KONVERGA

Databricks Asset Bundles

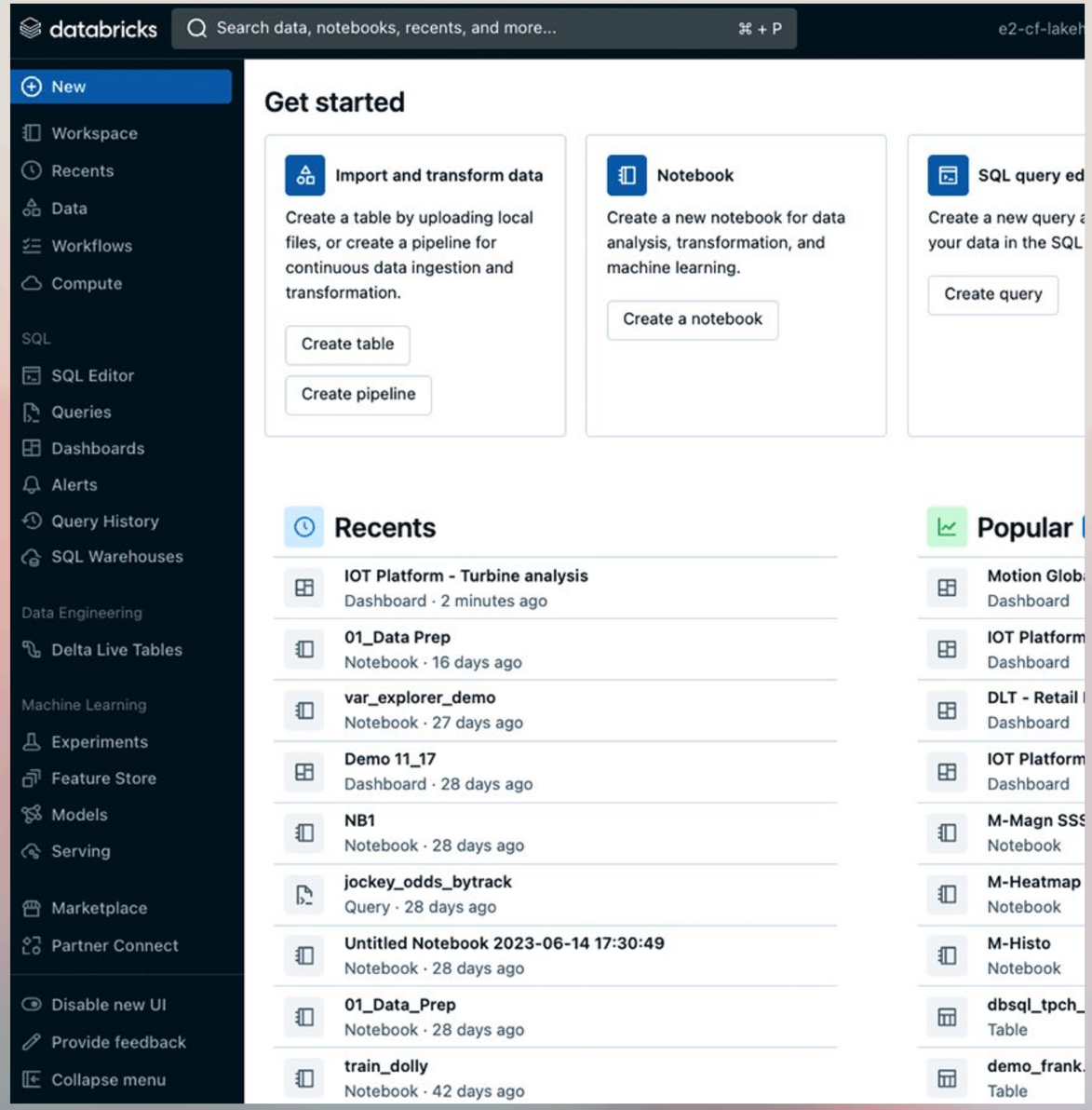
Lunch & Learn

Johan Hostens
30 January 2026

databricks

Databricks Intro

- Why Databricks?
- Azure resources
- Databricks workspace
- (Demo)

