



Databricks Asset Bundles

Lunch & Learn

Johan Hostens
30 January 2026



databricks

Databricks Intro

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

The screenshot displays the Databricks workspace interface. At the top, there is a search bar and a user profile icon. The left sidebar contains a navigation menu with the following items: New, Workspace, Recents, Data, Workflows, Compute, SQL, SQL Editor, Queries, Dashboards, Alerts, Query History, SQL Warehouses, Data Engineering, Delta Live Tables, Machine Learning, Experiments, Feature Store, Models, Serving, Marketplace, Partner Connect, Disable new UI, Provide feedback, and Collapse menu. The main content area is titled 'Get started' and features three primary action cards: 'Import and transform data' (with buttons for 'Create table' and 'Create pipeline'), 'Notebook' (with a 'Create a notebook' button), and 'SQL query editor' (with a 'Create query' button). Below these cards, there are sections for 'Recents' and 'Popular'. The 'Recents' section lists recent items such as 'IOT Platform - Turbine analysis' (Dashboard, 2 minutes ago), '01_Data Prep' (Notebook, 16 days ago), 'var_explorer_demo' (Notebook, 27 days ago), 'Demo 11_17' (Dashboard, 28 days ago), 'NB1' (Notebook, 28 days ago), 'jockey_odds_bytrack' (Query, 28 days ago), 'Untitled Notebook 2023-06-14 17:30:49' (Notebook, 28 days ago), '01_Data_Prep' (Notebook, 28 days ago), and 'train_dolly' (Notebook, 42 days ago). The 'Popular' section lists popular items like 'Motion Glob' (Dashboard), 'IOT Platform' (Dashboard), 'DLT - Retail' (Dashboard), 'IOT Platform' (Dashboard), 'M-Magn SSS' (Notebook), 'M-Heatmap' (Notebook), 'M-Histo' (Notebook), 'dbsql_tpch_' (Table), and 'demo_frank.' (Table).

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	<u>Object Model</u>
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



UNITY CATALOG

 Subscription [DEV]

 Resource group [DEV]

 Databricks workspace

 Access connector

 Storage account

 Keyvault

 Managed resource group

 Storage account

 Subscription [QA]

 Resource group [QA]

 Databricks workspace

 Access connector

 Storage account

 Keyvault

 Managed resource group

 Storage account

 Subscription [PRD]

 Resource group [PRD]

