



# Databricks Roadmap

Q3 FY25

---

10 de setembro de 2024



# Confidencial e sujeito a alterações

Esta apresentação descreve a direção geral do produto da Databricks, que pode mudar à medida que continuamos inovando.

Quaisquer serviços, recursos, funcionalidades ou aprimoramentos em potencial mencionados nesta apresentação estão sujeitos a alterações a critério da Databricks. Esses itens podem não ser entregues conforme planejado ou não ser entregues.

As informações nesta apresentação são confidenciais e não podem ser compartilhadas sem primeiro receber permissão por escrito da Databricks.

O Databricks permite que você  
realmente  
**Democratize dados + IA**



# Reconhecida como líder no setor

## Líder do Forrester Wave : Data Lakehouse



## Líder no Quadrante Mágico do Gartner para Ciência de Dados e Machine Learning

Figure 1: Magic Quadrant for Data Science and Machine Learning Platforms



Source: Gartner (June 2024)



# Data Intelligence Platform

Unity Catalog

Delta Lake UniForm

# Databricks Data Intelligence Platform

## Tópicos

---

Segurança e governança unificadas

Data Warehousing & BI

Engenharia de dados e streaming

Data Sharing and Colaboração

Armazenamento e gerenciamento unificados de dados

Experiência do desenvolvedor

Administração Empresarial

Mosaic AI

# Databricks Data Intelligence Platform

Segurança e governança unificadas

## Tópicos

---

Unity Catalog

Segurança corporativa

# Unity Catalog: Próximos lançamentos

## IA para governança

-  Expectativas unificadas no Lakehouse Monitoring (Public Preview)
-  Classificação de dados (Public Preview)

## Governança para IA

-  Linhagem de endpoint de modelo (GA)
-  AI Tools Catalog (Private Preview)

## Acessibilidade aberta

-  Gravações de tabela gerenciadas em APIs REST de UC (Private Preview)

## Federação de Lakehouse

-  Salesforce Data Cloud (Public Preview)
-  Conectores Oracle e Teradata (Private Preview)
-  Federação do Hive – interna, externa, Glue (Private Preview)

## Recursos avançados de governança

-  Controles de acesso baseados em atributos (ABAC) (Private Preview)
-  Tags Governadas (Private Preview)

## Administração simplificada

-  Disaster Recovery (Private Preview)

# Classificação de dados

Classificar automaticamente dados para PII

**Detecte e classifique automaticamente os dados de PII**

Classificadores integrados para 16 tipos de PII

Alertas para detecção de PII em uma coluna

Suporte para regras personalizadas e políticas ABAC em breve

Private Preview



## Catalog Explorer

Catalog Explorer > main > default >

main.default.my\_table

Overview Sample Data Details Permissions History Lineage Insights Quality

Filter columns...

Column	Type	Comment	Tags
CREDIT_CARD	string	⋮	pii_credit_card
DATE_OF_BIRTH	string	⋮	pii_date_of_birth
DATE_TIME	string	⋮	pii_date_of_birth
EMAIL_ADDRESS	string	⋮	pii_email
IBAN_CODE	string	⋮	pii_email

LOI  
NAI  
prod-monitoring@databricks.com  
to me



Alertas de detecção de PII

Lakehouse Monitoring detected new classification tags for table main.default.my\_table

### Message

Lakehouse Monitoring has detected the following classification tags on table main.default.my\_table:

- pii\_location on column EMAIL\_ADDRESS
- pii\_date\_of\_birth on column DATE\_OF\_BIRTH
- pii\_us\_bank\_number on column US\_BANK\_NUMBER
- pii\_location on column LOCATION
- pii\_race on column RACE
- pii\_us\_drivers\_license on column US\_PASSPORT
- pii\_name on column NAME

# Controles de acesso baseados em atributos

Dimensionar o gerenciamento de acesso

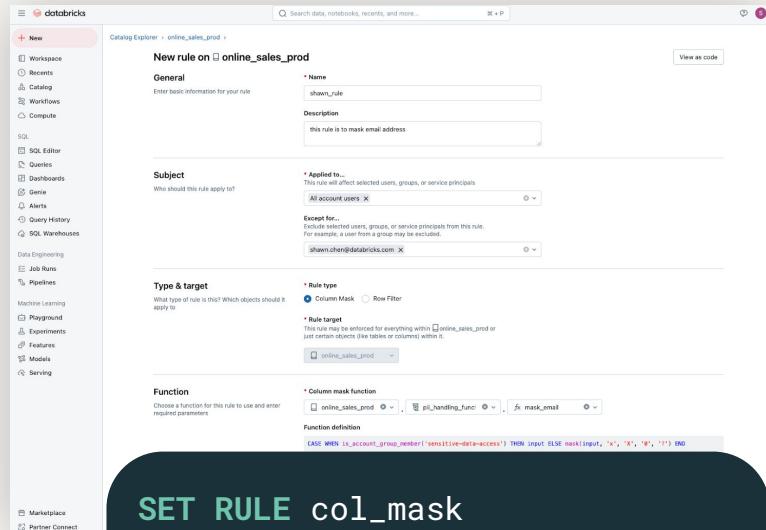
**Políticas escaláveis em todos os dados + ativos de IA usando tags e atributos controlados**

Fases:

**Fase 1 – RLS/CLM em uma abstração mais alta do que uma tabela, regras baseadas em tags controladas**

**Fase 2 – Regras GRANT/DENY.**

Private Preview



```
SET RULE col_mask
ON CATALOG turbine_prod
COLUMN MASK masking_function
TO `account users`
FOR tables
WHEN COLUMN tagged(`pii`)
```

# Metastore do Hive e Federação do Glue



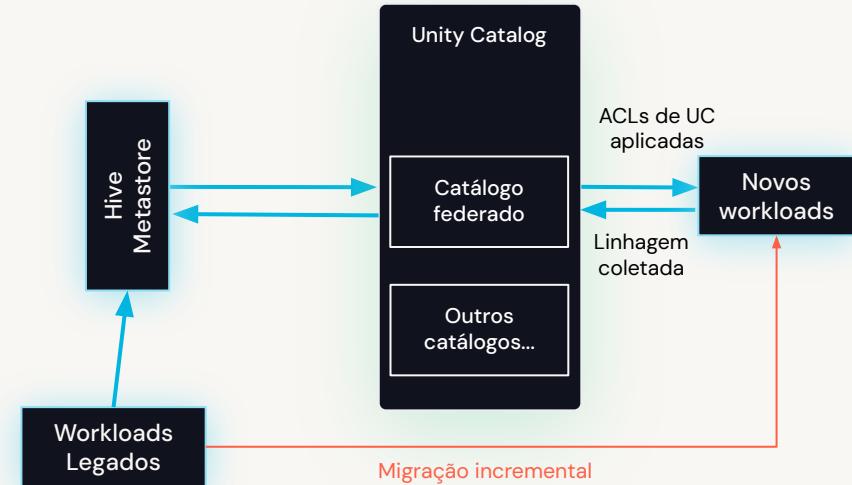
Simplifique a migração do Hive Metastore (HMS), Glue

**Descubra, controle e accesse dados de  
HMS interno/externo, AWS Glue**

Monte o Metastore interno ou externo do Hive, AWS Glue no Unity Catalog.

Acesse seus ativos existentes e aplique a governança de UC para novas cargas de trabalho.

Suporta ACLs, acesso entre workspaces, linhagem, pesquisa e muito mais.



Private Preview

# Melhorias no desempenho do parquet



**Desempenho mais rápido lendo  
tabelas Parquet externas no Catálogo  
do Unity**

Desempenho de leitura acelerado

Supporte para formatos de arquivo Parquet e ORC

Integração simplificada ao Unity Catalog

GA Q4

**30x mais  
rápido\***

Lendo o Parquet no  
Unity Catalog

# Unity REST APIs

Interoperabilidade entre diferentes mecanismos de computação

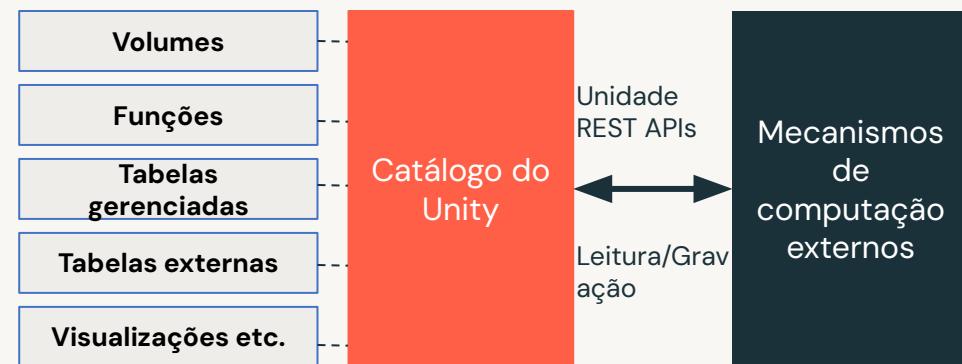
**Leitura e gravação em tabelas gerenciadas pelo UC através de mecanismos de computação externos**

