databricks

Destination Lakehouse

Powering the next era of business efficiency with Data and AI

Nicholas Eayrs
Vice President, Field Engineering

©2023 Databricks Inc. — All rights reserved

9000+
global customers

$1B+
in revenue

$3B
in investment

Inventor and pioneer
of the **data lakehouse**

databricks
The Lakehouse Company

Gartner-recognized Leader
Database Management Systems
Data Science and Machine Learning Platforms

Creator of

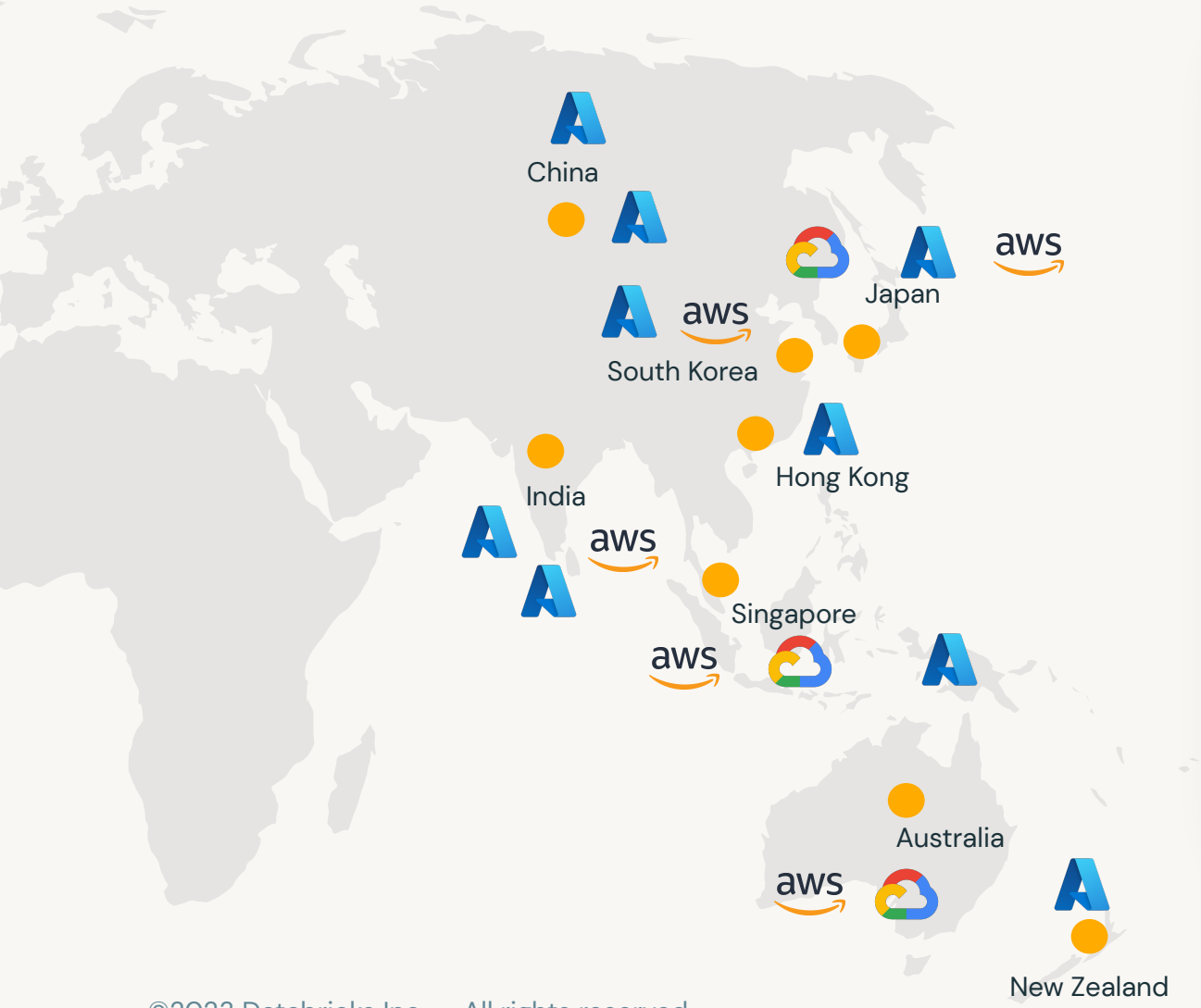
DELTA LAKE

APACHE Spark™

mlflow™

©2023 Databricks Inc. — All rights reserved

Databricks in Asia Pacific & Japan



- 90% YoY business growth
- 600 employees, 80% YoY growth
- 9 Azure regions, 5 AWS regions, 3 GCP regions