 Databricks workspace

 Access connector

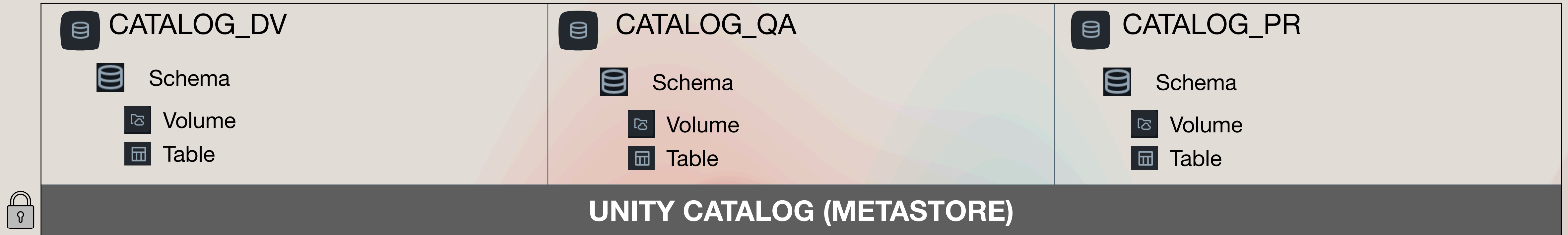
 Storage account

 Keyvault

 Managed resource group


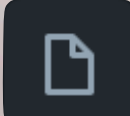








 Storage account

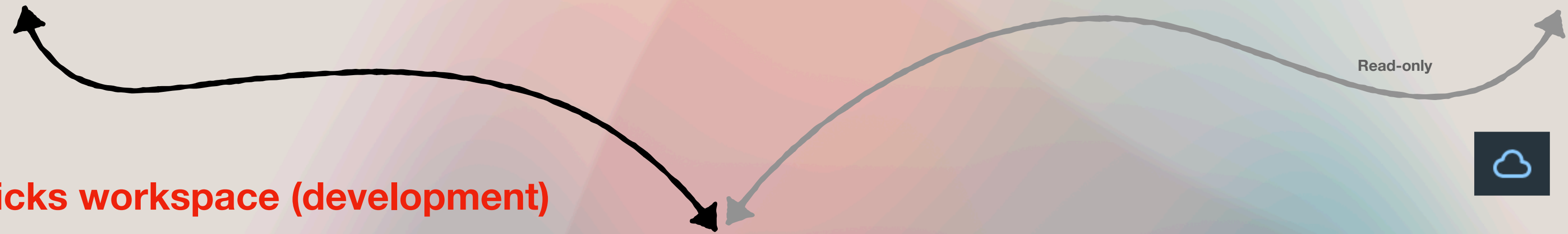
Databricks Workspace



 **Databricks workspace (development)**

 **Serverless Compute**

SCRIPTS / NOTEBOOKS	DATA	WORKFLOWS	COMPUTE
<ul style="list-style-type: none">  Notebooks (*.ipynb)  Scripts (*.py)  Packages (*.whl) 	<ul style="list-style-type: none">  CATALOG_DV  CATALOG_PR 	<ul style="list-style-type: none">  Jobs (Procedural) <ul style="list-style-type: none"> - Notebook tasks - Python Script/whl tasks - dbt tasks - Power BI tasks - ...  Pipelines (Declarative) <ul style="list-style-type: none"> - Ingestion pipelines - ETL pipelines 	<ul style="list-style-type: none">  All-purpose compute  Job compute  SQL warehouse