Leitura e gravação em tabelas gerenciadas pelo UC por mecanismos externos integrados às APIs abertas de UC

Próximas integrações de clients: OSS Spark, DuckDB, Daft, PuppyGraph, StarRocks

O acesso do client externo é governado pelo Unity Catalog

Private Preview



# Databricks Data Intelligence Platform

## Data Warehousing & BI

### Tópicos

---

Databricks SQL

AI/BI

# Databricks SQL

Construído sobre uma arquitetura de lakehouse

**Arquitetura unificada**  
para todos os seus dados, BI e  
IA

**Governança unificada**  
para todos os seus dados +  
ativos de IA

**Equipes de dados unificadas**  
com SQL + Python para todos

**Compartilhamento de dados**  
**abertos**  
e marketplace

**Reduza custos** com o único warehouse  
que não cria um novo silo de dados

**Simplifique a governança** com uma única  
fonte de verdade

**Melhore a produtividade** das pessoas  
trabalhando com SQL e Python juntos

**Melhore a colaboração**  
compartilhando dados + ativos de IA com  
segurança com qualquer pessoa

# Databricks SQL: Upcoming Releases

## Intelligent Experiences with AI

- ➡️ LLM via SQL (Public Preview)
- ➡️ Vector Search via SQL (Public Preview)
- ➡️ AI Forecast via SQL (Public Preview)
- ➡️ Inline Assistant in SQL Editor (GA)

## Enterprise Readiness

- ➡️ Budget Policies
- ➡️ Permissões de monitoramento
- ➡️ Melhorias nas system tables
- ➡️ APIs for Queries, Dashboards and Alerts
- ➡️ Improved PowerBI & Tableau integration

## Core Warehousing

- ➡️ Materialized Views E Streaming Tables (GA)
- ➡️ Variant data type (Public Preview)
- ➡️ SQL Scripting (Private Preview)
- ➡️ Collations Support (Private Preview)
- ➡️ Spatial SQL enhancements (coming soon)



# Análise poderosa de LLM/IA, simplificada



## Funções SQL para casos de uso de IA



Funções de IA integradas  
para análise pronta para uso



Conecte-se a modelos  
personalizados ou externos  
usando ai\_query()



Taxa de transferência  
eficiente para inferência de  
modelo em lote

Classificação

Forecasting (novo\*)

Correção gramatical

Máscara

Reconhecimento de entidade nomeada

Análise de Sentimento

Similaridade

Resumo

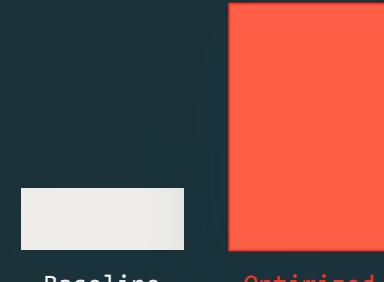
Geração de texto

Tradução

```
select ai_query(  
    endpoint => 'model-endpoint',  
    request => concat('Summarize  
    the following document.\n',  
    doc)  
) as summary  
from table
```

Melhoria de 4x no throughput  
usando ai\_query() últimos 3 meses

tokens per second



\* with the exception of ai\_forecast() in Gated Public Preview

Public Preview



# Vector Search no Databricks SQL

## Construa RAG em uma única instrução SQL



Implemente facilmente um Vector Search para RAG, recomendações, análises e muito mais



Conecte a pesquisa aos seus aplicativos com a API SQL Statement Execution



Utilização do Databricks Mosaic AI Vector Search



Totalmente seguro e governado pelo Unity Catalog

Gated Public Preview

```
// query vector search endpoint
with search as (
    select array_join(collect_set(Note), '\n') as results
    from VECTOR_SEARCH(
        index => 'demo.medical.notes',
        query => :question,
        num_results => 3
    )
    group by all
)
// summarize the results
select ai_gen(concat_ws(
    '',
    "INSTRUÇÕES: Você é um as",
    "DOCUMENTOS:\n", :question
)) as Answer
from search
```

Quais são os tratamentos recentes para artrite em pacientes?

Responder

1. Byron Sampson, um homem de 63 anos com osteoartrite, recebeu prescrição de prednisona 20Mg por dia durante 1 semana para possível artrite reumatóide (01/13/2024)

2. Miles Ray, um homem de 58 anos, recebeu paracetamol conforme necessário para o alívio da dor nas articulações. (11/15/2023)

# UDFs do Python (funções definidas pelo usuário)

## Aprimore seus workloads SQL usando Python



Traga lógica avançada e o poder do Python para suas cargas de trabalho SQL



Armazenado e controlado no Unity Catalog sob o namespace de 3 níveis



O isolamento do Lakeguard garante que a governança de dados seja totalmente aplicada na camada de computação



Definir dependências personalizadas (PyPi ou .whl)

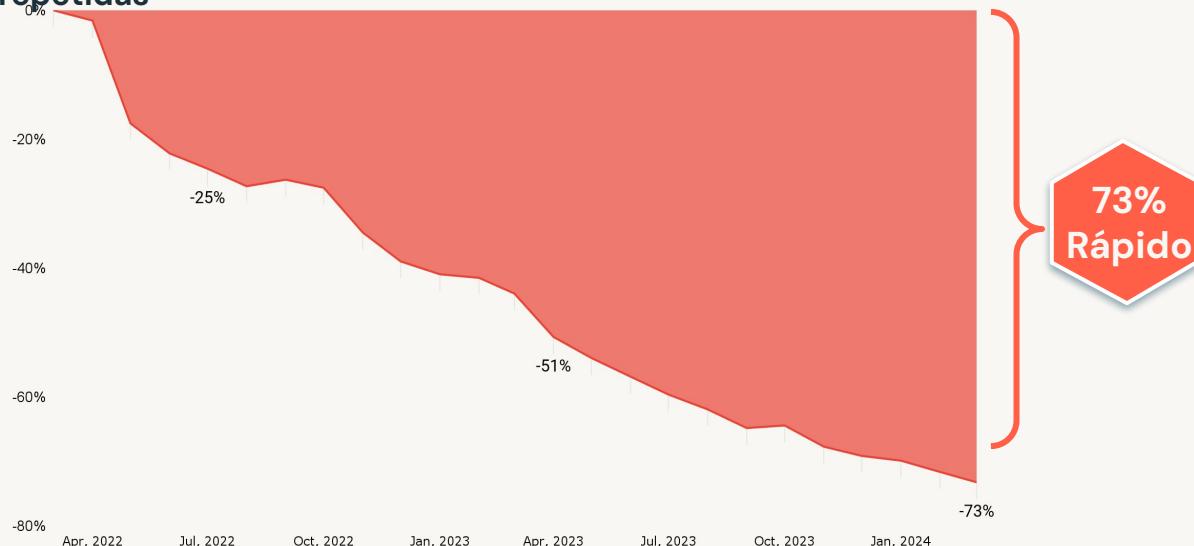
```
CREATE FUNCTION redact(a STRING)
RETURNS STRING
LANGUAGE PYTHON
AS $$

import json
keys = ["email", "phone"]
obj = json.loads(a)
for k in obj:
    if k in keys:
        obj[k] = "REDACTED"
return json.dumps(obj)
$$;
```

GA até Q4

# Melhorias de desempenho baseadas em IA

## Melhorias de desempenho de todas as cargas de trabalho de BI do cliente repetidas



Investimentos em recursos baseados em IA para melhorar continuamente o desempenho de cargas de trabalho de BI, exploratórias e ETL

- Simultaneidade aprimorada
- Latência de consulta curta
- Desempenho pronto para uso (ou seja, sem ajuste manual)

- 73% de melhoria automática no desempenho nos últimos 2 anos

- + 5% desde março

# Databricks Data Intelligence Platform

## Data Warehousing & BI

### Tópicos

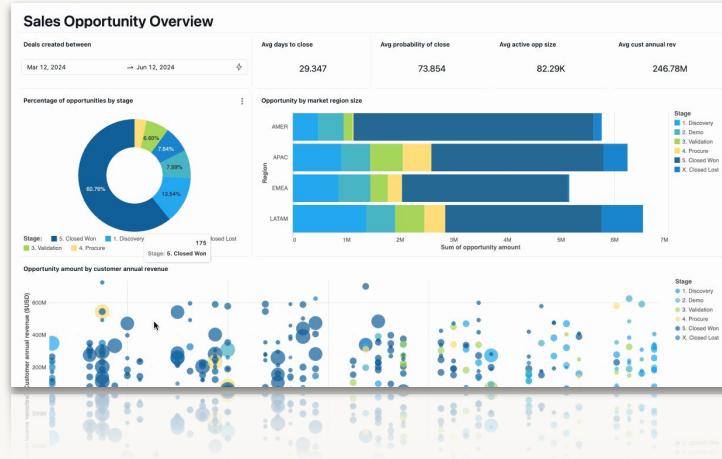
---

Databricks SQL

AI/BI

# Introduzindo Databricks AI/BI

Análise inteligente para dados do mundo real



## Dashboards

Crie e compartilhe rapidamente insights visuais interativos

GA

The Genie Spaces interface for Frontier Sports Sales Data. It includes:

- Header: databricks Genie Spaces, Share button.
- Section: Frontier Sports Sales Data (with a small orange icon).
- Description: Gold data set for sales performance by customer, product, and employee.
- Actions: Sell-through rates by category, RFM by customer segment, Top line KPI's for current FY, Explain the data set.
- Input fields: Ask your question... and say how do you... (with a microphone icon).