ATLASSIAN Grab truedigital

PETRONAS AIA HEALTHIER, LONGER, BETTER LIVES SCB ไทยพาณิชย์

malaysia airlines standard chartered DREAM11

©2023 Databricks Inc. — All rights reserved

78% of CIOs say scaling AI to create business value is their top priority for data strategy by 2025.

They want to become Data Forward

Source: MIT Technology Review Global CIO survey

©2023 Databricks Inc. — All rights reserved

The Data-Forward Enterprise

Data, analytics and AI working in concert

Customer lifecycle

Product development

Workforce productivity

Business operations

Data powering every aspect of the business

The Data-Forward Enterprise

Data and AI working in concert

Becoming data forward

- Grab's in-house Consumer 360 platform enables them to develop more innovative features to continually enhance customer-centric experiences
- Grab ingests and optimizes 1000s of user-generated signals and data sources from various applications, to enhance data integrity and security

Impact

- Gained a unified understanding of customers to **personalize recommendations to millions of users everyday**
- New customer features can be developed faster, **lowering the costs of experimentation**, and **accelerating innovation**

https://www.databricks.com/customers/grab

Data-Forward requires multiple roles

To create quality data, insights and predictions that drive future actions

Data Engineering

Ingest & transform

Data Streaming

Real-time insights

Data Warehousing

SQL & business insights

Data Science

ML to predict outcomes

Data Governance

Management and secure discovery of data assets, with auditing and lineage

→ Data Sharing

©2023 Databricks Inc. — All rights reserved

Data-Forward: from siloed to unified teams

Transforming to product-based outcomes

Specialized team, working in silos
Limits innovation and slows initiatives

Unified teams, shared responsibility
Data agility, acceleration & innovation

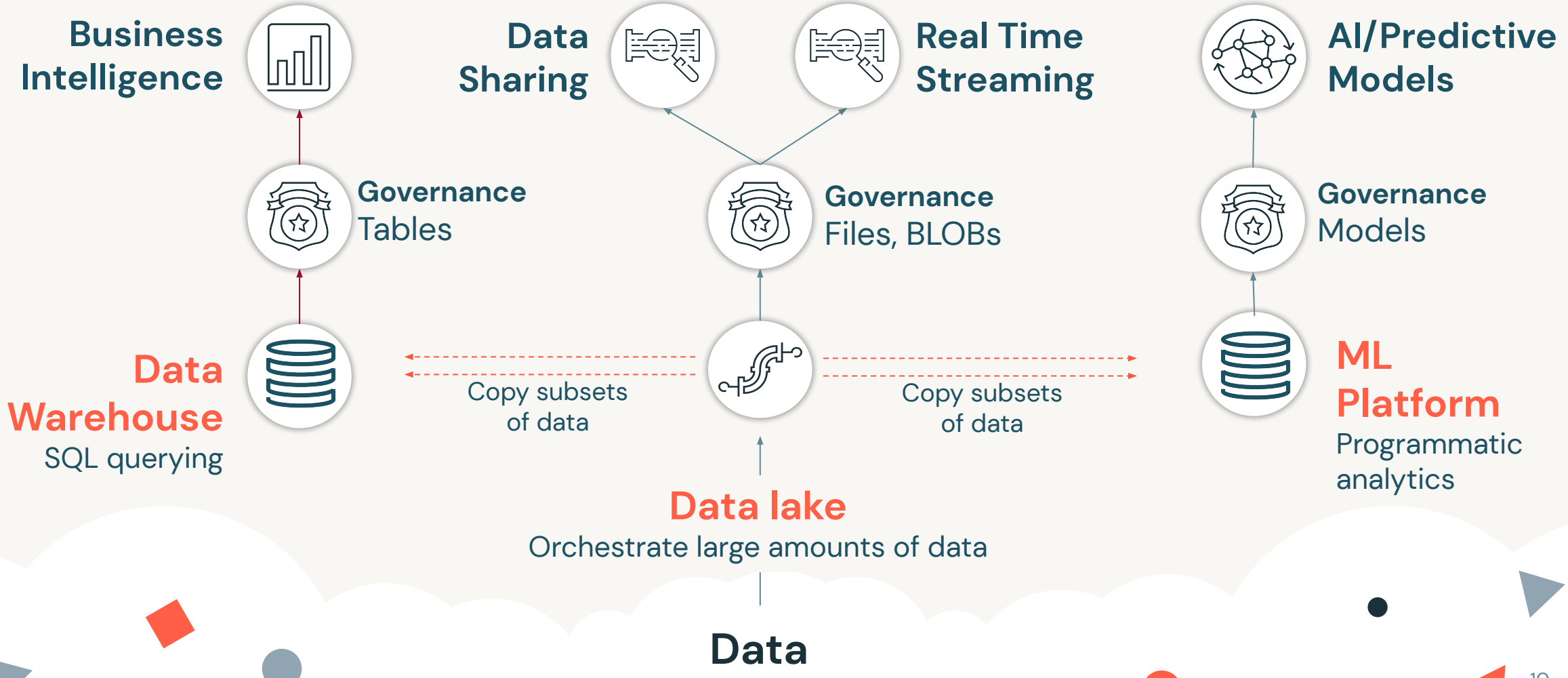
©2023 Databricks Inc. — All rights reserved