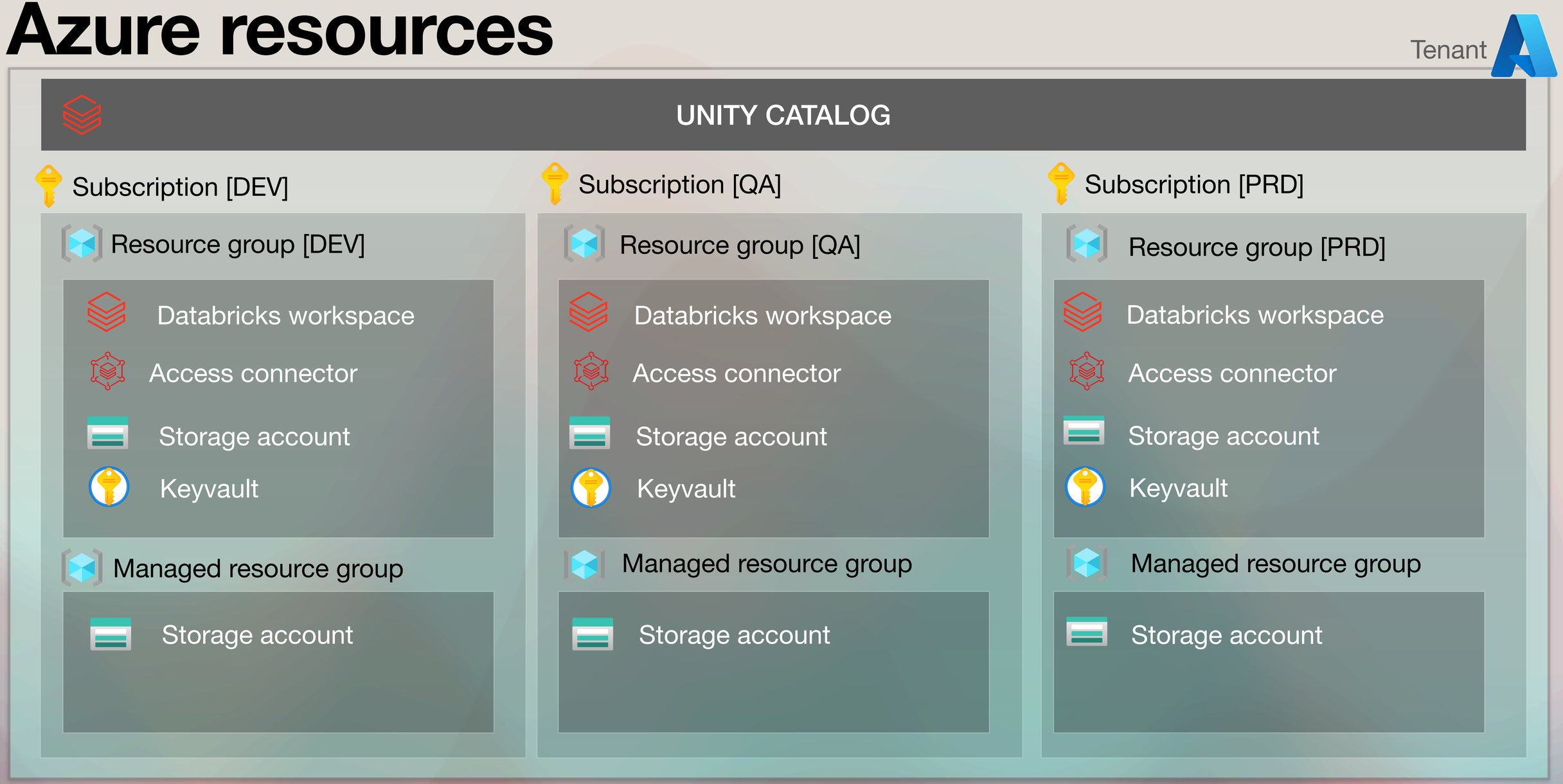


Why Databricks?

