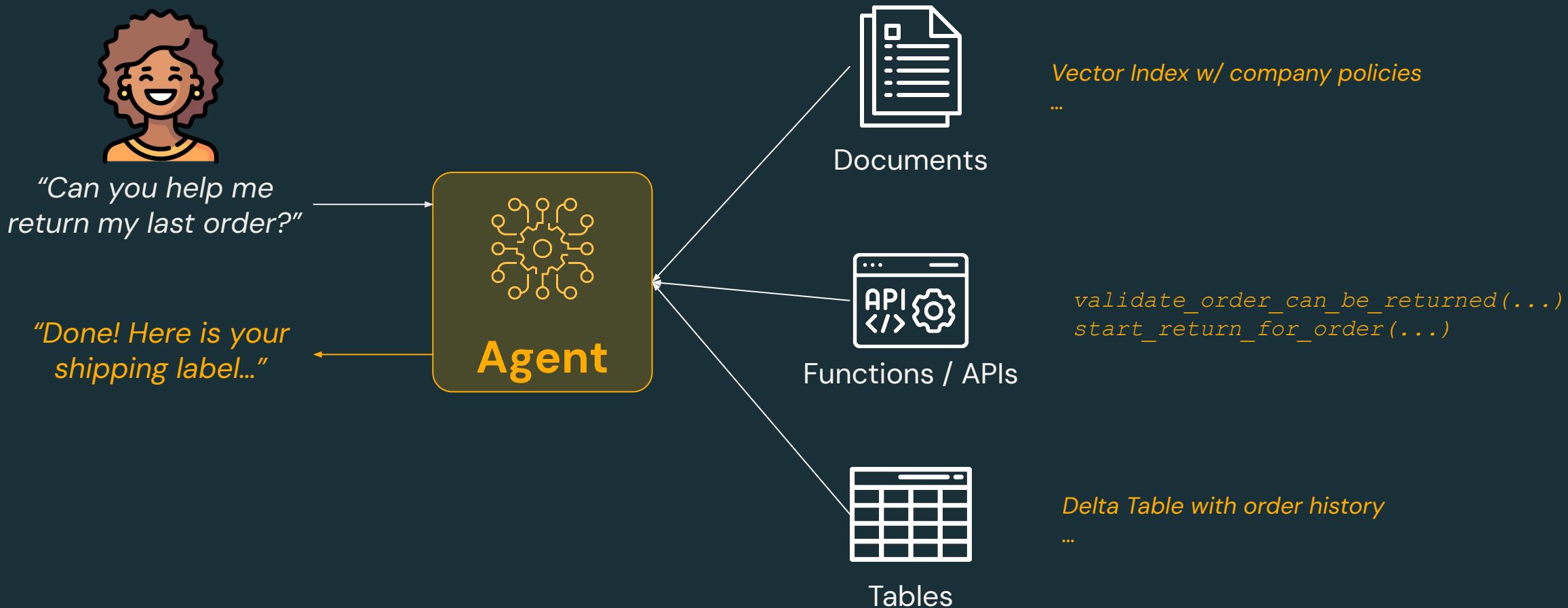
Start our return process

I should let the user know I've started their return & send the shipping label.

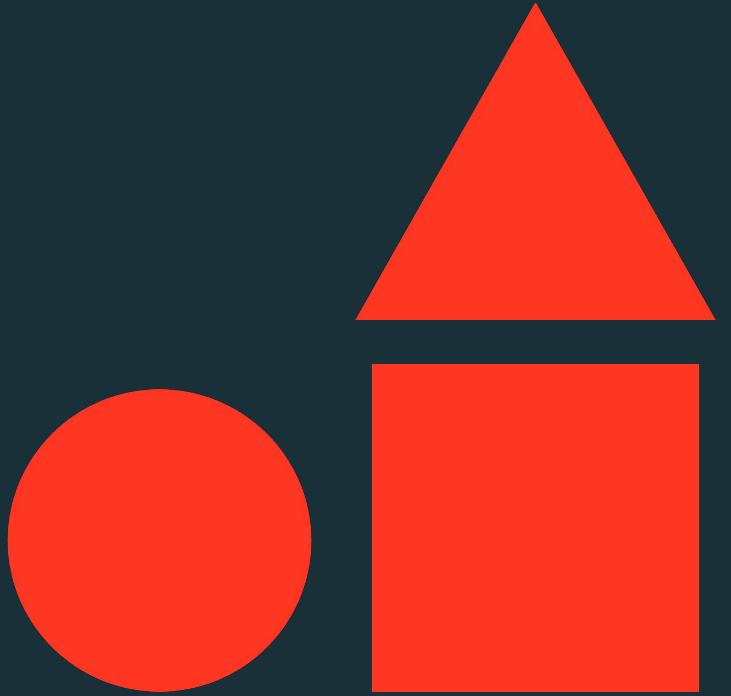


What is an Agent?