Many organizations struggle because their data platforms don’t support their data strategy

©2023 Databricks Inc. — All rights reserved

Historically, multiple platforms

Creates disconnection and duplication costs



Lakehouse paradigm: single platform

One unified platform to support multiple personas

Data Science & ML

Data Streaming

Business Intelligence

Data Engineering

Governance and Security
Files, Blobs and Table ACLs

Data Lakehouse

All Raw data, structured and unstructured

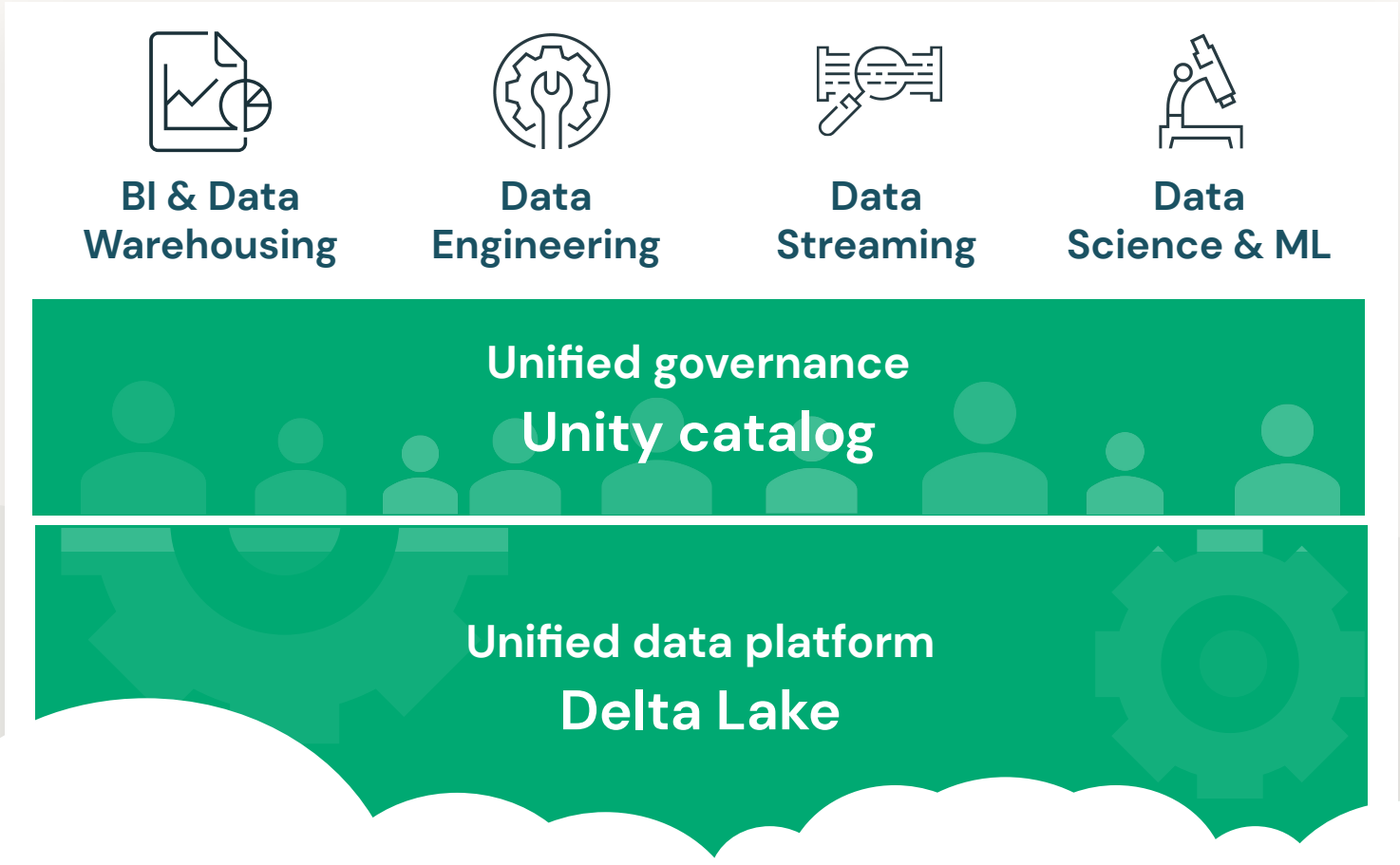
Unified development
Open formats and native support for mainstream languages