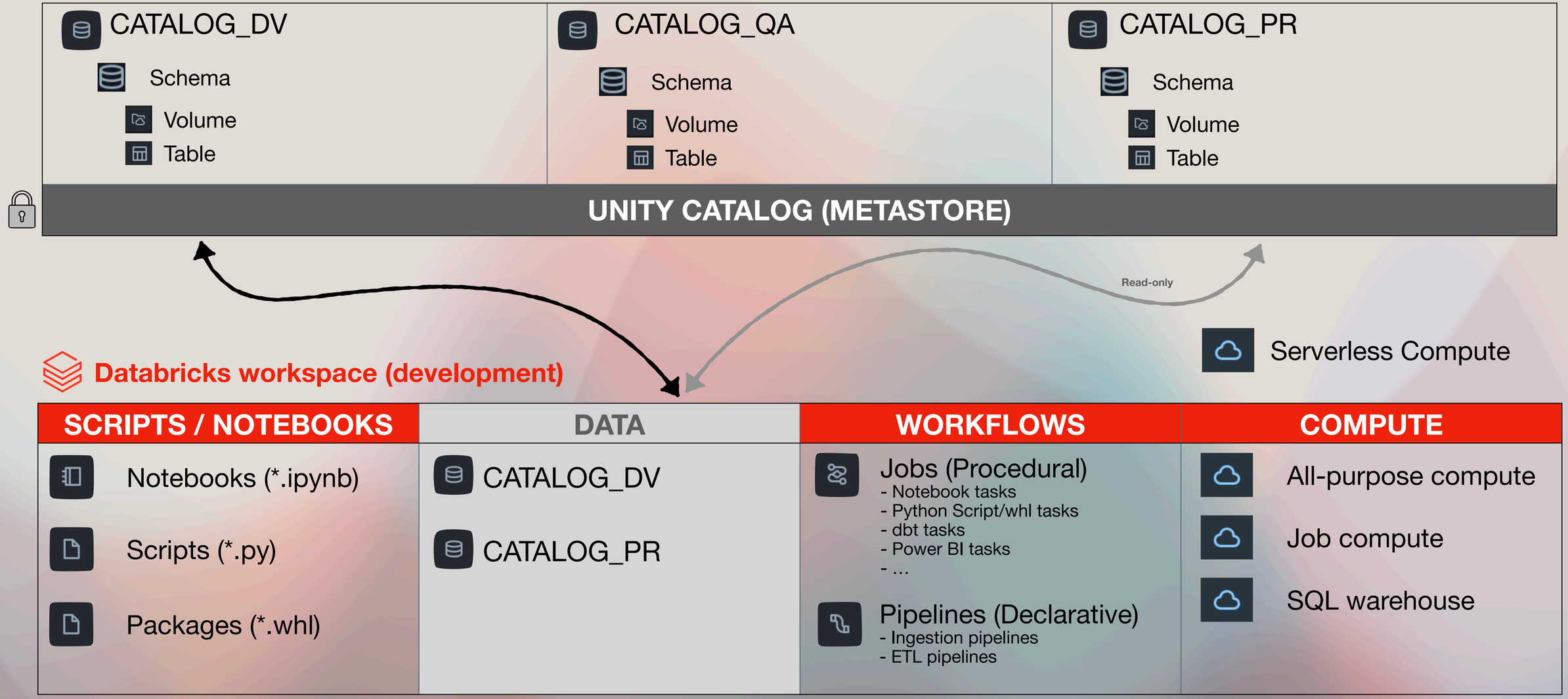
Stable, scaleable and secure platform	Try it yourself 😉
Mature	Databricks GA: 2015 (same year as Power BI) Azure Databricks GA: 2018
Well-thought-out architecture	Object Model
Innovative	Spark Delta-format …
Clear pricing model	Pay-as-you-go DBU (Databricks Units)
Built-in governance and easy auditing	Unity Catalog
Access control **Ownership** **Version control** **Promotion** [dev - qa - prod]	Personal folders for notebooks / access control workflows Service principals for workflows GIT folders **Databricks Asset Bundles**
Good integration with Azure and other cloud platforms	Entra, Storage Accounts, SQL server,.. Lakeflow connect - managed connectors
Multi-cloud	Azure, AWS, GCP

Azure resources

Tenant



Databricks Workspace



[1] Databricks introduction

Jobs

- Activity types
- Trigger types
- Compute
- Notifications

Catalogs

- Descriptions
- Lineage
- Audit tables

Databricks Asset Bundles

- What is a DAB?
- Test project summary
- Visual Studio Code configuration
- (Demo)

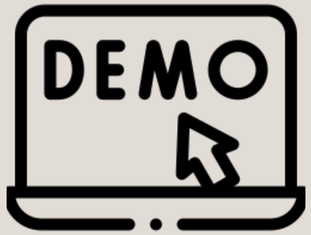
What is a DAB?