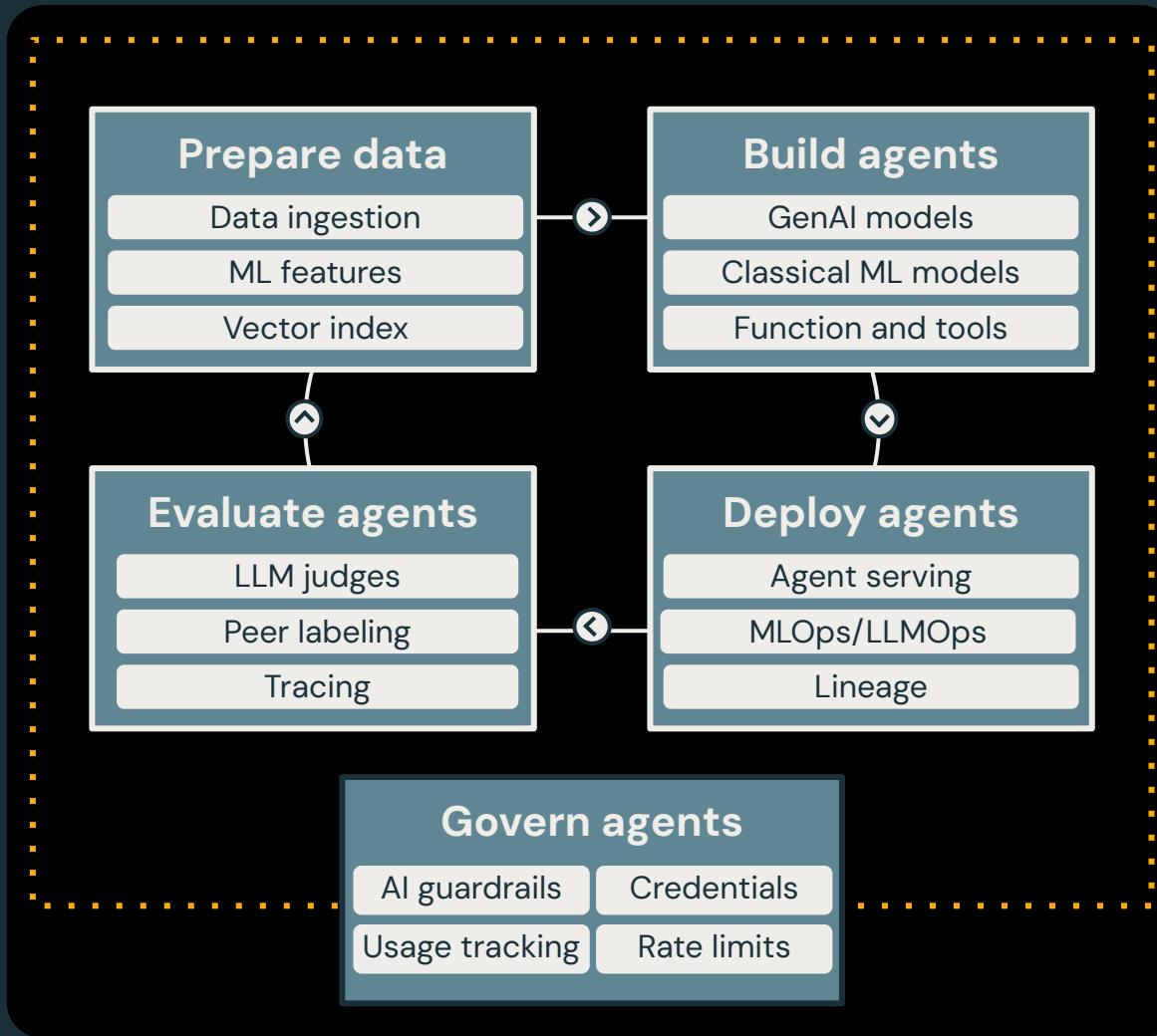
An Agent does this by retrieving your data and calling your APIs



