DEEPER BI INTEGRATIONS WITH DBSQL:

BI Integrations & Partner Ecosystem

Can Efeoglu – Sr. Staff Product Manager
Jade Wang – Staff Software Engineer

June 11, 2024

# Product safe harbor statement

This information is provided to outline Databricks' general product direction and is for **informational purposes only**. Customers who purchase Databricks services should make their purchase decisions relying solely upon services, features, and functions that are currently available. Unreleased features or functionality described in forward-looking statements are subject to change at Databricks discretion and may not be delivered as planned or at all
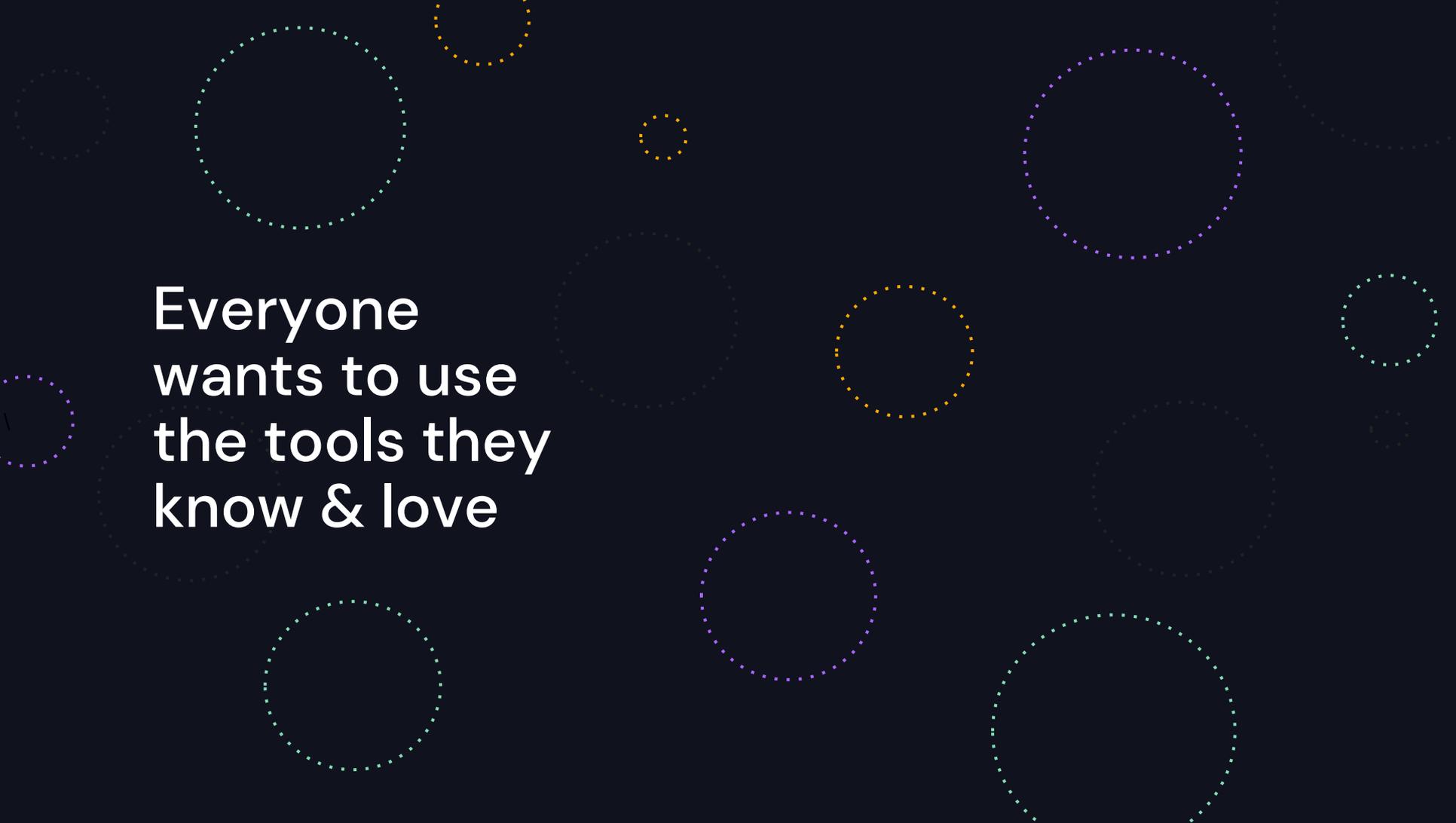
Everyone wants to use the tools they know & love