## Genie

Converse com seus dados em linguagem natural

Public Preview

# AI/BI: Upcoming Releases

## AI/BI Dashboards

-  Cross-filtering (GA)
-  Maps Visual – Plot Points/Markers (GA)
-  Box Plot Visual (GA)
-  Launch Genie from Dashboard (Public Preview)
-  Publish Notebook SQL Cell (incl viz) to Dashboard (GA)
-  Dashboard Embedding (Private Preview)
-  Multi-Page Dashboards (Private Preview)

## AI/BI Genie

-  GCP Support (Public Preview)
-  Clone Genie space (GA)
-  Genie Audit Logs (GA)
-  Genie Questions API (Private Preview)
-  Share space with non-Workspace user (Private Preview)
-  Genie Benchmarks (Private Preview)

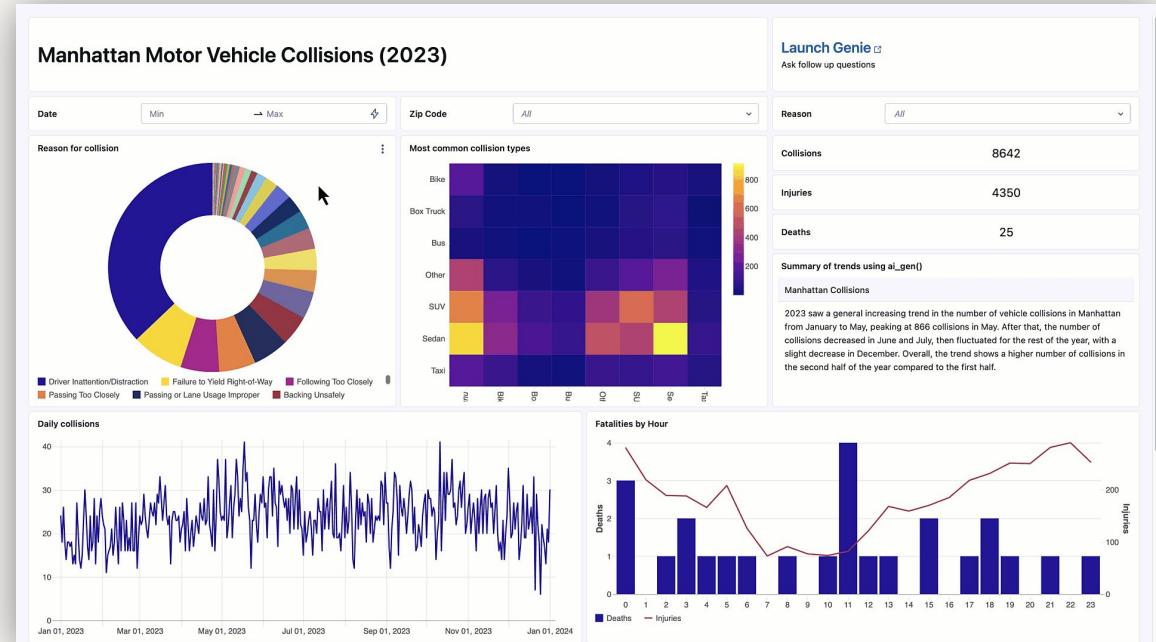


# Dashboard cross-filtering

Os consumidores do painel podem interagir com visualizações para ver o efeito em outras visualizações.

Clique no componente do gráfico e veja os resultados atualizados instantaneamente em todo o seu painel.

Identifique padrões, correlações e valores discrepantes com mais facilidade.

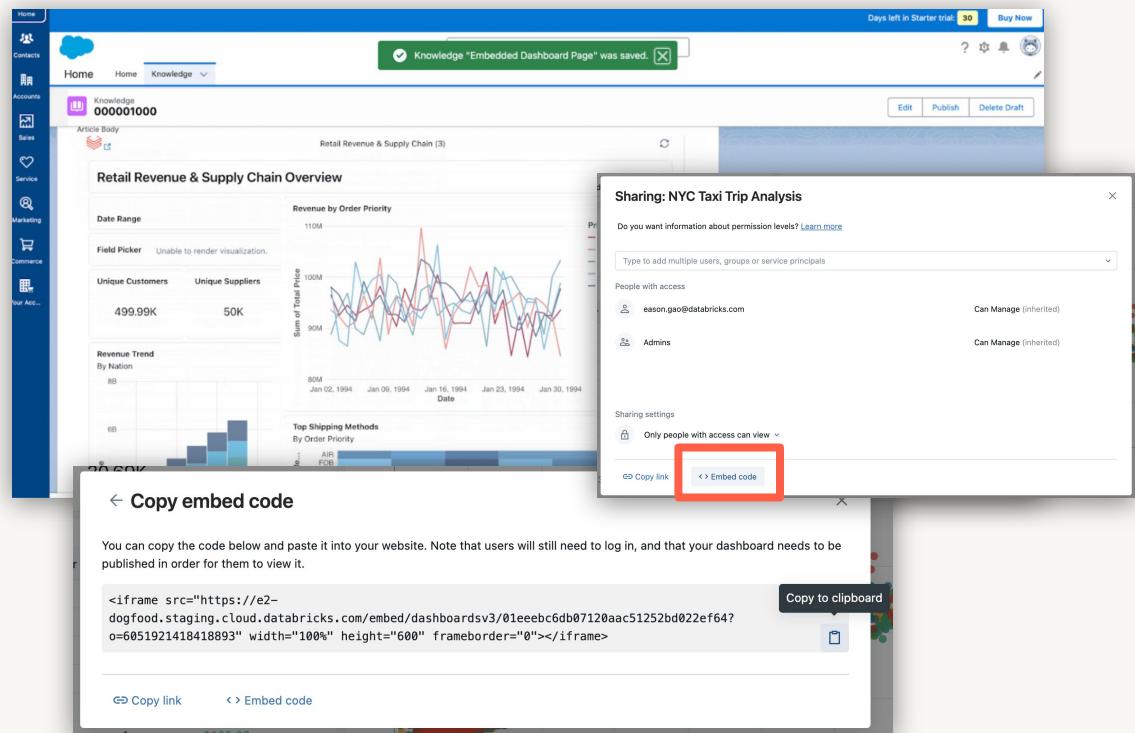


GA

# Dashboard embedding

Incorpore com segurança seus painéis de IA/BI em sites externos e aplicativos de terceiros (por exemplo, Confluence, Salesforce, Sharepoint)

Para incorporar seu painel, basta clicar em "Embbbed code" na caixa de diálogo do painel, copiar o trecho de código e inseri-lo na superfície incorporada!



Private Preview

# Genie Conversation APIs

## Uma maneira programática de interagir com AI/BI Genie



Faça perguntas e receba respostas por meio da nova API de conversação do Genie

Crie sua própria interface de usuário que interaja com o mecanismo Genie

Visualização inicial limitada apenas a operações de perguntas e respostas. Operações CRUD (por exemplo, criar espaço Genie ou salvar como instrução) no roteiro de longo prazo

Private Preview

The screenshot displays the Databricks API documentation for the Genie Conversation APIs. It features two main endpoints:

- Start conversation**:
  - Request samples**: Shows a POST request to /api/2.0/genie/spaces/{space\_id}/start-conversation with a JSON payload: { "content": "Give me top sales for last month" }.
  - Response samples**: Shows a 200 OK response with a JSON object containing statement\_id, status, state, error, and message.
- Execute SQL query in a conversation message**:
  - Request samples**: Shows a POST request to /api/2.0/genie/spaces/{space\_id}/conversations/{conversation\_id}/messages/{message\_id}/execute-query with a JSON payload: { "content": "SELECT \* FROM sales WHERE month = 'last month'" }.
  - Response samples**: Shows a 200 OK response with a JSON object containing statement\_response, total\_chunk\_count, chunks, and byte\_count.