Demo



The most complete agent systems platform



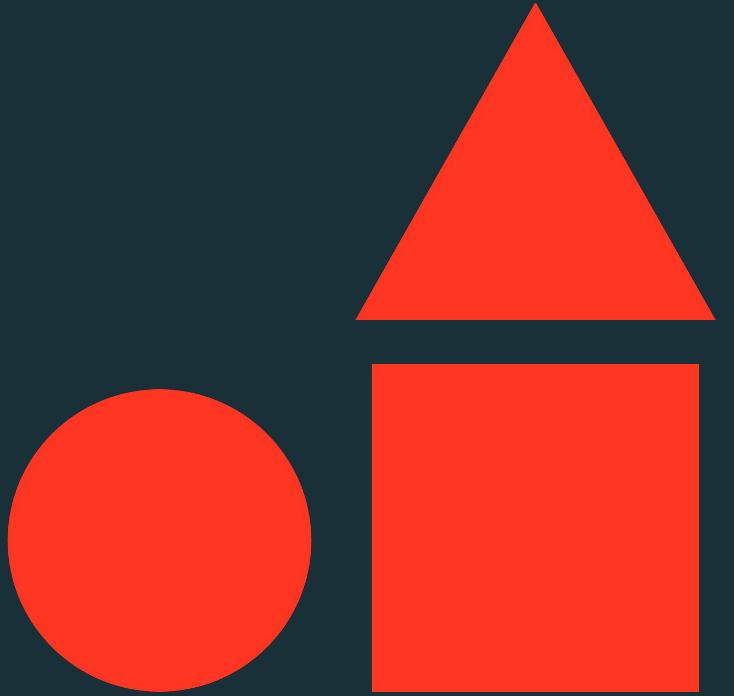
Agents that reason over your data across every enterprise system

Support for all existing and future AI models

Build trust with guardrails, evaluation, and monitoring



Demo



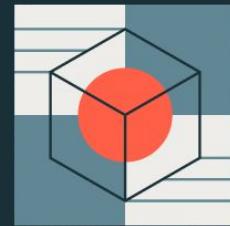
Databricks Apps

Quickly Build Secure Data Applications



Simple

Simplified app development in the technologies developers already know makes it easy to build instead of being bogged down in learning other technologies