[1] Databricks introduction

Jobs

- Activity types
- Trigger types
- Compute
- Notifications

Catalogs

- Descriptions
- Lineage
- Audit tables







Databricks Asset Bundles

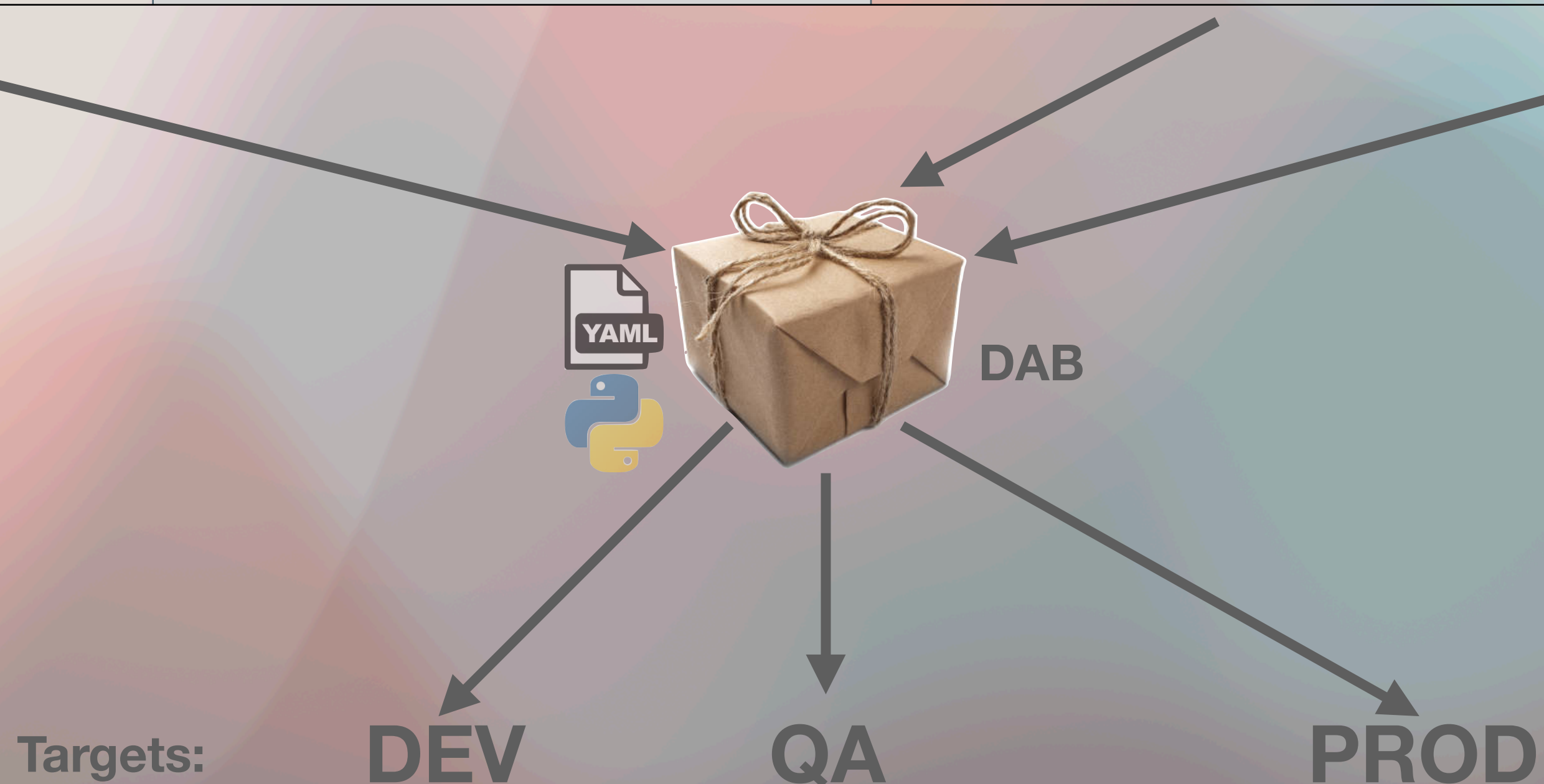
- **What is a DAB?**
- **Test project summary**
- **Visual Studio Code configuration**
- **(Demo)**

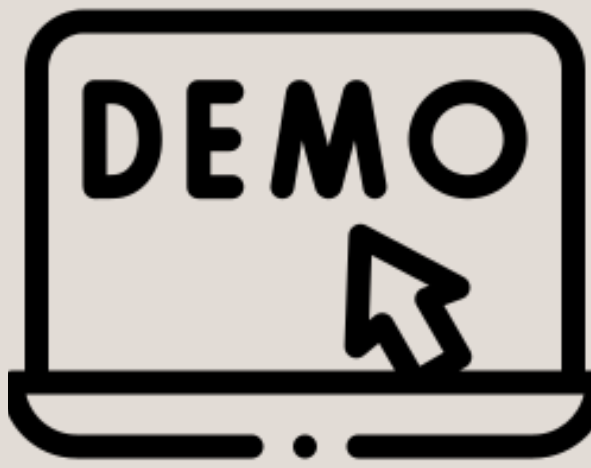


What is a DAB?

 Databricks workspace (development)