# Databricks Data Intelligence Platform

## Engenharia de dados

### Tópicos

---

LakeFlow Connect

Delta Live Tables

Streaming Engine

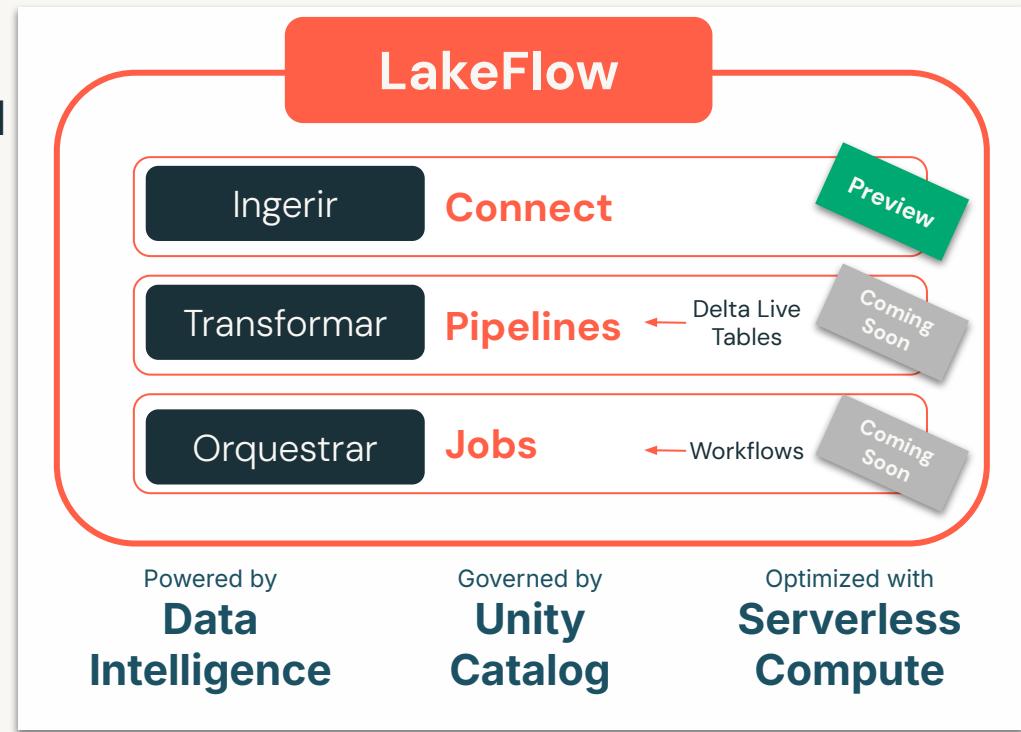
Databricks Workflows

# LakeFlow

Uma solução unificada de engenharia de dados alimentada por inteligência de dados

**LakeFlow Connect:** Disponível agora em Preview!

**LakeFlow Pipelines** e o **LakeFlow Jobs** serão 100% compatíveis com o **Delta Live Tables** e **Workflows**, ambos disponíveis hoje!



# Data Engineering: próximos lançamentos

## LakeFlow Connect



- New connectors for:  
ServiceNow, SharePoint, Google Analytics, Oracle  
NetSuite, SQL Server on-prem, and PostgreSQL  
(Private Preview)

## Serverless Compute



- Improved cost management for serverless  
compute (Private Preview)



- Serverless network (Public Preview)



- Serverless compute in GCP (Public Preview)

## Databricks Workflows



- ForEach Task Looping (GA)



- Table Triggers (Public Preview)



- SQL task results reference (Private Preview)



- PyDABs (Private Preview)

# Data Engineering: próximos lançamentos

## Delta Live Tables (DLT)

-  Unity Catalog integration (GA)
-  Direct Publishing Mode (Private Preview)
-  Support for Apache Kafka and external Delta table sinks (Private Preview)

## Streaming Engine

-  Support for Scala streaming workloads in shared clusters (GA)
-  Custom Python sources/sinks (batch and streaming support) (GA)
-  Improved Stateful Processing APIs (Public Preview)
-  Real-time mode for operational workloads and sub-second processing (Private Preview)



# LakeFlow Connect

Conectores de ingestão eficientes para todos

**Conectores adicionais para bancos de dados populares e aplicativos corporativos**

Ingestão eficiente e incremental

Fácil de configurar, com API ou interface visual

Nativo no Databricks, com integração e observabilidade do UC

Private Preview em Q3

Private Preview em Q3



PostgreSQL



Google Analytics



SharePoint



ORACLE  
NETSUITE



Public Preview Today!



Cloud based  
SQL Server



salesforce



# Federação

O que já é GA?

- MySQL
- PostgreSQL
- Amazon Redshift
- Salesforce Data Cloud
- Snowflake
- Microsoft SQL Server
- Azure Synapse (SQL Data Warehouse)
- Google BigQuery
- Databricks



Private Preview em Q3



Public Preview Today!



# Modo de publicação direta no DLT



Publique tabelas em catálogos e esquemas distintos a partir de *um único pipeline DLT*

Não há necessidade de implantar vários pipelines para cada catálogo e esquema → Custos de DLT mais baixos e menos carga operacional

Sintaxe unificada com DBSQL

DLT detecta dependências automaticamente! LIVE.\* não é mais necessária.

```
CREATE STREAMING TABLE
catalog1.schema1.table1 AS
SELECT * FROM
cloud_files(`path`)
```

```
CREATE MATERIALIZED VIEW
catalog2.schema2.table2 AS
SELECT day, sum(sales)
FROM catalog1.schema1.table1
GROUP BY day
```

Private Preview in Q3

# Integração DLT + UC

Adicionar governança de UC a tabelas criadas por pipelines DLT

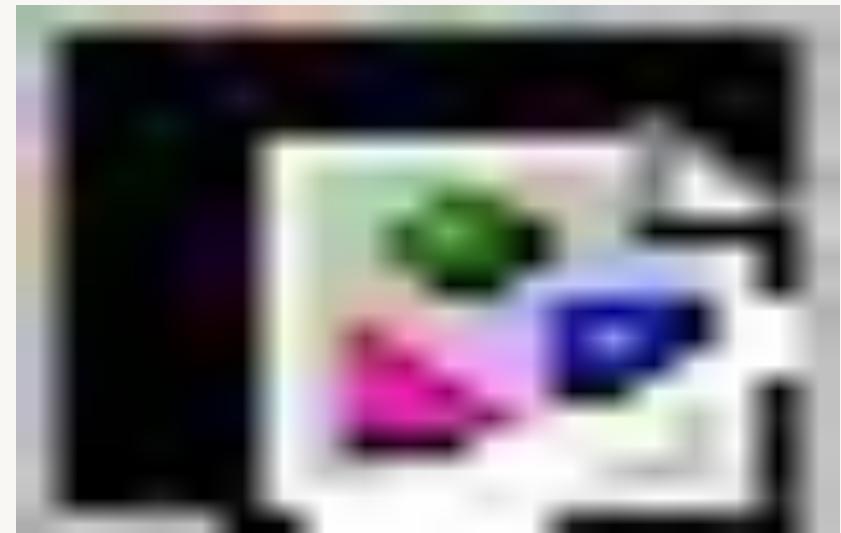
Log de eventos DLT acessível para não proprietários

Permitir a migração de tabelas para pipelines DLT

Fechar lacunas de paridade entre tabelas ST/MVs e Delta: RLS/CM, liquid clustering, ler um CDF de STs

Permitir a leitura de MVs e STs de mecanismos externos (não Databricks)

Torne MVs e STs compatíveis com UniForm



GA in Q3

# Table Triggers

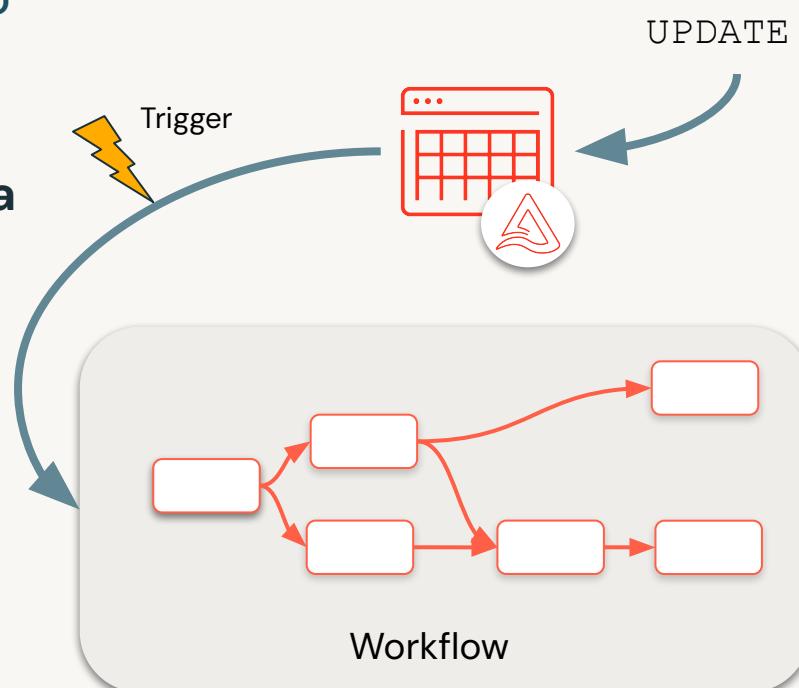
Executar um trabalho exatamente quando necessário