Centralised catalog
Fine-grained governance for data and AI

Reliable data
Data integrity and performance with ACID transactions

©2023 Databricks Inc. — All rights reserved

Data-forward with Databricks Lakehouse



Open and collaborative environment to **accelerate** ML, SQL, BI, and Streaming use cases

One security and governance approach for **all data assets** - files, tables, models, dashboards...

A **reliable** data platform to efficiently process all data types with **integrity** at scale

Why Databricks Lakehouse?

Deliver faster innovation at lower cost

1

Economics matter

2

Governance is key

3

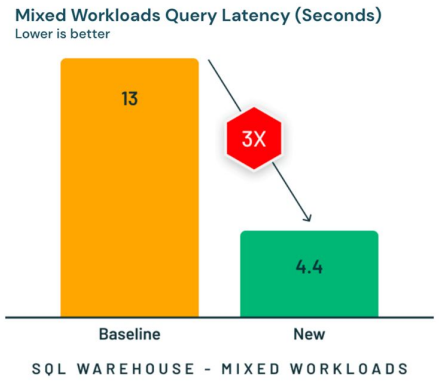
Simplicity is a necessity

©2023 Databricks Inc. — All rights reserved

1 Economics matter

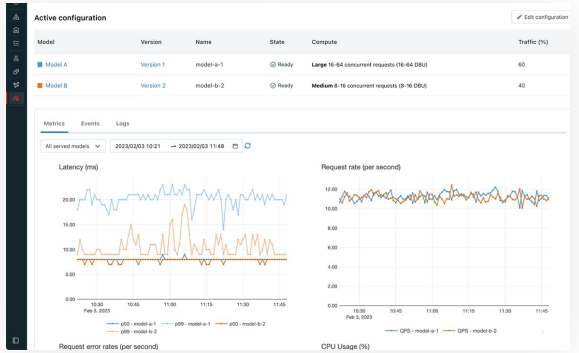
Delivering reduced TCO and efficient growth

Be more responsive



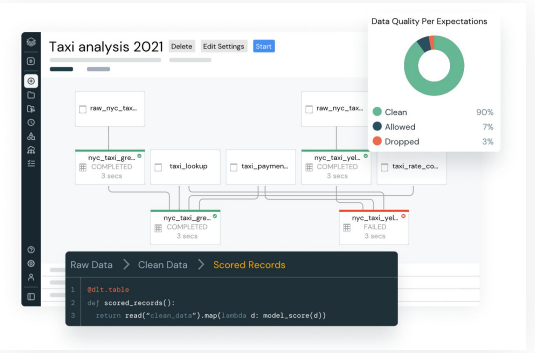
Serverless Warehouse for efficient querying and instant starts