Databricks workspace (development)

SCRIPTS / NOTEBOOKS		WORKFLOWS	COMPUTE
Notebooks (*.ipynb) Scripts (*.py) Packages (*.whl)		Jobs (Procedural) - Notebook tasks - Python Script/whl tasks - dbt tasks - Power BI tasks - ... Pipelines (Declarative) - Ingestion pipelines - ETL pipelines	Job compute configuration

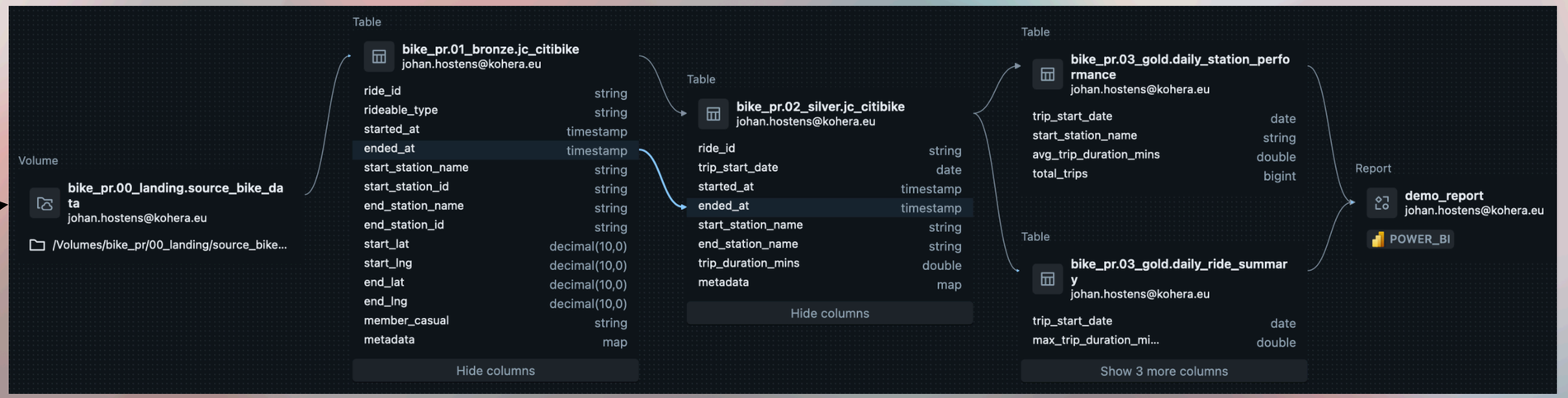
Targets: DEV QA PROD

Test project summary



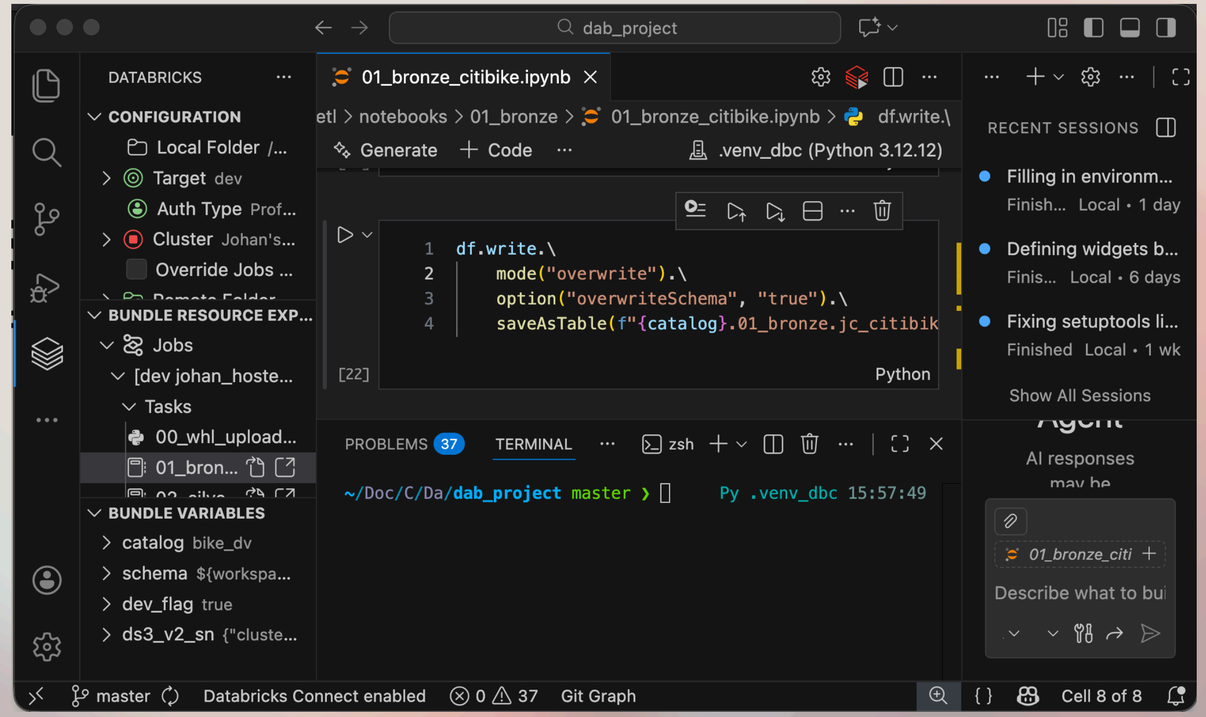
Databricks Asset Bundle		
Wheel file		
(1) Notebooks + Job	**(2) Python files + Job**	**(3) Lakeflow Spark Declarative pipelines (Delta Live Tables)**
Notebooks (ipynb)	Python files (py)	Notebooks (ipynb)
Job (yml)	Job (yml)	Pipeline (yml)