Acionar uma execução de trabalho quando ocorrer uma atualização em uma tabela delta especificada

Mantenha os dados atualizados com latência reduzida

Reduza as execuções de trabalho redundantes

Suporte para dependência de várias tabelas



Public Preview em Q3

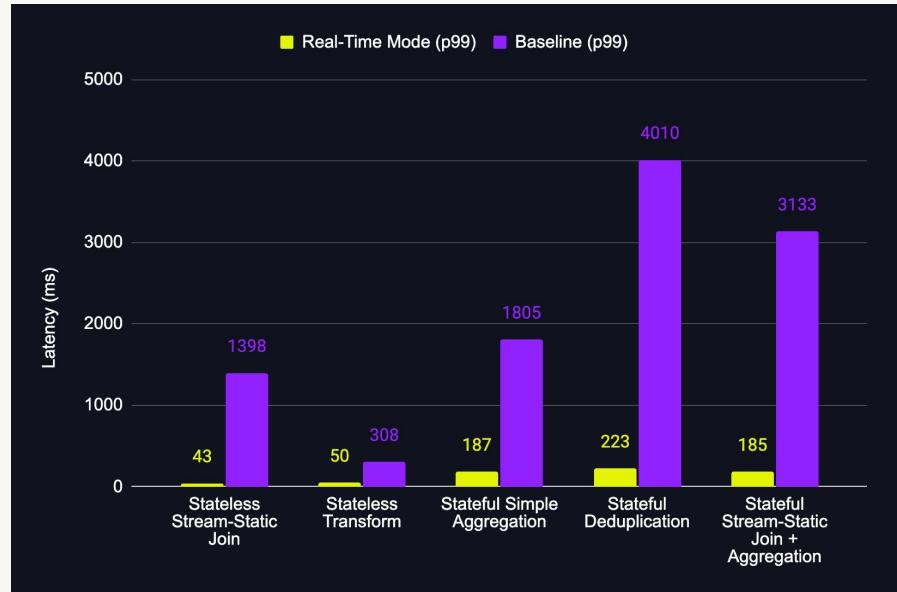


# Real-time mode

Latência em milissegundos para cargas de trabalho operacionais e processamento de menos de um segundo

Basta apertar um botão e funciona – sem alterações de código

Fornece latência em milissegundos e latência p99 entre 100–200ms em todos os casos de uso



Private Preview em Q3

# Task Looping

Fazer um loop em uma única tarefa com  
For Each no Databricks Workflows

Simplifica a criação para lidar com tarefas  
repetitivas

Crie fluxos de trabalho mais flexíveis

Inicialmente, repetindo uma única tarefa com  
fluxos de trabalho filhos chegando mais tarde



GA em Q3

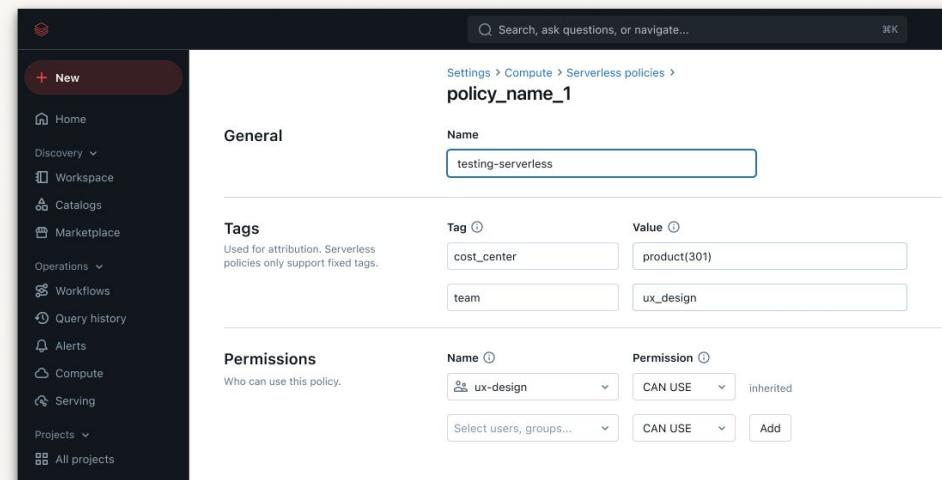
# Gerenciamento de custos aprimorado para computação serverless

Mais flexibilidade de custo e controle para computação serverless com Notebooks, Workflows e Delta Live Tables

Defina uma preferência para otimizar custo versus desempenho

Defina políticas de orçamento que imponham o tagueamento de recursos para relatórios de custos precisos

Private Preview em Q3



# Databricks Data Intelligence Platform

Confiabilidade e desempenho dos dados

Tópicos

---

Delta Lake UniForm

# Delta Lake UniForm: Upcoming Releases

## AI-powered Performance

-  Automatic Liquid Clustering Key Selection (Private Preview)
-  Automatic Statistics (Private Preview)
-  Fine-grained Predictive Optimization Enablement (Coming soon)

## Functionality

-  Default storage (Private Preview)
-  Type widening (Public Preview)
-  Liquid clustering performance improvements (Coming soon)
-  UniForm support for v1 readers (Coming soon)

# Liquid Clustering

**Clustering de alto desempenho e fácil de usar – não tenha mais dores de cabeça com partições!**

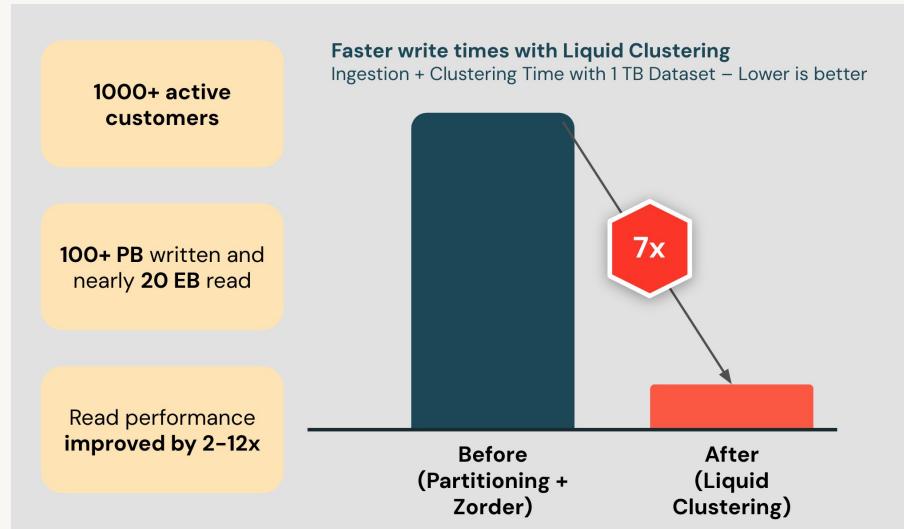
Reclustering de dados existentes com OPTIMIZE FULL

Integração: DLT, delta sharing

Suporte ao structured streaming

GA Q3

\*Also available in AWS & Azure GovCloud



"O Delta Lake Liquid Clustering **melhorou nossas consultas de séries temporais em até 10 vezes** e foi extremamente simples de implementar em nosso Lakehouse. Ele nos permite agrupar em colunas **sem nos preocupar com cardinalidade ou tamanho do arquivo** e reduz significativamente a quantidade de dados que precisa ser lido – algo que sempre tivemos que gerenciar com o particionamento Delta e o ajuste fino de ordem z."

– Bryce Bartmann (Chief Digital Technology Advisor, Shell)

# Seleção automática da chave do Liquid Clustering

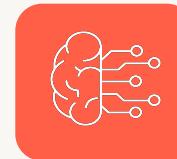
**Aplique ou evolua automaticamente chaves de clustering com base nos dados e no padrão de consulta dos clientes.**

Elimina completamente as decisões de layout de dados

Otimização de layout de dados totalmente autônoma

Desenvolvido por Otimização Preditiva

Private Preview



Liquid Clustering alimentado pela engine da Data Intelligence Platform

## Telemetria

Quais são os padrões de consulta de suas cargas de trabalho?