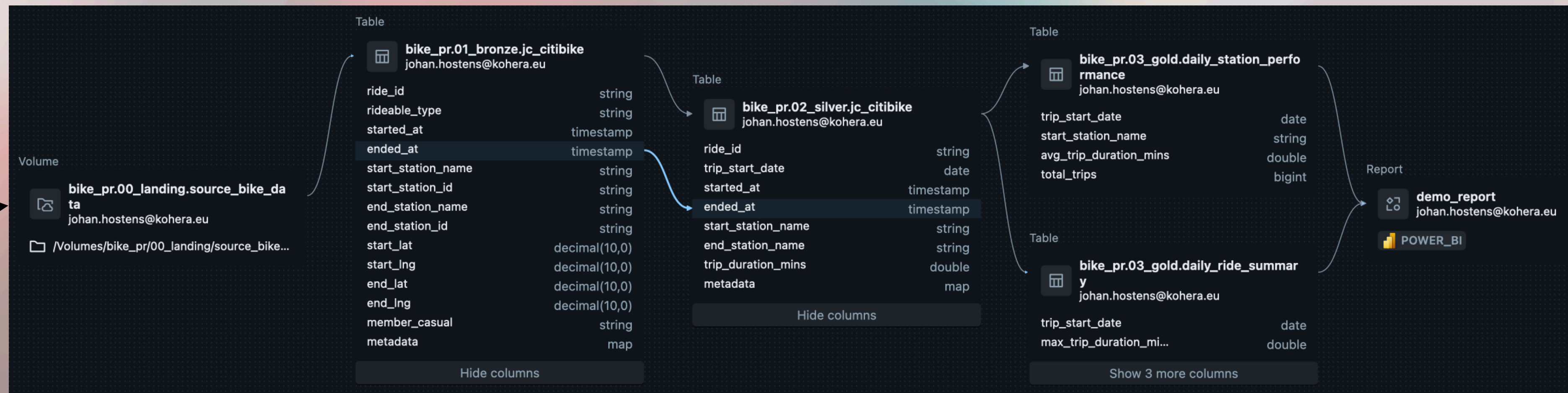
SCRIPTS / NOTEBOOKS		WORKFLOWS	COMPUTE
<ul style="list-style-type: none"> Notebooks (*.ipynb) Scripts (*.py) Packages (*.whl)		<ul style="list-style-type: none"> Jobs (Procedural)<ul style="list-style-type: none">- Notebook tasks- Python Script/whl tasks- dbt tasks- Power BI tasks- ... Pipelines (Declarative)<ul style="list-style-type: none">- Ingestion pipelines- ETL pipelines	<ul style="list-style-type: none"> Job compute configuration