But the ecosystem is wide and ever changing
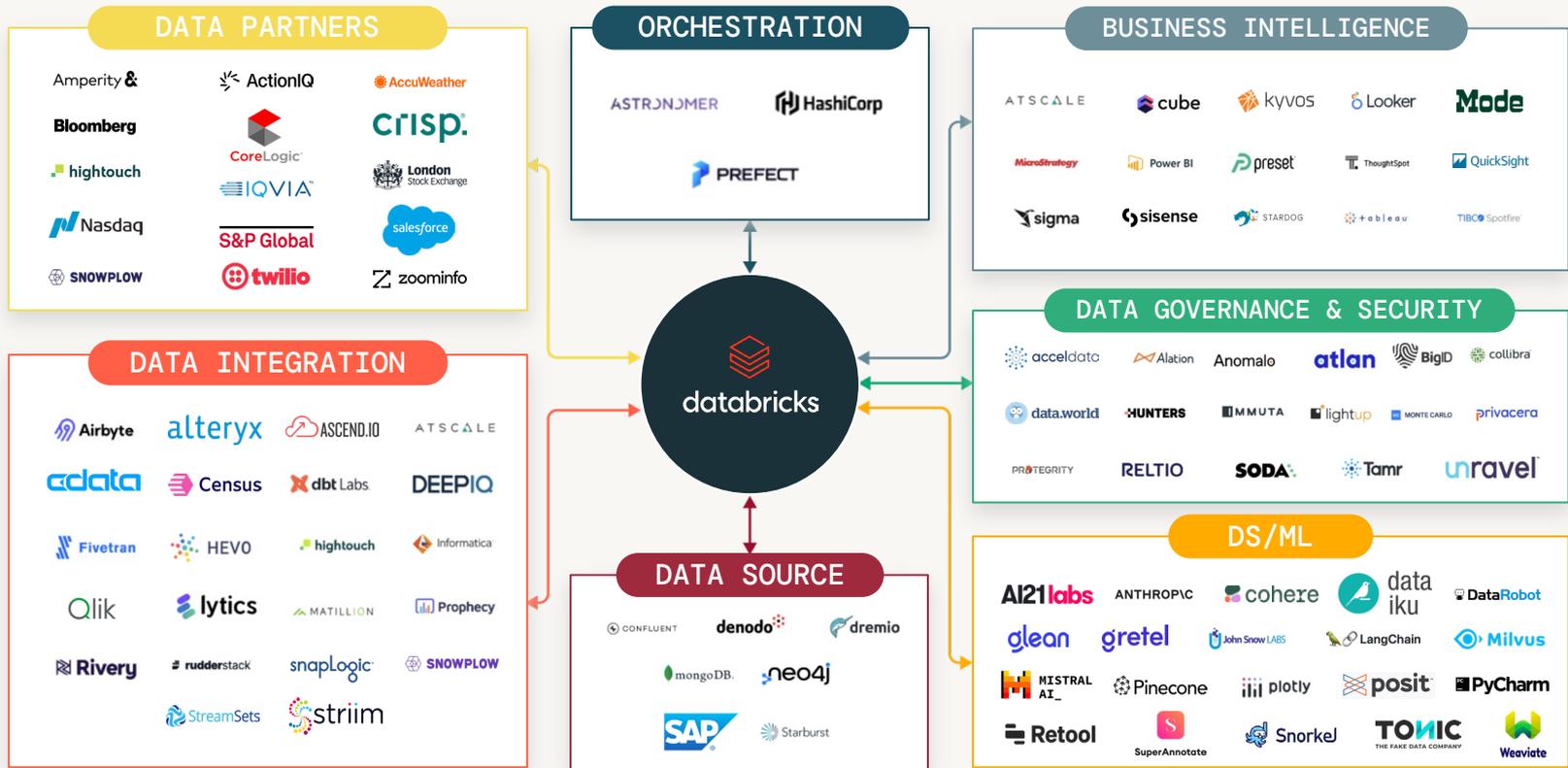
Navigating breadth AND depth