## Avaliação do modelo

Como a configuração do Liquid clustering deve ser atualizada?

## Optimização

Atualizar automaticamente as colunas de clustering e executar o clustering

# Estatísticas automáticas

**Coleta eficiente de estatísticas  
para melhorar o desempenho da  
consulta**

Execuções durante a ingestão + por meio da Otimização  
Preditiva

O modelo de IA prioriza tabelas para maximizar o ROI

Mais eficiente do que executar ANALYZE separadamente

Private Preview



**33% mais  
rápido\***

Em média, para  
consultas com  
estatísticas

# Armazenamento default para o Catálogo do Unity

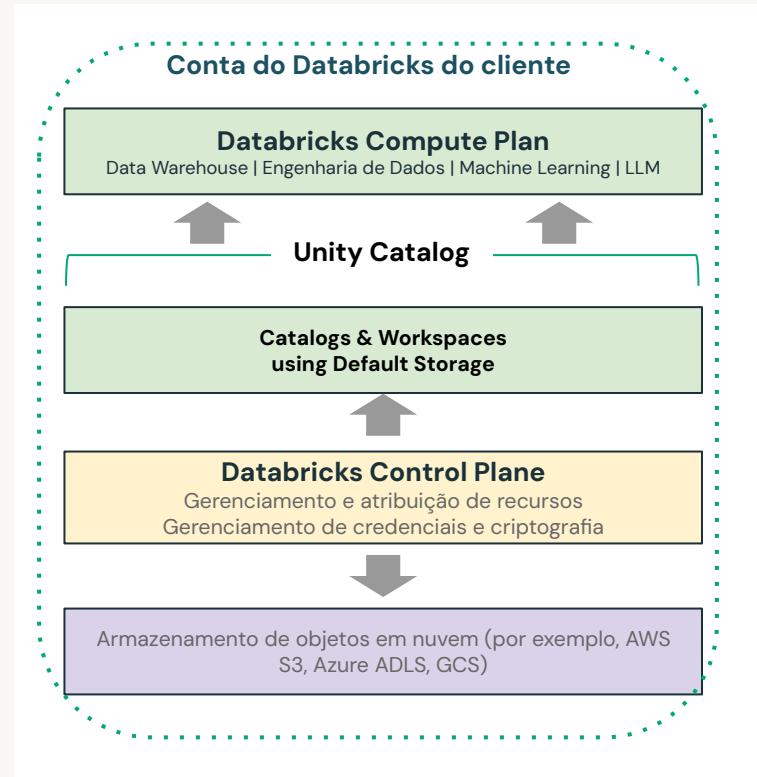
## Configuração com um clique para novos workspace e catálogos

Acesso unificado a dados, faturamento e observabilidade

Segurança pronta para uso, sem precisar configurar ou gerenciar o armazenamento em nuvem

Juntamente com a computação sem servidor, cria uma experiência de lakehouse SaaS

Private Preview



# Databricks Data Intelligence Platform

## Experiência do desenvolvedor

### Tópicos

---

Produtividade do desenvolvedor

Capacidades de produção

# Developer Experience: Upcoming Releases

## Databricks Assistant

-  Databricks Hosted Models (Public Preview)
-  Assistant Autocomplete (GA)
-  Optimize Query (Public Preview)
-  Usability Improvements: Quick Fix, inline chat history, diagnose job errors (GA)
-  Improved quality for Delta Live Tables code (GA)

## Intelligent Search and Discovery

-  Personalized Search Results (GA)
-  Blended search results across types (Public Preview)

## Databricks Notebook

-  Notebook Debugger on shared clusters (Public Preview)
-  Command Palette (GA)
-  Notebooks UX Enhancements – Improved Comments (GA), Result Table Improvements (GA), Tabbed UI (Public Preview)

## Custom Apps

-  Custom Apps (Public Preview)

## IDE Integrations

-  Pycharm Integration (GA)
-  DABS in Workspace and IDE (Public Preview)



# Otimizar consulta



Nova opção "Otimizar" do Assistente para otimizar e alertar sobre ineficiências em seu código SQL

As otimizações incluem chaves de partição recomendadas, cláusulas where em colunas particionadas, eficiências de cláusulas WHERE, eficiências ORDER BY, transformações recomendadas para alta cardinalidade GROUP BY, uso correto de tipo de dados e muito mais!

Disponível em Notebooks e Editor SQL

Public Preview

## Antes da otimização

```
1 %sql
2 -- Non-performant query example
3 SELECT *
4 FROM main.nyctaxi.trips
> See performance (1)
```

## Após a otimização

```
1 %sql
2 -- Optimized query example
3 SELECT tpep_pickup_datetime, tpep_dropoff_datetime, trip_distance, fare_amount, pickup_zip, dropoff_zip
4 FROM main.nyctaxi.trips
5 WHERE tpep_pickup_datetime >= date_sub(current_date(), 30) -- Optimization: Filter records from the last 30 days to reduce data volume
6 LIMIT 1000 -- Optimization: Limit the number of rows returned to avoid performance issues
```

# Assistente com modelos hospedados pelo Databricks

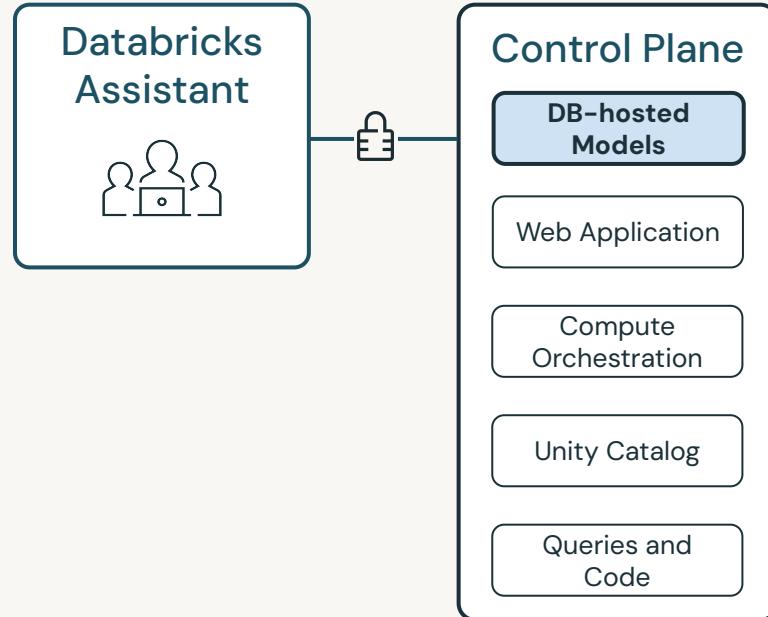


Use Databricks Assistant with Databricks-hosted and managed models.

Uses Databricks Model Serving to serve models. Eliminates cross-cloud traffic.

Models are stateless and don't log customer data. Customer data is protected by multiple layers of isolation.

Public Preview



# Assistant Quick Fix

Correções automatizadas de linha única  
para erros comuns em SQL e Python

Correção rápida de erros simples,  
incluindo erros de sintaxe, nomes de  
tabelas e colunas com erros  
ortográficos, importações ausentes,  
tipos de dados incompatíveis e  
correção de instruções GROUP BY.

Otimizado para velocidade. Sugestões  
em 1-3 segundos para que você não  
precise tirar as mãos do teclado.

Generally Available



A screenshot of a Jupyter Notebook interface. The cell contains the following SQL code:

```
%sql
SELECT * FROM users.samantha_banchik.nyctaxitrips
WHERE tpep_pickup_datetime >= 2016-02-01 AND tpep_pickup_datetime < '2016-02-14'
LIMIT 10
```

Below the code, there is a message: "[Shift+Enter] to run and move to next cell [Esc H] to see all keyboard shortcuts".