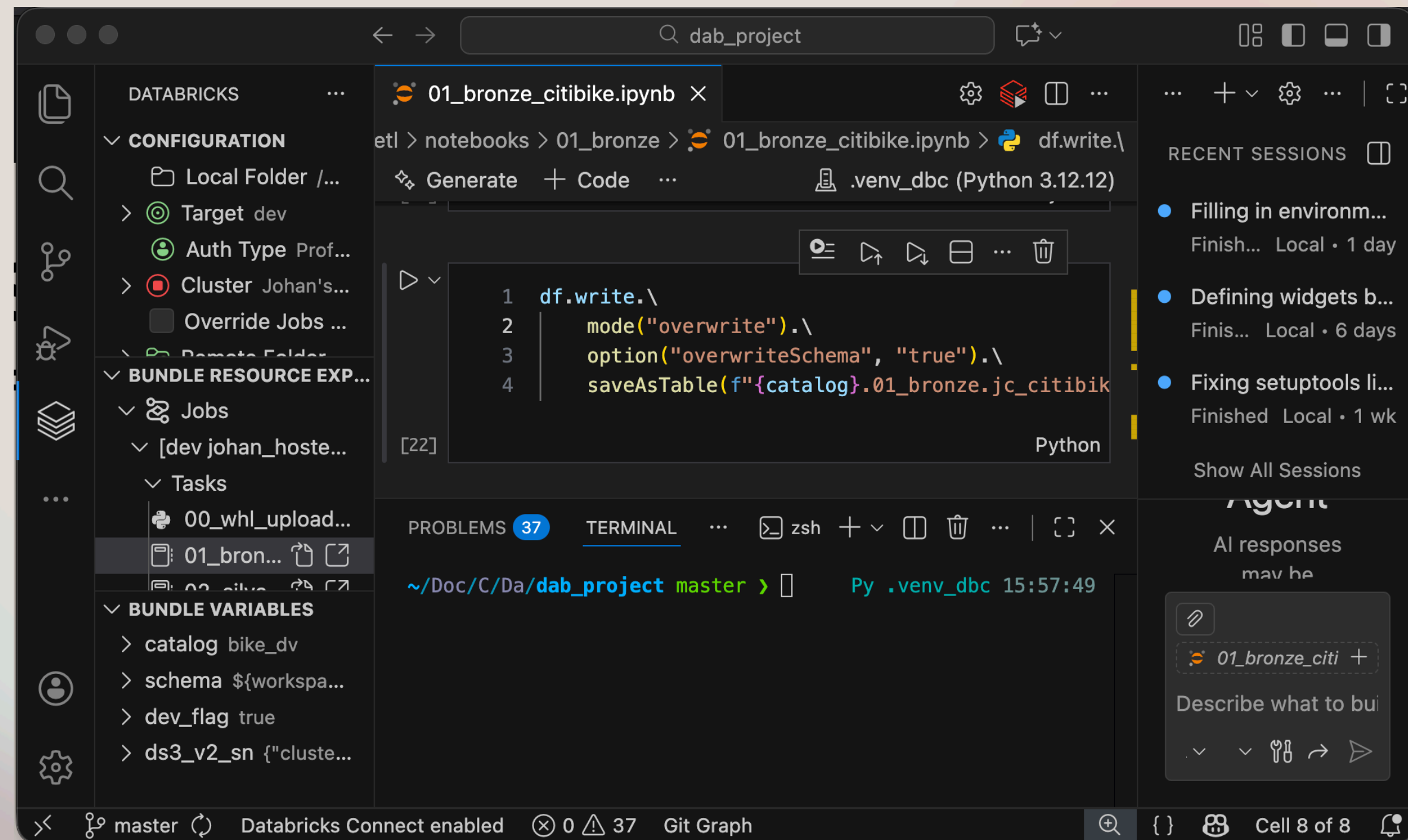
Test project summary

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 (yaml)	Job (yaml)	Pipeline (yaml)



Visual Studio Code

Databricks Plugin



Combined with

Formatter: **Black** 

Linters: **Ruff** 

AI-assisted coding: **GitHub Copilot** 

Type checker: **mypy** : my[py]



Auto-documentation: **Sphinx** 

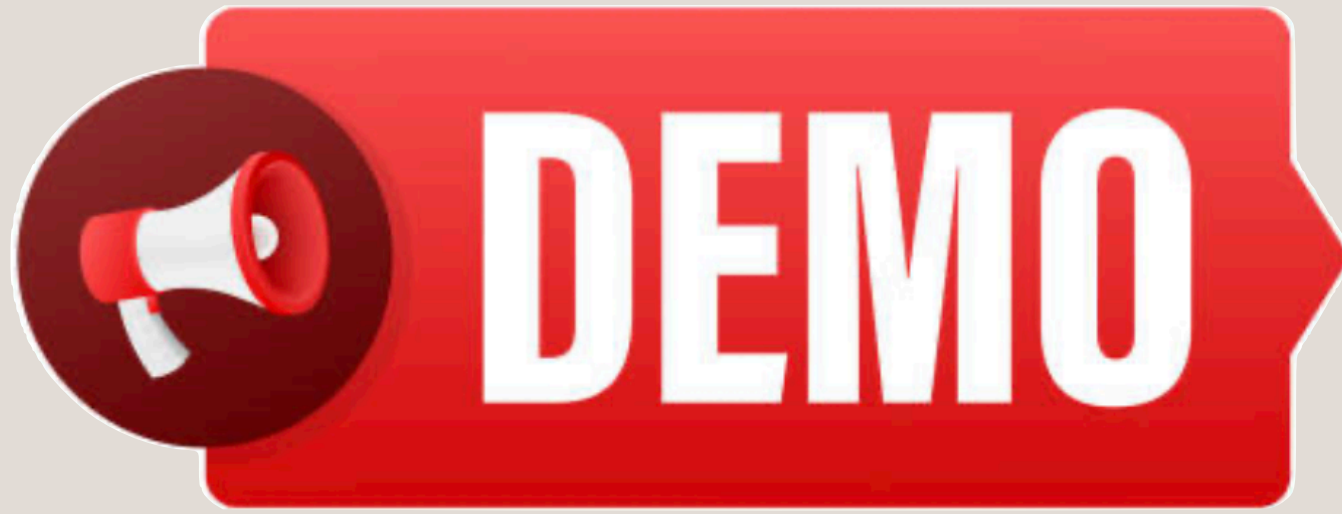
Testing framework: **pytest** 

Azure resources 

dbt projects 

SQL databases 

...