JC-202503-citibike-tripdata.csv



Visual Studio Code

Databricks Plugin



git

Auto-documentation: Sphinx

Testing framework: pytest

Combined with

Formatter: Black

Linter: Ruff

AI-assisted coding: GitHub Copilot

Type checker: mypy

—

Azure resources

dbt projects

SQL databases

...

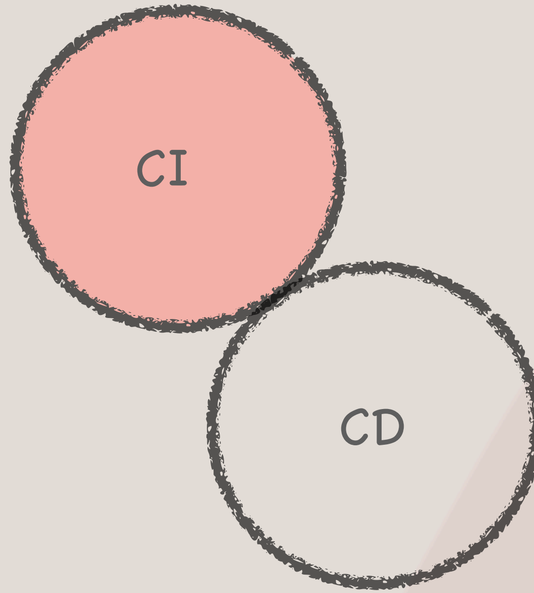
[2] Databricks in VS Code

- VS Code Databricks plugin
- Databricks CLI
- DAB project
- Deployed resources in development / production
- Automated documentation with Sphinx
- Pytest results