# Databricks Custom Apps



Aplicativos Python personalizados usando Gradio, Streamlit e Dash

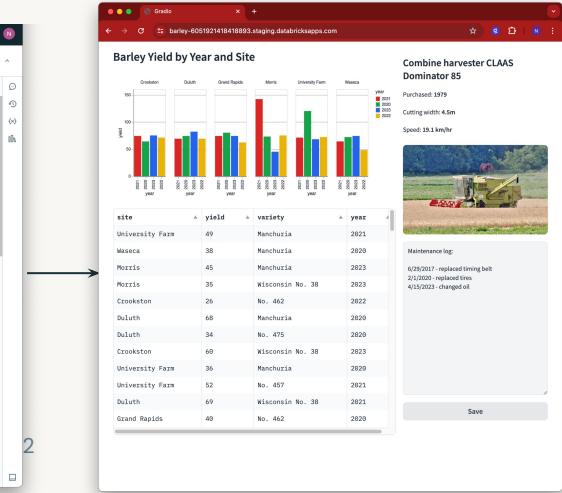
Criar soluções de dados para usuários não técnicos

Aproveite todo o poder da plataforma de inteligência de dados

Segurança e governança integradas

Compartilhe facilmente com usuários fora do Databricks

```
1 import gradio as gr
2 import numpy as np
3 import pandas as pd
4 import random
5
6 simple = pd.DataFrame([
7     {
8         "a": ["PA", "NY", "CT", "DE", "MD", "MD", "IL"],
9         "b": [28, 25, 43, 91, 81, 53, 19, 87, 52],
10    }
11 )
12
13 barley = pd.DataFrame([
14     {
15         "site": [
16             random.choice([
17                 "University Farm",
18                 "Waseca",
19                 "Morris",
20                 "Crookston",
21                 "Grand Rapids",
22                 "Duluth"
23             ])
24         ,
25         "yield": [
26             random.randint(25, 75) for _ in range(120)
27         ],
28         "variety": [
29             random.choice([
30                 "Manchuria",
31                 "Wisconsin No. 38",
32                 "Glabron",
33                 "No. 42",
34                 "No. 475",
35             ])
36         ,
37         "year": [
38             random.randint(1920, 2020) for _ in range(120)
39         ]
40     }
41     for _ in range(120)
42 )
43
44
45
46
47
48 ])
```



Public Preview em Q3



# Databricks Data Intelligence Platform

## Administração Empresarial

### Tópicos

---

Budget Alerts (Preview)

AI/BI Cost Dashboard (Preview)

System Tables (GA)

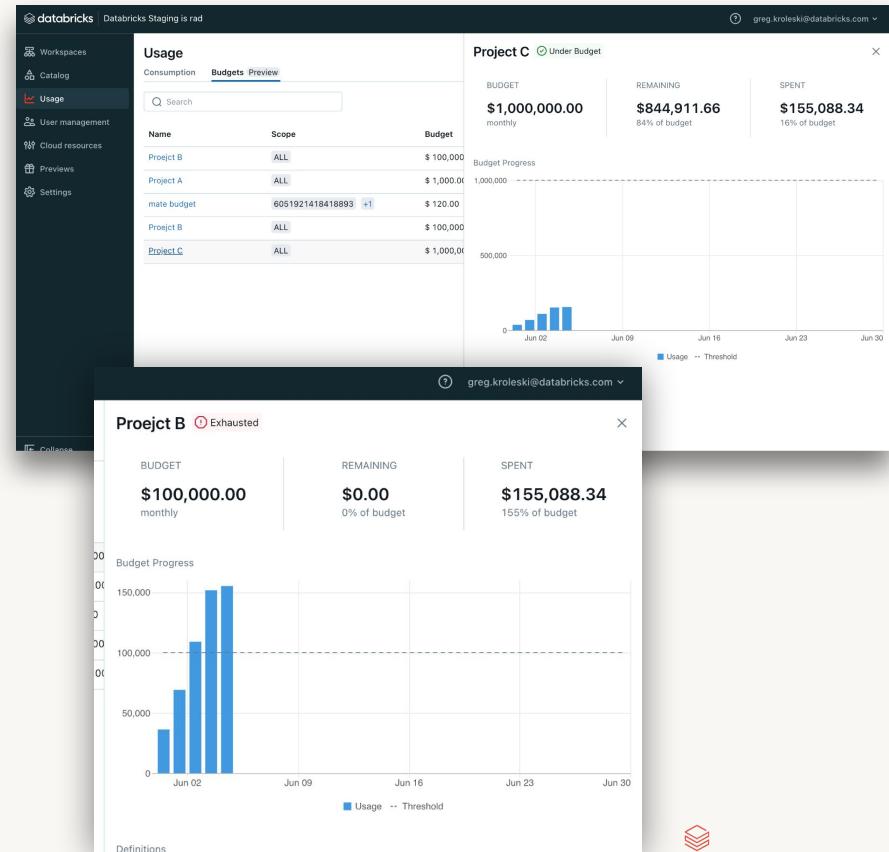
# Databricks Budget Alerts



Use tags para organizar os gastos e rastreá-los em relação a um orçamento

Definido no nível da conta ou para qualquer combinação de espaços de trabalho e/ou tags  
Receba alertas por e-mail quando os gastos ultrapassarem o limite

Interface de interface do usuário fácil e acesso à API



Public Preview em Q3



# AI/BI Cost Dashboard



Exibir tendências de uso em sua conta ou espaços de trabalho

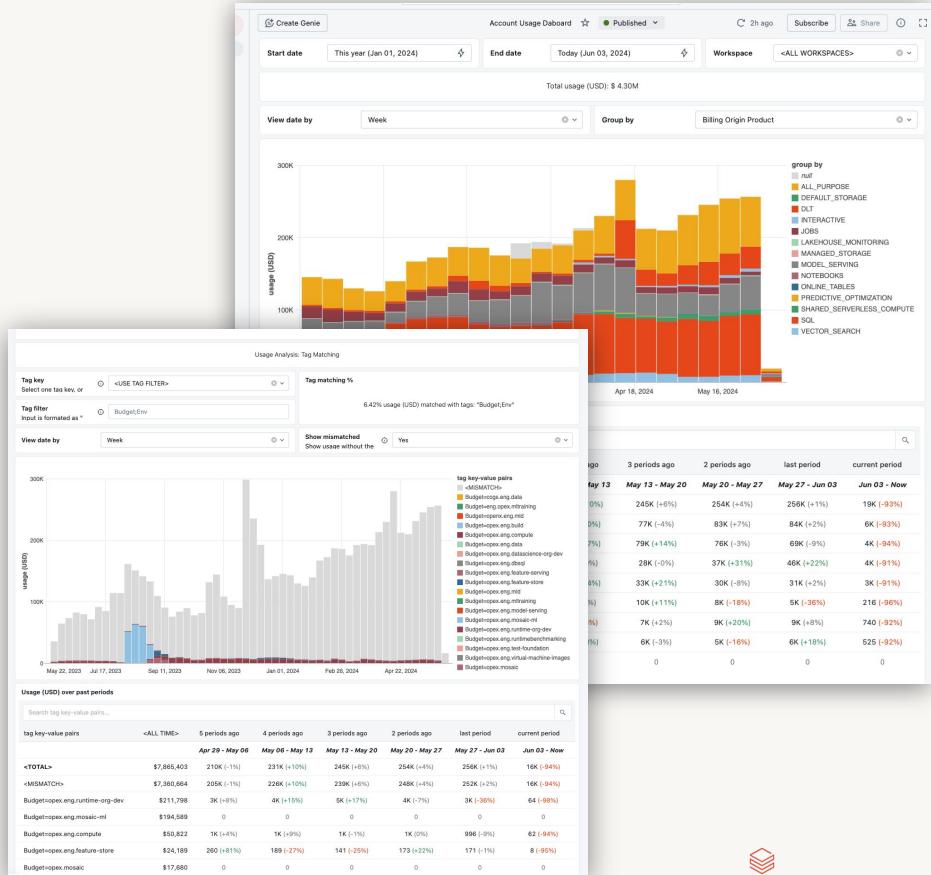
Filtros de interface do usuário fáceis de usar para detalhar por produto, espaço de trabalho e muito mais

Identifique rapidamente as cargas de trabalho, usuários, endpoints, etc. com os maiores gastos

Atribua custos usando tags e verifique se há integridade

Compartilhe com outras pessoas publicando-o com credenciais incorporadas

Public Preview in Q3



# System Tables em GA



Tabelas de faturamento (system.billing.usage e system.billing.list\_prices)

- Retenção de 365 dias incluída Grátis
- 365 dias extras à medida que desenvolvemos a retenção configurável