[2] Databricks in VS Code

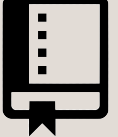







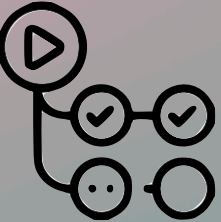
- VS Code Databricks plugin
- Databricks CLI
- DAB project
- Deployed resources in development / production
- Automated documentation with Sphinx
- Pytest results

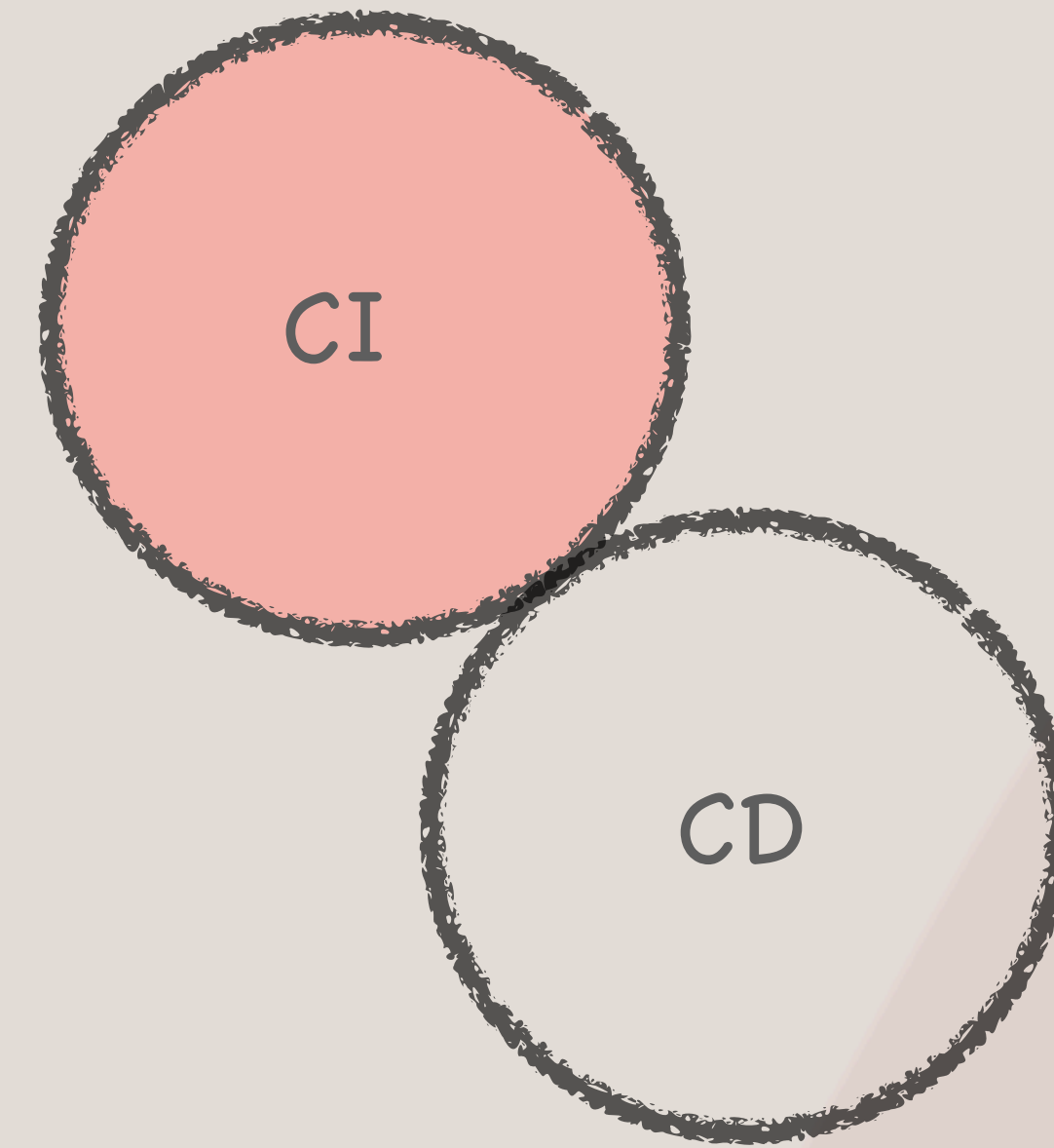
Github: CI/CD

- **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

The screenshot shows the GitHub Actions interface for a workflow named 'service principals #3'. The workflow is in a 'Completed' state, indicated by a green checkmark. The main area displays a list of job steps, all of which have completed successfully. The steps include setting up the job, checking out code, setting up Python 3.12, installing dependencies, running tests, and uploading reports. A sidebar on the left provides navigation options for the workflow run, including a summary, usage, and workflow file.