Production-Ready

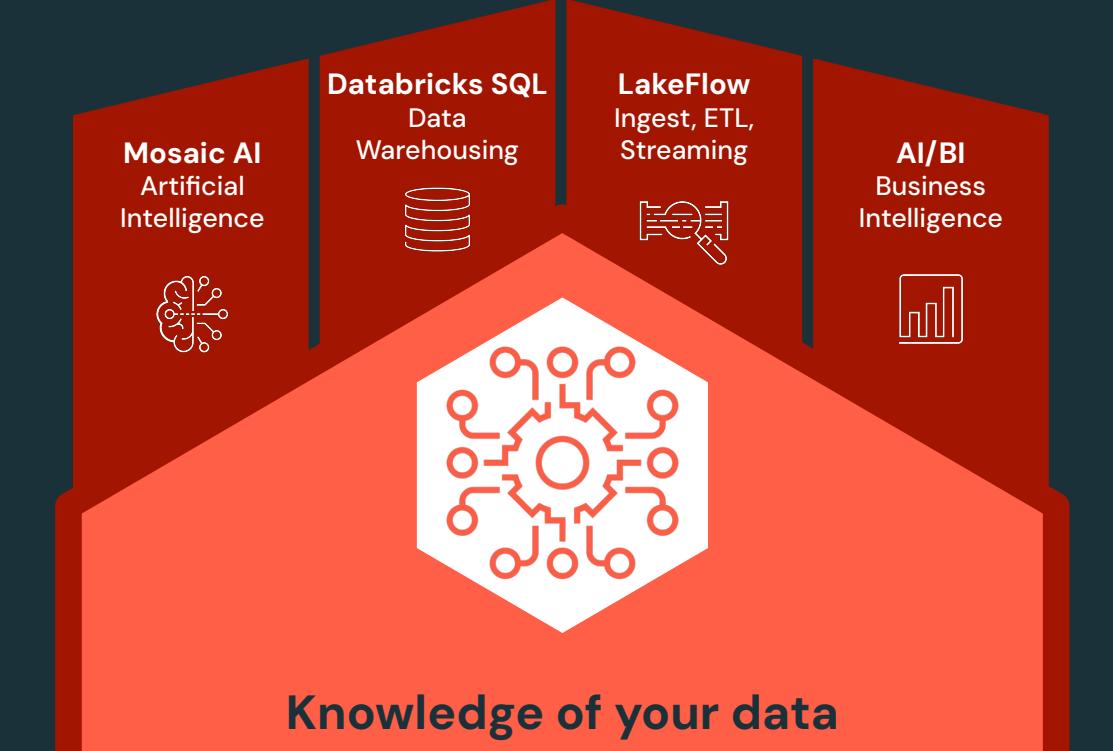
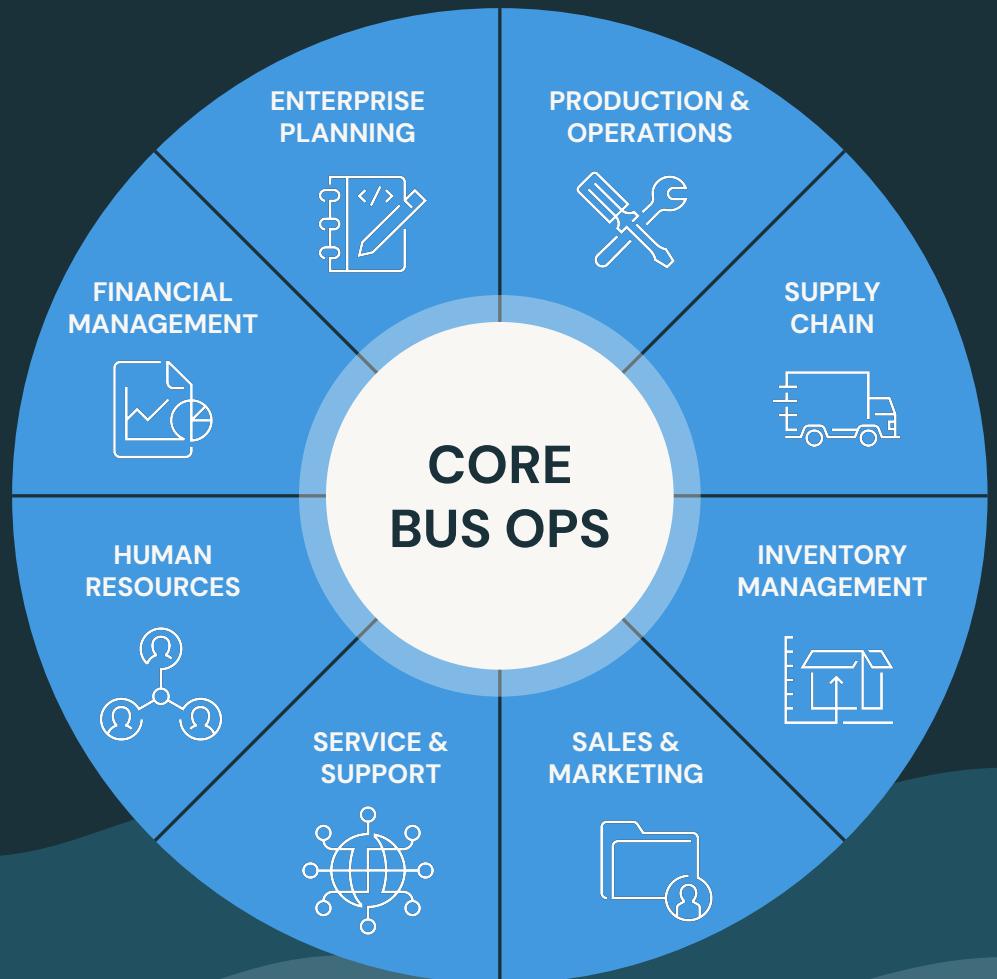
IT has high confidence in the production readiness of the apps developers are building and these can be deployed much faster, and get to market quickly

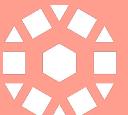


Built-in Governance

Data applications are inherently secure and well-governed, leading to greater data security