Generally Available

# Databricks Data Intelligence Platform

## Mosaic AI

### Tópicos

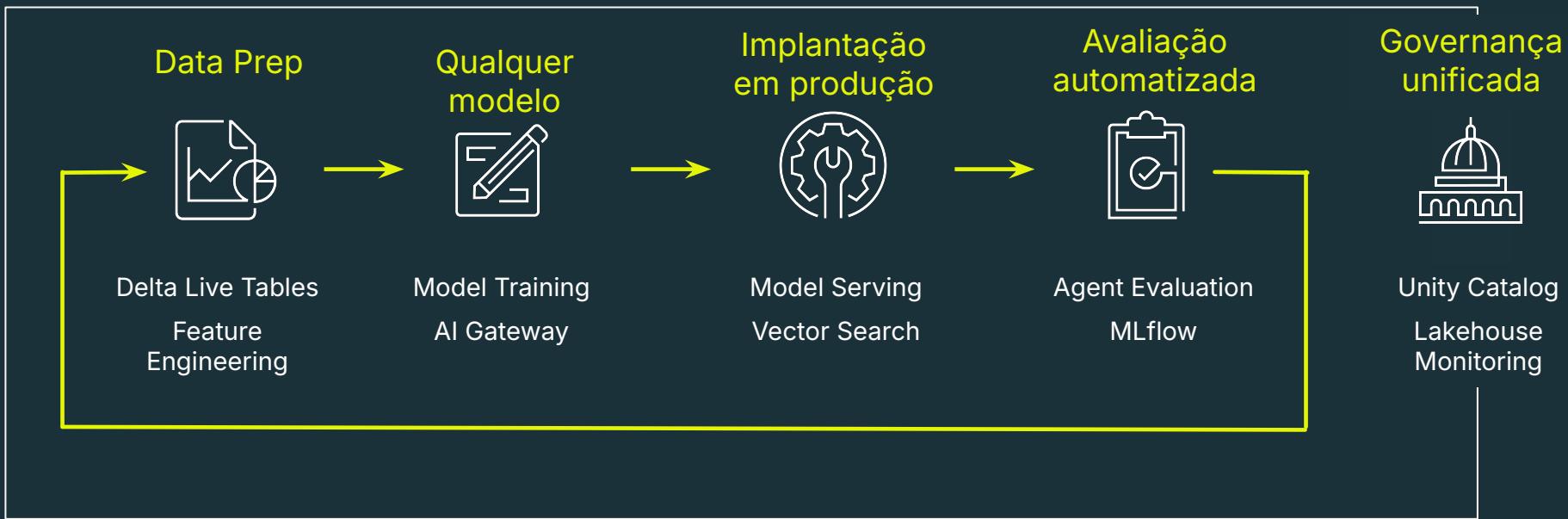
---

Meta Llama 3.1 em Model Serving  
AI Gateway

Gen AI Model Training



# Crie sistemas compostos de IA com o Mosaic AI



# Crie sistemas compostos de IA com o Mosaic AI

## ✓ Avaliação e operações automatizadas

Agent Evaluation | MLflow

## ✓ Personalização com dados corporativos

Model Training | Vector Search

## ✓ ✓ Quaisquer dados e qualquer modelo

AI Gateway | Model Serving

## ✓ Governança unificada

Unity Catalog | Lakehouse Monitoring



# Mosaic AI – Próximos lançamentos

## Generative AI



Model Training (Public Preview)



AI Gateway (Public Preview)



Agent Framework and Evaluation (Public Preview)

## Built in Governance for AI lifecycle



UC Tool Catalog Private Preview (Remote Function support)



UC Open Source Read/Write Models

## Model Serving



Meta Llama 3.1 (GA)



Batch Inference (Public Preview)



GCP Support (Public Preview)



Shutterstock ImageAI (Public Preview)



Foundation Model API (GA)

## MLOps & LLMOps



MLFlow 3.0



# Meta Llama 3.1 em Model Serving



## O novo padrão em modelos LLM OSS

Os modelos OSS alcançaram os modelos proprietários

O Mosaic AI pode servir e realizar fine-tuning com os modelos Meta Llama 3.1

Meta Llama 3.1 tem 3 variantes:

- 405B: ideal para casos de uso avançados que exigem raciocínio complexo
- 70B: destaca-se em chatbots, assistentes virtuais, fluxos de trabalho com agentes, geração de código
- 8B: excelente modelo pequeno que oferece respostas rápidas a um custo imbatível.

The screenshot shows the Databricks workspace interface. On the left is a sidebar with navigation links like 'New', 'Workspace', 'Recents', 'Catalog', 'Workflows', 'Compute', 'SQL', 'Data Engineering', 'Machine Learning', and 'Marketplace'. The main area is titled 'Welcome to Databricks' with a search bar. Below it, there are sections for 'Newly added' and 'Featured models'. Under 'Newly added', there are cards for 'DBRX Instruct' (Databricks), 'Llama 3 70B Chat' (Databricks), and 'Mistral 8x7B Instruct' (Databricks). Each card includes a brief description and a 'Try in Playground' button. Under 'Featured models', there are cards for 'BGE Large (En)' (Databricks), 'GPT 4' (OpenAI), 'GPT 3.5 Turbo' (OpenAI), 'PaLM 2 Chat Bison' (PaLM), 'Claude 3 (Anthropic)' (AWS Bedrock), and 'Jurassic-2 Ultra' (AI21 Labs). Each card also has a 'Try in Playground' button.

Disponível hoje



# AI Gateway

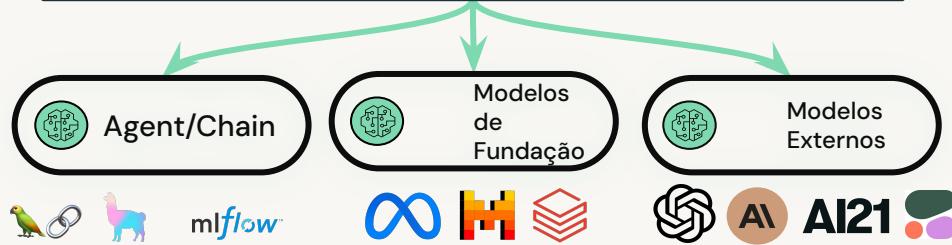
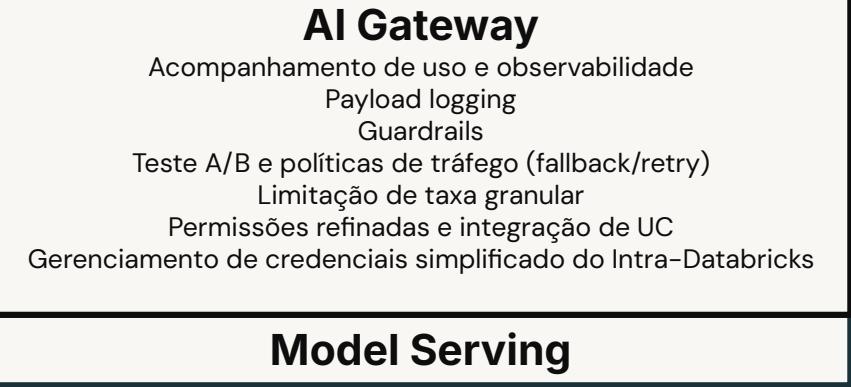


**Solução única para implantar, controlar e gerenciar todos os seus modelos de IA**

Permissão e limitação de taxa para controlar quem tem acesso e quanto acesso

Registro de carga útil e rastreamento de uso para auditar os dados que estão sendo enviados para APIs de modelo e quem as está chamando

AI Guardrails para evitar vazamento de dados e solicitações/respostas inseguras.



Public Preview em Q3

# Model Training

**Especialize modelos para tarefas específicas, contextos de negócios ou conhecimento de domínio para melhorar a qualidade**

Melhore os modelos de código aberto para ter um bom desempenho em tarefas corporativas especializadas para obter maior qualidade

Estenda a base de dados de conhecimento de um modelo adicionando documentos internos, idiomas ou eventos recentes atualizados

O modelo resultante é registrado no Catálogo do Unity e veiculado no Model Serving com um clique.



Experiments > Mosaic AI Model Training Preview

General

Task

- Chat Completion  
Finetune your model on chat logs between a user and an AI assistant
- Continued Pre-training  
Train your model with additional text data to add new knowledge to a model
- Instruction Finetuning  
Finetune your model on structured prompt-response data to adapt the model to a new task

Select Foundation Model  
Models trained in Model Training may be subject to license requirements and/or use policies. Learn more [↗](#)

Llama 3 8B Instruct

Training data

Training data\* ⓘ  
Select a dataset or enter a link to a Hugging Face dataset, for example databricks/databricks-dolly-15k

Browse training data [Browse](#)

Model registration

Register to location\* ⓘ  
Select the location in Unity Catalog where you want to register the trained model

Select a catalog [Choose schema](#)

Model name

ift-meta-llama-3-8b-instruct-ndc2mt

Public Preview

# Ofertas de parceiros para ajudar a implantar um RAG LLM POC em menos de 4 semanas

accenture



Aimpoint  
Digital



CELEBAL  
TECHNOLOGIES

Computomic

Deloitte.

ENTRADA

Infosys<sup>®</sup>

KOANTEK



Lovelytics.

slalom

Entre em contato com seu **executivo de contas**; informe-o de que você está interessado em "[Parceiro]'s RAG LLM POC"

O investimento começa em US\$ 35 mil para um engajamento de 4 semanas



# Data Intelligence Platform

Unity Catalog

Delta Lake UniForm



# databricks

# APPENDIX