Github: CI/CD

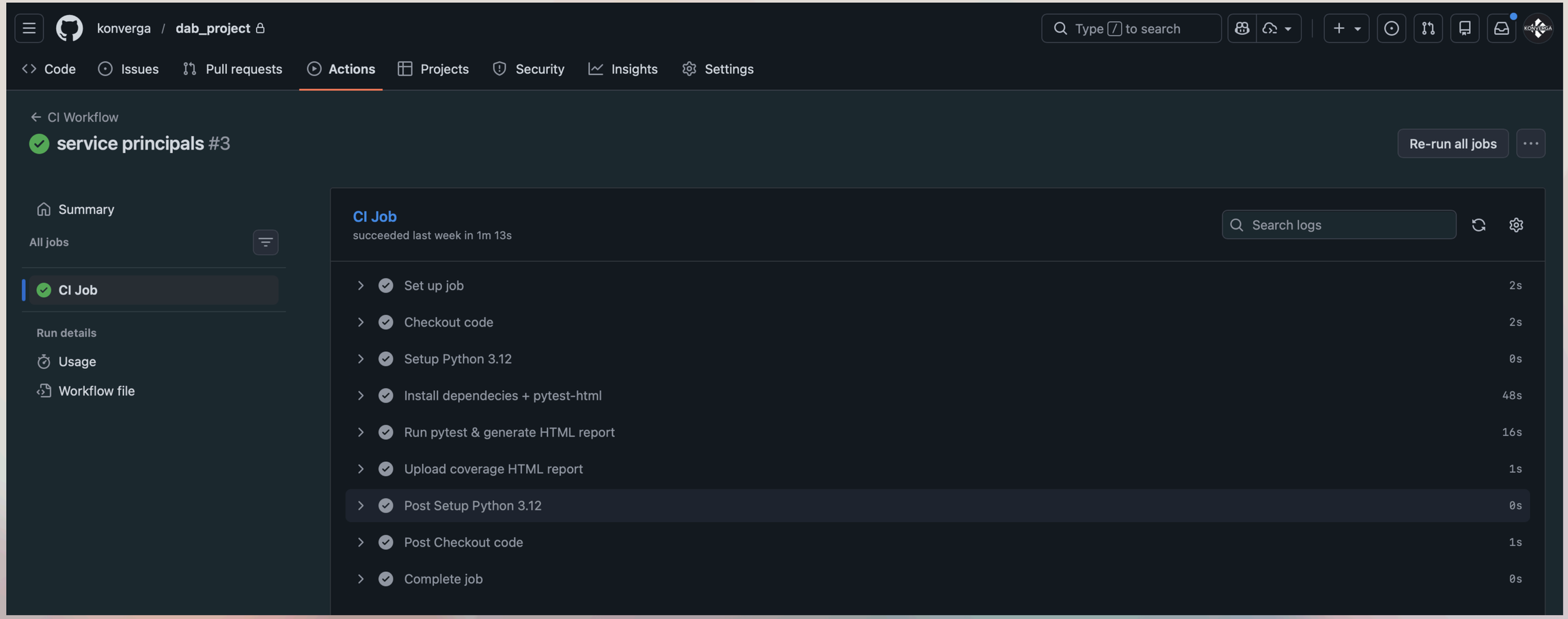
- Continuous Integration (CI)
- Continuous Deployment (CD)

Continuous Integration

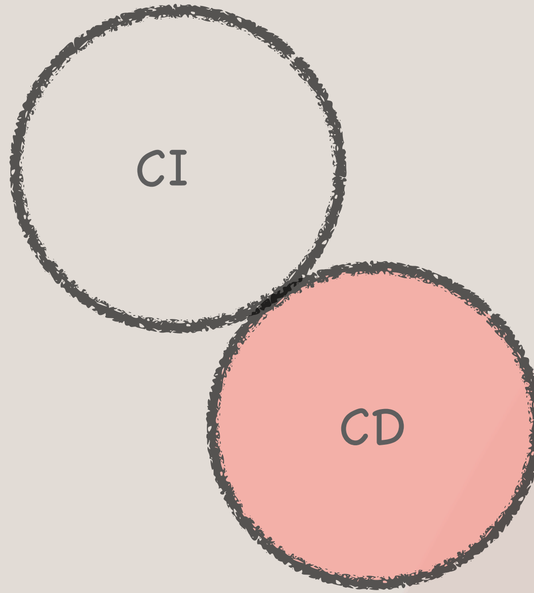


1. Clone Repository
2. Create local feature branch
3. Make changes
4. Local unit testing
5. Commit changes
6. Push feature branch to GitHub
7. Open Pull-Request
   CI workflow automatically triggered