konverga / dab_project

Code Issues Pull requests **Actions** Projects Security Insights Settings

← CI Workflow

✓ service principals #3 Re-run all jobs ⋮

Summary

All jobs

✓ CI Job

Run details

Usage

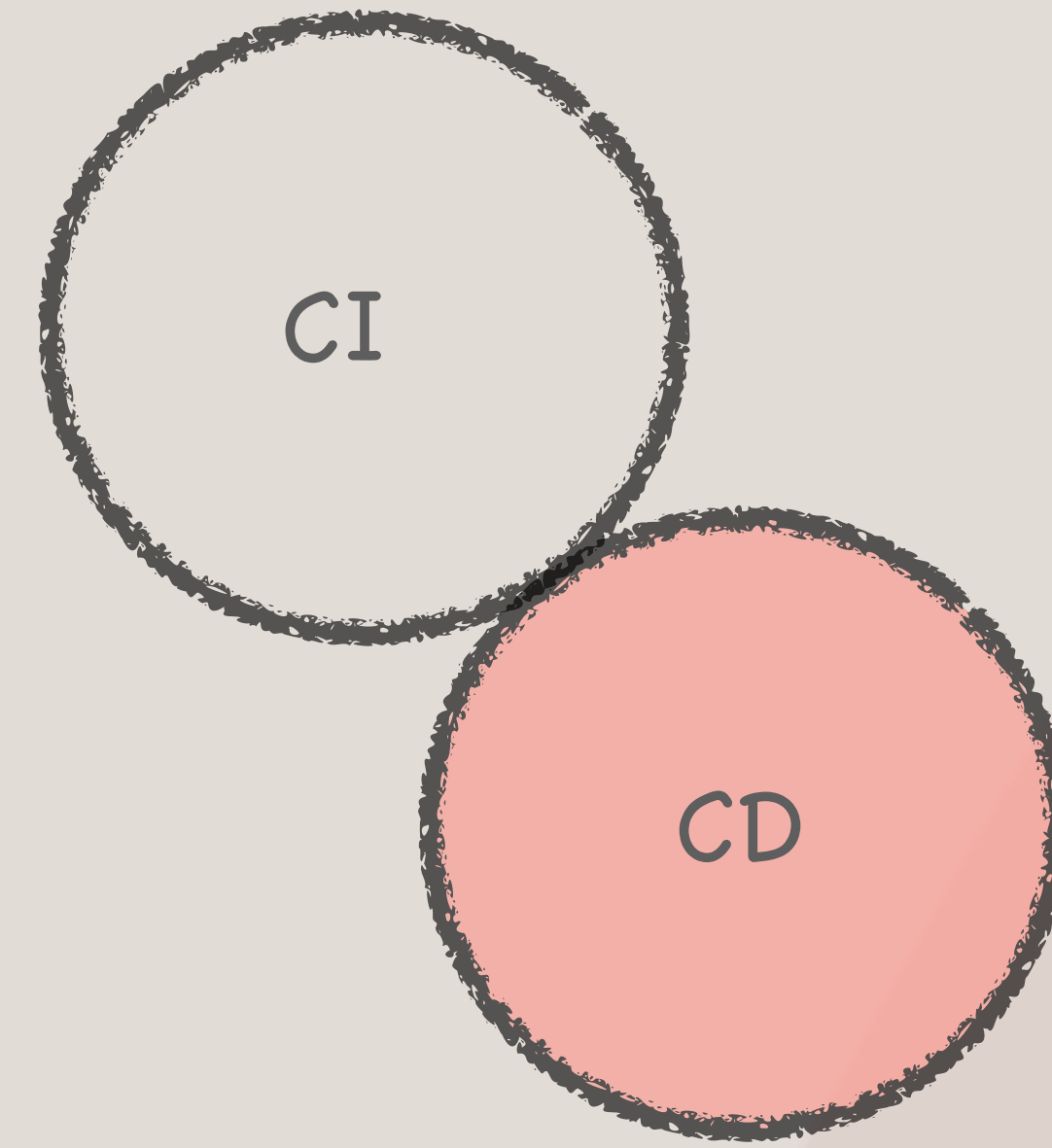
Workflow file

CI Job
succeeded last week in 1m 13s






Search logs

> ✓ Set up job	2s
> ✓ Checkout code	2s
> ✓ Setup Python 3.12	0s
> ✓ Install dependencies + pytest-html	48s
> ✓ Run pytest & generate HTML report	16s
> ✓ Upload coverage HTML report	1s
> ✓ Post Setup Python 3.12	0s
> ✓ Post Checkout code	1s
> ✓ Complete job	0s

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

The screenshot shows a GitHub Actions workflow run for the 'Merge pull request #2 from konverga/feature/service_principal #4'. The workflow is titled 'CD Workflow' and is currently in a 'Succeeded' state. The main job is 'Deploy to Quality Assurance', which succeeded last week in 26 seconds. The workflow consists of several steps, all of which are completed successfully.