Managed infrastructure



Serverless Model Serving for real-time model inferences

Optimise workloads

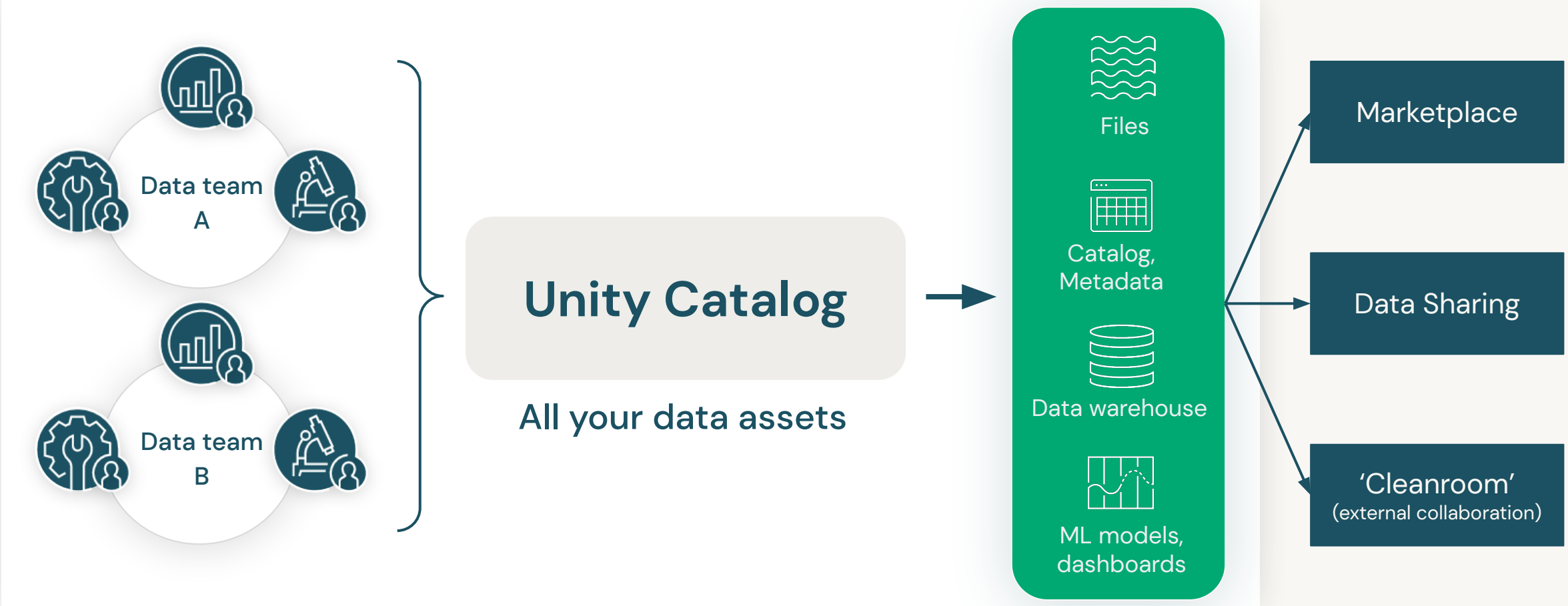


Best TCO for ETL workloads using **Delta Live Tables**

©2023 Databricks Inc. — All rights reserved

2 Governance is key

Unified for responsible and audited data discovery and access

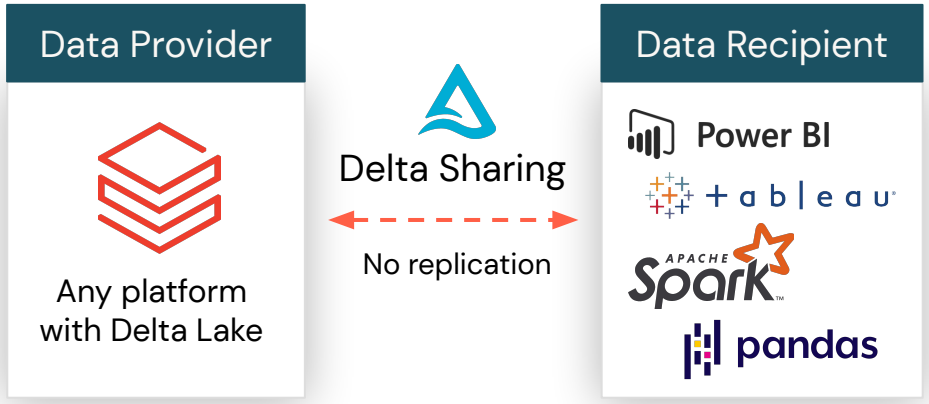


©2023 Databricks Inc. — All rights reserved

3 Simplicity is a necessity

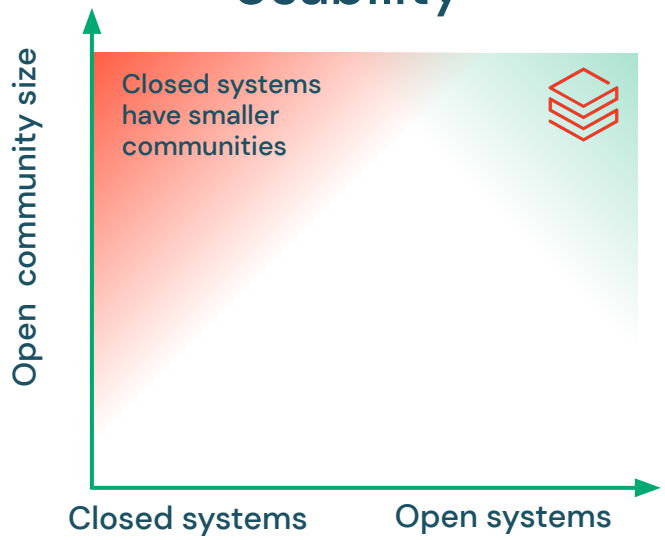
Open systems accelerate innovation and economies of scale

Interoperability



Open formats and protocols to securely consume data using any platform and tools of choice

Usability



Attract and retain talent with native support for the most popular big data technologies

©2023 Databricks Inc. — All rights reserved

Lakehouse delivers
lower costs and faster innovation
to create the
data-forward enterprise

©2023 Databricks Inc. — All rights reserved

Introducing Dolly

A democratised, open source large language model

Dolly: Tiny model. Big attitude. For anyone.

What does Dolly mean for me?

- Substantially **reduces** the **cost & complexity** of utilising generative language AI
- **20x smaller than leading LLMs**
- **Can run** on a single machine completely **within your organisation**
- **Can be fine tuned to have significant depth** on your use case