Github: CI flow screenshot



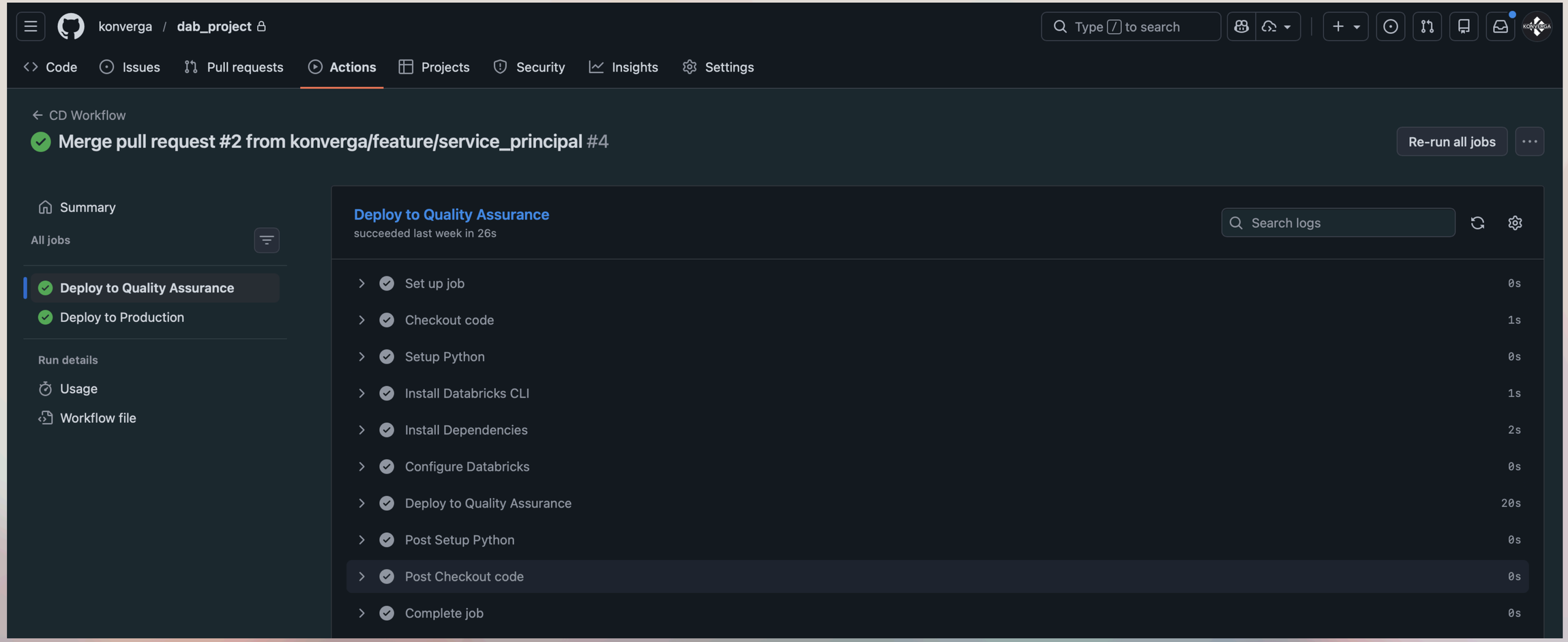
Continuous Deployment



Pull Request approved:

1. CD workflow is automatically triggered
2. DAB is deployed to QA Databricks workspace
3. Pause workflow — Integration tests & Data Quality checks
4. Manual approval
5. DAB is deployed to PRD Databricks workspace

Github: CD flow screenshot



Fun Events in the Spotlight

- **6:00 pm: Welcome with food & drinks**
  Doors open, grab a drink and a bite as we get started!
- **7:00 pm - 7:40 pm: From data to demand: Developing your own forecasting models & Custom AI Agents with Databricks**
  Turning client data into reliable demand forecasts is one thing, operationalizing it safely at scale is the real challenge.
  This meetup talk shows how we use Databricks as the backbone for demand forecasting, from data preparation to production-grade delivery. MLFlow sits at the core: tracking experiments, comparing runs, and registering the best model for controlled promotion and rollout. We'll demonstrate how the model is embedded in Databricks data pipelines so forecasts are produced automatically and consistently for downstream use. Forecasts naturally trigger follow-up questions, so we'll also introduce our own secure agent that answers client questions grounded in their own data. You'll see how agent-level security and access controls keep every response scoped, compliant, and auditable, without slowing teams down.
  By Bram Verbeke, Analytics Consultant @ Plainsight
- **7:50 pm - 8:30 pm: Engineering the Future: Next-Generation Pipelines with Databricks Lakeflow**
  Databricks Lakeflow is a powerful set of tools for running data engineering workflows on Databricks. It allows data practitioners to build end-to-end, production-ready data pipelines. Yet, having the right tools is only half the job; knowing how to use them effectively is just as important. During this session, we'll dive into the most important components of the Lakeflow toolkit: Lakeflow Connect for seamless data ingestion, Lakeflow Declarative Pipelines for building robust batch and streaming pipelines, and Lakeflow Jobs for headache-free workload automation.
  Attendees will gain practical insights into building production-ready data pipelines. Focusing on proven data engineering best practices, we'll also look ahead to the future of Databricks Lakeflow and how it can help democratize data across teams
  By Oleksandra Bovkun, Sr. Developer Advocate @ Databricks
- **8:30 pm - 10:30 pm: Networking & Drinks**
  End the night with casual drinks and great conversations!