Deeper Integrations: BI and More
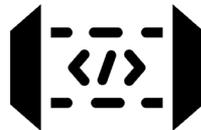
Open APIs & Drivers

Robust Platform Capabilities

Deeper Integrations: Business Intelligence & more

BEFORE:
From Databricks Gold Layer to BI Tool
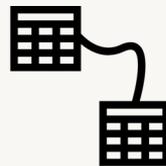
	CREATE GOLD LAYER	**CONNECT FROM BI TOOL**	**MODEL & PUBLISH**	**AUTHOR REPORT & CONSUME**	**REPEAT!**
Task	**Transform** data from **bronze & silver layer** into **modeled gold tables**	**Enter credentials** & **connect** to Databricks from BI Tool	Select tables, **model data & its relationships** (again...) and **publish** to Online platform	Using the published datasets: **author reports & consume data**	**Repeat** as dataset and modeling relationships **evolve**.
Where?		BI Tool – **Desktop**	BI Tool – **Desktop**	BI Tool – **Online**	

How could this be better?

- **Direct integration:** integrated into end of data pipelines
- **No desktop hops:** between desktop & online BI platforms
- **No manual connection management:** Remove need to manage connections & credentials manually
- **Do entity modeling once:** No need to re-do table/entity relationships across platforms.
- **Automatically up-to date:** Push changes to underlying tables & their relationships automatically.

AFTER:
From Databricks Gold Layer to BI Tool

	CREATE GOLD LAYER	**PUBLISH (KEEP IN SYNC!)**	**AUTHOR REPORT & CONSUME**
Task	**Transform** data from **bronze & silver layer** into **modeled gold tables**	Publish dataset metadata, including entity relationships (PK/FK). Keep synced as desired.	Using the published datasets: **author reports & consume data**
Where?			BI Tool – **Online**

Announcing Preview:
Power BI & Tableau Integrations

Seamless catalog integration & data model sync

Power BI Integration *(Preview now)*

Publish UC datasets from Databricks UI, without PBI Desktop to Power BI Online.

Sync entire schemas including table relationships (PK/FK) to save time.
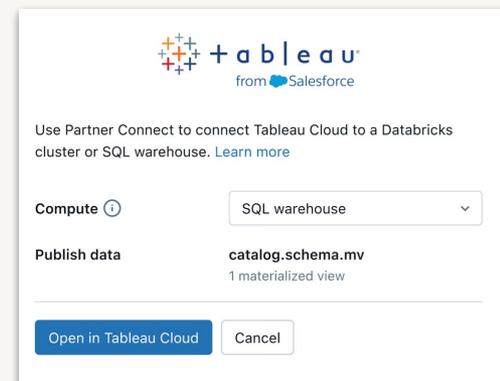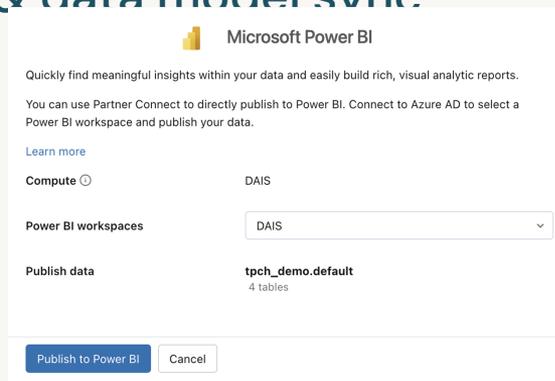
Tableau Integration *(Q3 Preview)*

Easily explore Unity Catalog datasets in Tableau Online with a single click from Data Explorer.