©2023 Databricks Inc. — All rights reserved

Dolly

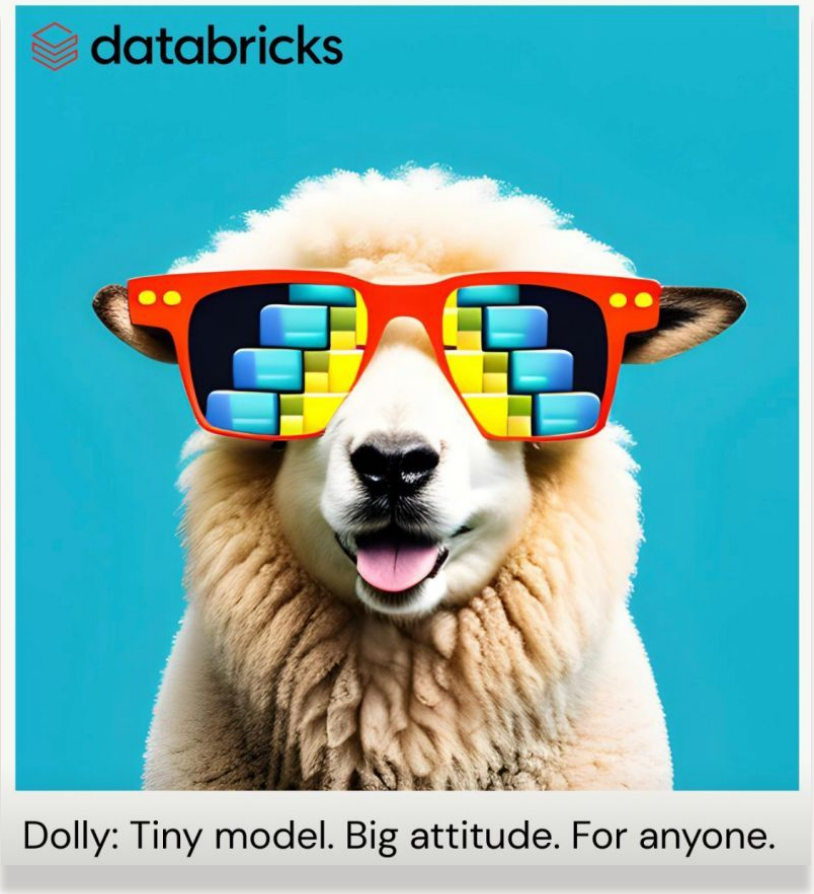
Live Demonstration

Pivot to test-dolly here

©2023 Databricks Inc. — All rights reserved

Build Your Own Large Language Model Like Dolly

Discover how to fine-tune and deploy your custom LLM



JOIN US ON

26 APRIL 2023, 12 PM SGT

©2023 Databricks Inc. — All rights reserved

Try & learn about Databricks Lakehouse

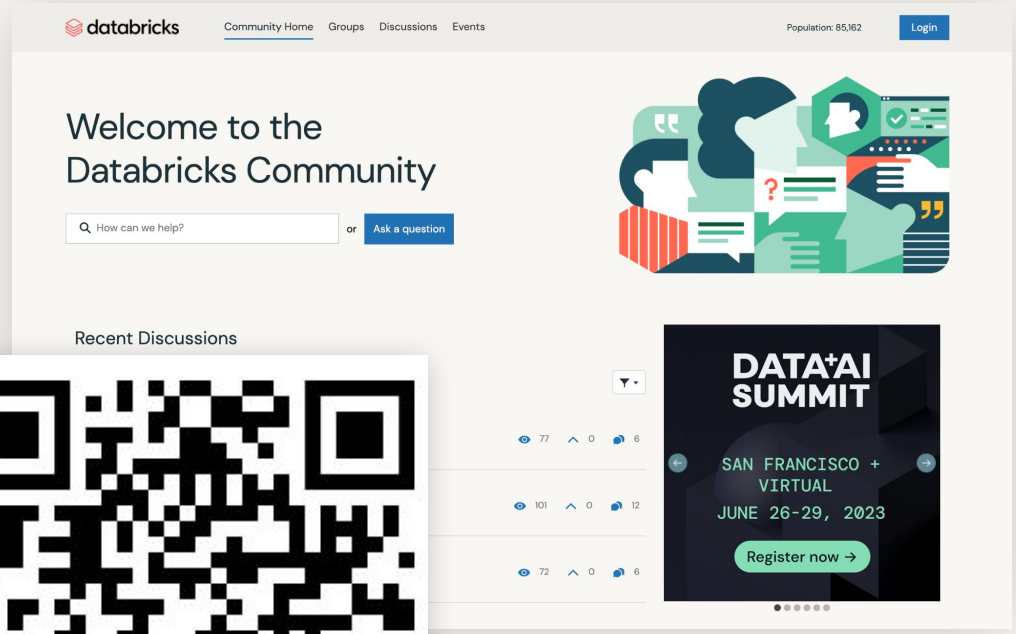
Databricks Training

Free training | **Live** Instructors | **Every** day

- Get started with **Data Analysis**
- Get started with **Data Engineering**
- Get started with **Machine Learning**
- Get started with **Platform Admin.**
- Databricks for **Business Leaders**

customer-academy.databricks.com

Databricks Community



bit.ly/databricks-community-sg

©2023 Databricks Inc. — All rights reserved

Discover how lakehouse brings together all your data

The world's largest data, analytics and AI conference returns live, to San Francisco and virtually. Four days packed with keynotes by industry visionaries, technical sessions, hands-on training and networking opportunities.