Workflow Summary:

- Merge pull request #2 from konverga/feature/service_principal #4 (Succeeded)
- Deploy to Quality Assurance (Succeeded)
- Deploy to Production (Succeeded)

Deploy to Quality Assurance Job Details:

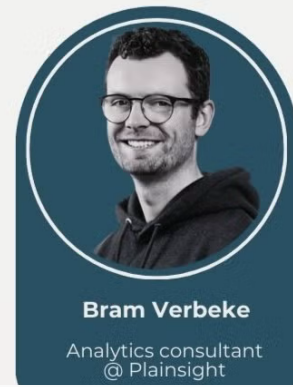
succeeded last week in 26s

Step	Duration
Set up job	0s
Checkout code	1s
Setup Python	0s
Install Databricks CLI	1s
Install Dependencies	2s
Configure Databricks	0s
Deploy to Quality Assurance	20s
Post Setup Python	0s
Post Checkout code	0s
Complete job	0s

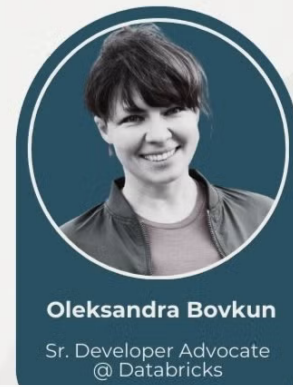
Fun Events in the Spotlight

12 TH DATABRICKS MEETUP BELGIUM

February 12, 2025
Plainsight, Mechelen



Bram Verbeke
Analytics consultant
@ Plainsight



Oleksandra Bovkun
Sr. Developer Advocate
@ Databricks

Hosted by:
PLAINSIGHT

- 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!



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 Clean Rooms (currently in preview), Cluster permissions, and budgets

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