dataMinds {saturdays}

Saturday, February 21, 2026

Tjomme Vergauwen

Data intake and processing made easy with Databricks Lakeflow Declarative Pipelines

Business users often wonder why it takes so long to ingest their data into the data lake. Maybe their development team is not using the right tools to get the job done.
In this session I will explain how Databricks AutoLoader and Lakeflow Declarative Pipelines can help to get the job done. AutoLoader is a good fit for quick data ingestion. Lakeflow Declarative Pipelines (formally known as Delta Live Tables) is ideal to process your data (both snapshot or streaming) into your data lake and track the history of changes (Slowly Changing Dimensions type 2). Depending on your data flow, different scenarios can be chosen, which I will explain in the session.
When this process is meta data driven, it becomes even more powerful. It can even be extended with data quality rules.
As this can only be done using SQL or python, some knowledge of these is recommended.
Data architects and data engineers will benefit the most from this session. Business users will learn that there are ways to get the data in faster :-)

ROOM 1 (D0.01A) Sat 2:25 pm - 3:25 pm

Reiner De Smet

The Different Levels of Securing a Self-Service Environment with Unity Catalog in Databricks

In this session, we will explore how to set up and maintain secure self-service environments within Databricks, focusing on key features of Unity Catalog, while keeping in mind how a team can maintain and monitor these environments.

The session starts with an intro on Unity Catalog and how it enables you to securely set up your self-service environment.

After the basics on Security with Unity Catalog we will discuss various levels of secure self-service environments:
- Basic security measures using Unity Catalog to manage and govern data within the same workspace.
- Intermediate strategies involving the creation of new workspaces and providing access through Unity Catalog and Delta Sharing.
- Advanced security setups leveraging the Trusted Research Environment ideology
- Extra Databricks features for securing and monitoring data collaboration within Databricks
Databricks Clean Rooms (currently in preview),
Cluster permissions, and bugets

ROOM 1 (D0.01A) Sat 3:50 pm - 4:50 pm

databricks AI DAYS

Live from Brussels

Thursday, April 2, 2026 | 8:00 AM–1:30 PM

Voco Brussels, Indringingsweg 1, 1800 Vilvoorde

Join us at Databricks AI Days, a half-day, in-person event designed to give practitioners the foundational skills and hands-on experience to build and operate Analytics, AI, and Apps. We provide a clear, hands-on introduction to data and AI, with tailored content for those early in their journey.

With a fantastic keynote session & deep-dive sessions with Databricks, you'll learn how:

- Your entire organization can gain insights from data using natural language
- You can develop quality AI applications on your data — without sacrificing data privacy or control
- Data intelligence simplifies governance, data engineering, data streaming, data warehousing and user experiences

Agenda

AI Days	8:00 AM - 9:00 AM	Breakfast & Registration
	9:00 AM - 9:30 AM	Welcome + Keynote
	9:30 AM - 9:45 AM	Customer Story
	9:45 AM - 10:15 AM	Analytics: Why Yesterday's BI Can't Answer Today's Questions
	10:15 AM - 10:45 AM	Break & Networking
	10:45 AM - 11:15 AM	AI Agents: How to build and deploy high-quality AI agents that actually work
	11:15 AM - 11:45 AM	Apps: The complete agent stack for building modern AI applications
	11:45 AM - 12:00 PM	Partner Story
	12:00 PM - 12:15 PM	Partner Story
	12:15 PM - 12:30 PM	Closing Remarks
	12:30 PM - 1:30 PM	Lunch & Networking

Speakers


Jeremy van Doorn
DATABRICKS
Senior Director, Field Engineering


Alena Paselle
DATABRICKS
Senior Solutions Engineer


Stef Creemers
DATABRICKS
Senior Solutions Engineer


Sander Lam
DATABRICKS
Senior Solutions Engineer