Microsoft Power BI

Quickly find meaningful insights within your data and easily build rich, visual analytic reports.

You can use Partner Connect to directly publish to Power BI. Connect to Azure AD to select a Power BI workspace and publish your data.

Learn more

Compute ⓘ	DAIS
Power BI workspaces	DAIS
Publish data	tpch_demo.default
	4 tables

Publish to Power BI   Cancel

tableau from Salesforce

Use Partner Connect to connect Tableau Cloud to a Databricks cluster or SQL warehouse. Learn more

Compute ⓘ	SQL warehouse
Publish data	catalog.schema.mv
	1 materialized view

Open in Tableau Cloud   Cancel

Publish to BI Tool: Looking Ahead – Roadmap

- **API based publishing & management**
  - Publish new dataset(s) directly from API
  - List & manage existing datasets
  - Schedule automatic publishing
- **Lineage support**
- **Enhanced admin capabilities**
  - Admin view for managing user published datasets
  - Publishing of arbitrary tables/views across catalogs & schemas
  - Control which tools are available to publish to per workspace
- **Expand into more BI platforms**
  - Support for other 3rd party platforms (given partner has API support)

Delta Sharing Tableau Connector

Easily share data to external partners & customers without Databricks accounts

New data connector expands Delta Sharing open ecosystem for business intelligence use cases

Easily share live data with your external partners, who can consume it directly in Tableau

As a data provider, expand the reach of your data by meeting end-users where they are in Tableau. As a data consumer, easily tap into the rich Databricks ecosystem and Databricks Marketplace

Leverage built-in data governance and security with Unity Catalog

"Together, we're delivering a powerful data-sharing capability to help customers seamlessly share data between Tableau and Databricks. Real-time sharing quickly refreshes analyses, helping teams collaborate and make data-driven decisions based on the most up-to-date information."

— Kevin Glover,
Director, Product Management, Tableau

More BI Integrations:

New Sigma, Looker Studio and updated Looker Connector

Looker

Revamped connector with improved pushdowns & incremental PDT support.

Looker Studio

NEW: Databricks connection support for Looker Studio.

sigma

NEW: Databricks connection support for Sigma now Generally Available

Familiar spreadsheet like experience right on the Lakehouse – including Input Table support for a more application-like experience.

Deeper dbt integration

Enhanced dbt workflows with incremental ingest & materializations

- **Incremental ingest & smart materializations:** Enhanced **MV/ST** support with better configurability & stability.

- **Flexible compute management:** Per model compute selection for SQL vs. Python models.

- **Lakehouse Federation Integration:** Directly query any data source supported by Lakehouse Federation from your dbt pipelines.
- **Support for core platform features:** liquid clustering, tags and more.

```
CREATE STREAMING TABLE report
AS SELECT SUM(profit)
FROM cloud_files(prod.sales)
```
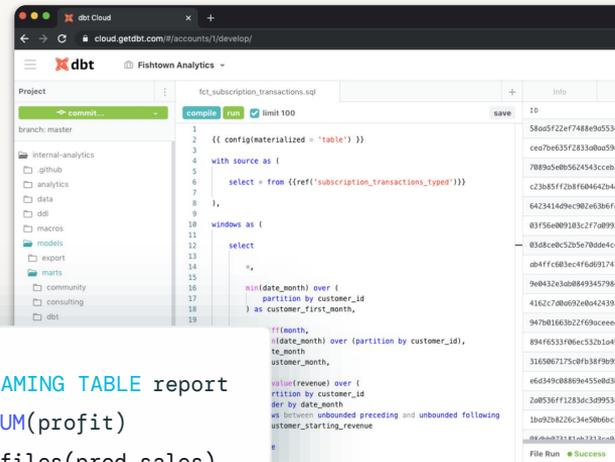
Open APIs and Drivers

# Open APIs & Drivers – beyond tool specific connectivity

Connect, query & build on the Lakehouse from any stack

Connect, query & build
Interact with Databricks SQL programmatically by automating tasks, integrating with other systems and build custom applications.