- Hear from data and AI thought leaders about latest trends and innovations, including Apache Spark, Delta Lake, MLflow, TensorFlow, dbt and many others.
- See how others are applying the data lakehouse paradigm to unify data, analytics, and AI on one platform
- Catch up on the latest Databricks developments, including SQL Analytics, Unity Catalog, and Delta Lake

BY databricks

©2023 Databricks Inc. — All rights reserved

DATA+AI SUMMIT

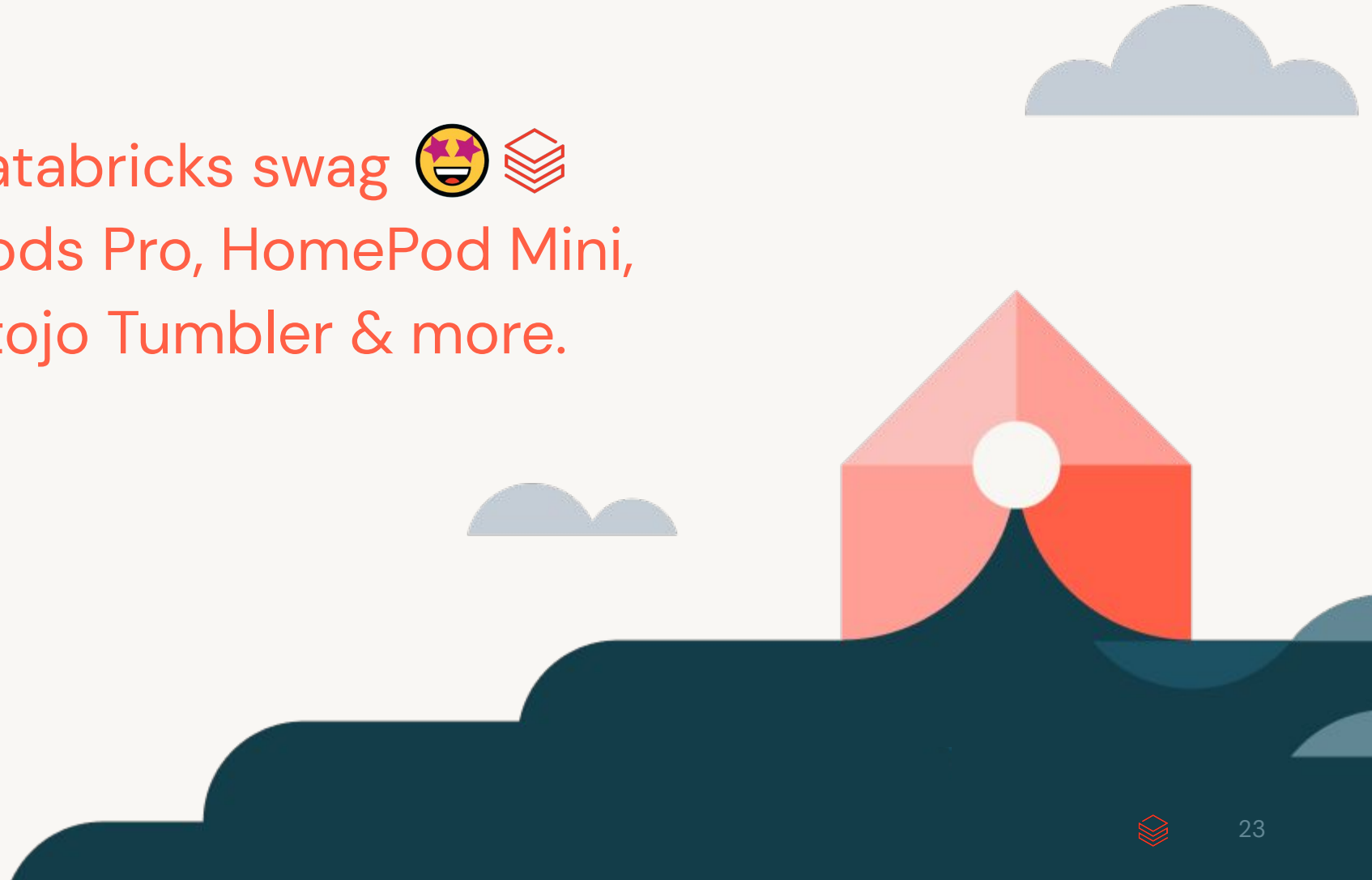
SAN FRANCISCO + VIRTUAL
JUNE 26-29, 2023

Register Now

Thank you

Grab a sure-win Databricks swag

Prizes include AirPods Pro, HomePod Mini,
Balmain Pen Set, Stojo Tumbler & more.



©2023 Databricks Inc. — All rights reserved

©2023 Databricks Inc. — All rights reserved