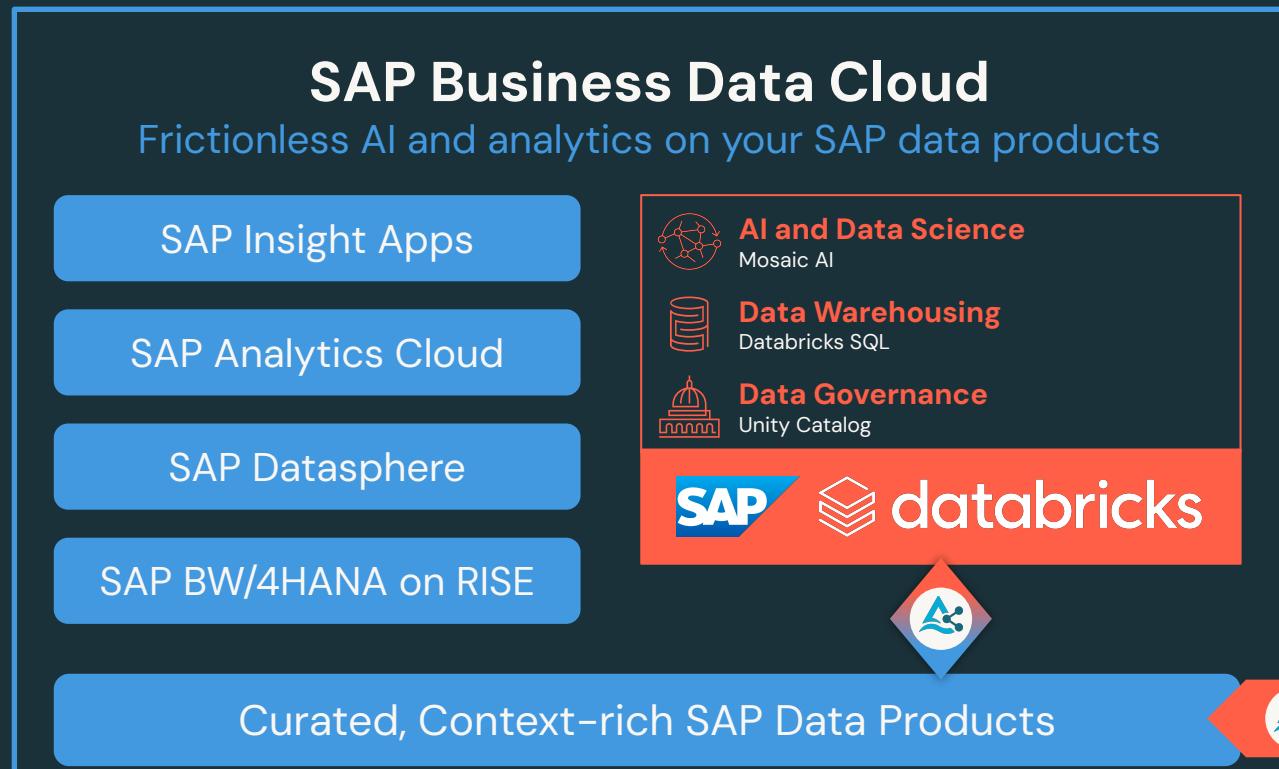


 **Unity Catalog**

Your data stored in an
open, broadly accessible
lakehouse format



Get the best of both worlds





Multi-cloud Flexibility



Data & AI is a journey



how can I try out the data intelligence platform?

⌚ Pro Search

✓ Search for information on how to try out a data intelligence platform

✓ Look for specific platforms that offer trial or demo versions

✓ Find user reviews or experiences of trying out data intelligence platforms

✖ Sources

What is a data intelligence platform | Databricks Blog

databricks · 1

Databricks Data Intelligence Platform

databricks · 2

BigID Data Intelligence Platform: How It Works

bigid · 3



View 3 more

✳ Perplexity

To try out a data intelligence platform, you can follow these general steps based on the platforms mentioned in the search results:

1. Databricks Data Intelligence Platform:

- Databricks offers a free trial of its Data Intelligence Platform. You can sign up for this trial to explore its features and capabilities. The platform is built on a lakehouse architecture, combining data and AI to provide a unified foundation for data management and governance ².
- Visit the Databricks website and look for the "Try for free" option to start your trial ².





databricks