Any stack
Pick from a selection of Python, GO, Node.js, ODBC (now supporting ARM systems), JDBC, REST, dbx CLI & ADBC (.NET)*.
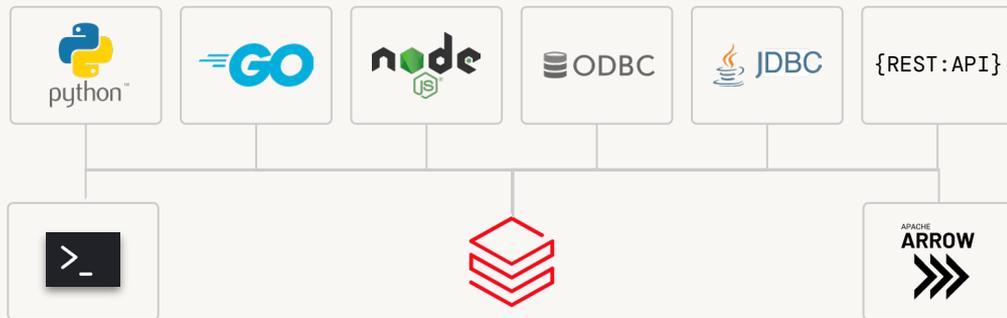
SQL Statement Execution API

Now Generally Available

Empower users to execute SQL statements over HTTP, providing seamless integration with a wide range of applications and technologies.

New features available in the GA release:

- <u>Parameterized statements</u> for safe and dynamic SQL queries
- <u>*Result retention*</u> for refetching statement results
- <u>Multiple result formats</u> for enhanced interoperability
- <u>Byte and row limits</u> to manage output sizes
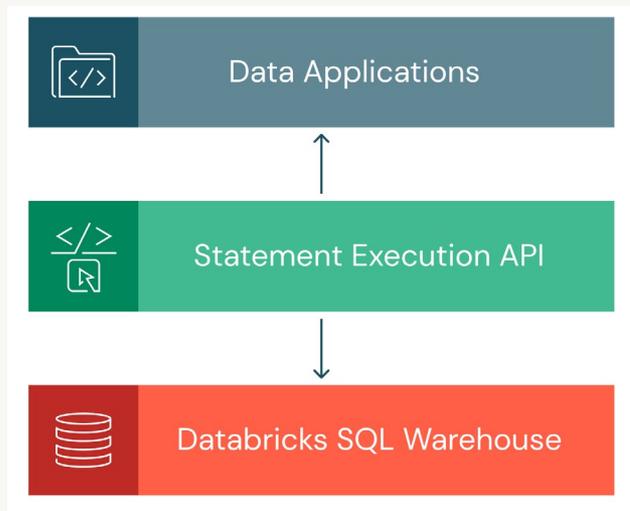
Python, GO & Node.js

Now Generally Available

Providing robust and reliable connectivity options for a wide range of development environments.

- <u>Parameterized statements</u> for safe and dynamic SQL queries
- <u>Smart connection management and retries</u> ensuring more reliable connectivity and resilience
- <u>Cloud Fetch support</u> ensuring high throughput data transfers
- <u>UC Volumes – Ingest:</u> Ingest and upload local files into UC staging locations to flexibly manage file staging from any environment

What else?

Preview: SQLAlchemy v2 and ADBC (.NET)

SQLAlchemy 2.0 Dialect (Preview)

We've upgraded our dialect to support v2

- Improved compatibility and performance
- Ability to use <u>v1 or v2</u> with Python Connector 3.0+

Arrow Database Connectivity (Coming Soon)

Partnering with Microsoft to offer ADBC driver (.NET) based connectivity within Power BI and larger Microsoft ecosystem:

- Improved performance: better DAX operator pushdowns & reliability.
- Faster model refreshes on import mode.

Robust Platform Capabilities

Partner Connect

Evolving with the ecosystem

Discover new tools/partners
10+ new partners

Quick trial experience

Time to production

1. Select the partner
2. Initiate partner **trial sign-up** from Partner Connect
3. Partner Connect automatically **connects the partner product** to your Databricks workspace
4. Complete partner trial sign-up, select your data source connectors
5. Ingest data into Delta Lake

Partner Solution

PARTNER

Partner connectors
(applications, databases, files/storage)
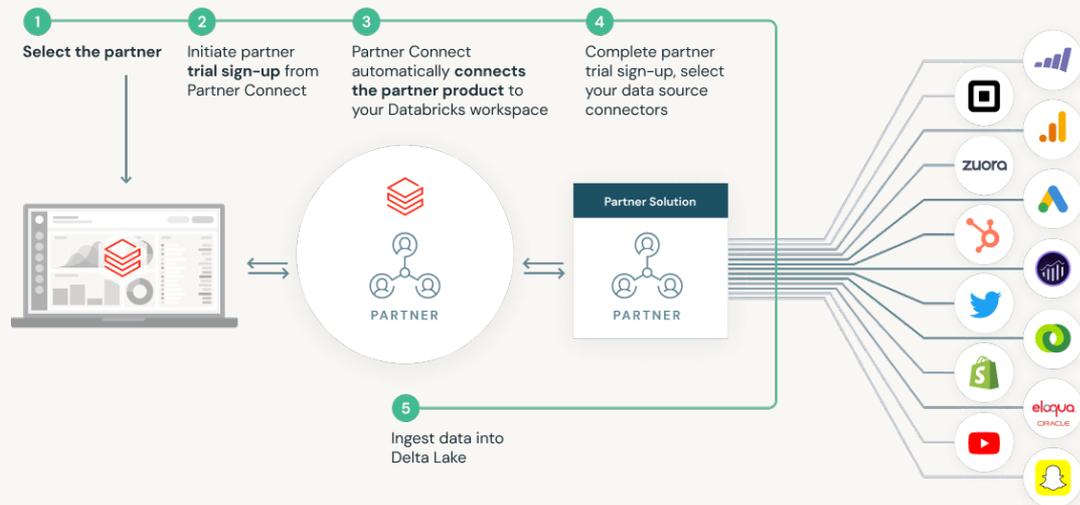
Security: OAuth by default

OAuth on all drivers & APIs

OAuth on All Drivers/APIs
OAuth is now available across all Databricks APIs and drivers.

OAuth on Partner Connectors
Power BI, Tableau, dbt, Alteryx, Thoughtspot, Dataiku and more.

Developer Experience
Administration UI for managing OAuth integrations, providing a more intuitive and user-friendly experience for developers.

DEMO

Wrap Up

Deeper Integrations: BI and More
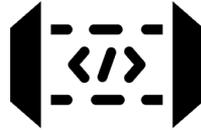
Open APIs & Drivers

Robust Platform Capabilities

Learn more at the summit!

Databricks Events App

Tells us what you think

- We kindly request your valuable feedback on this session.

- Please take a moment to rate and share your thoughts about it.

- You can conveniently provide your feedback and rating through the **Mobile App**.

What to do next?

- Discover more related sessions in the mobile app!

- Visit the Demo Booth: Experience innovation firsthand!

- More Activities: Engage and connect further at the Databricks Zone!

Get trained and certified

- Visit the Learning Hub Experience at Moscone West, 2nd Floor!

- Take complimentary certification at the event; come by the Certified Lounge

- Visit our Databricks Learning website for more training, courses and workshops! databricks.com/learn