



Generative AI with Databricks

From Concept to Deployment



Today's Speakers



JK Wong
Sr. Partner Solutions
Architect

jk.wong@databricks.com

 **databricks**



Andrew Kennedy
Sr. Data & AI
Consultant

andrew.kennedy@thorogood.com

THOROGOOD®

Agenda



- Who is Databricks?
- What is Generative AI and the impact on Global Economy
- Databricks Data Intelligent Platform



- Use Case References
 - Knowledge discovery and search (Demo)



10,000+
global customers

\$1.5B+
in revenue

\$4B
in investment

Inventor of the **lakehouse**
&
Pioneer of **generative AI**



databricks

The data and AI company

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

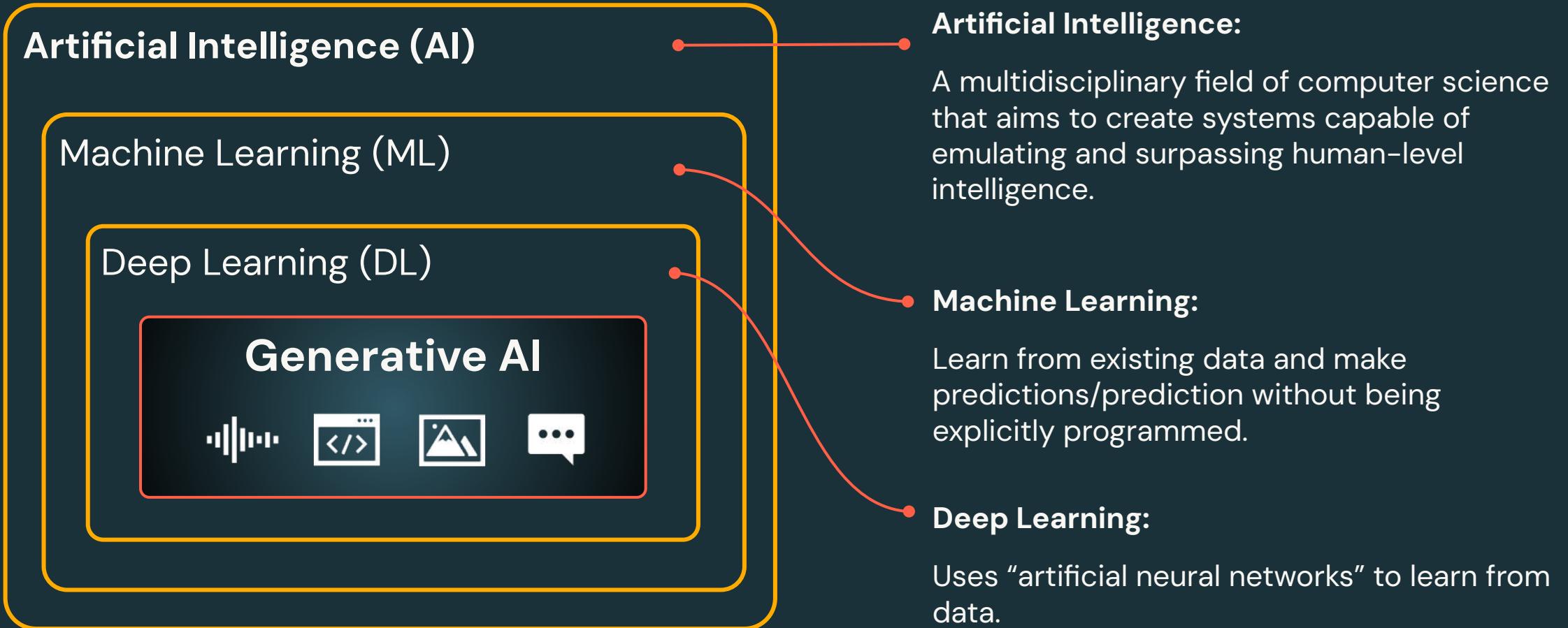
Creator of



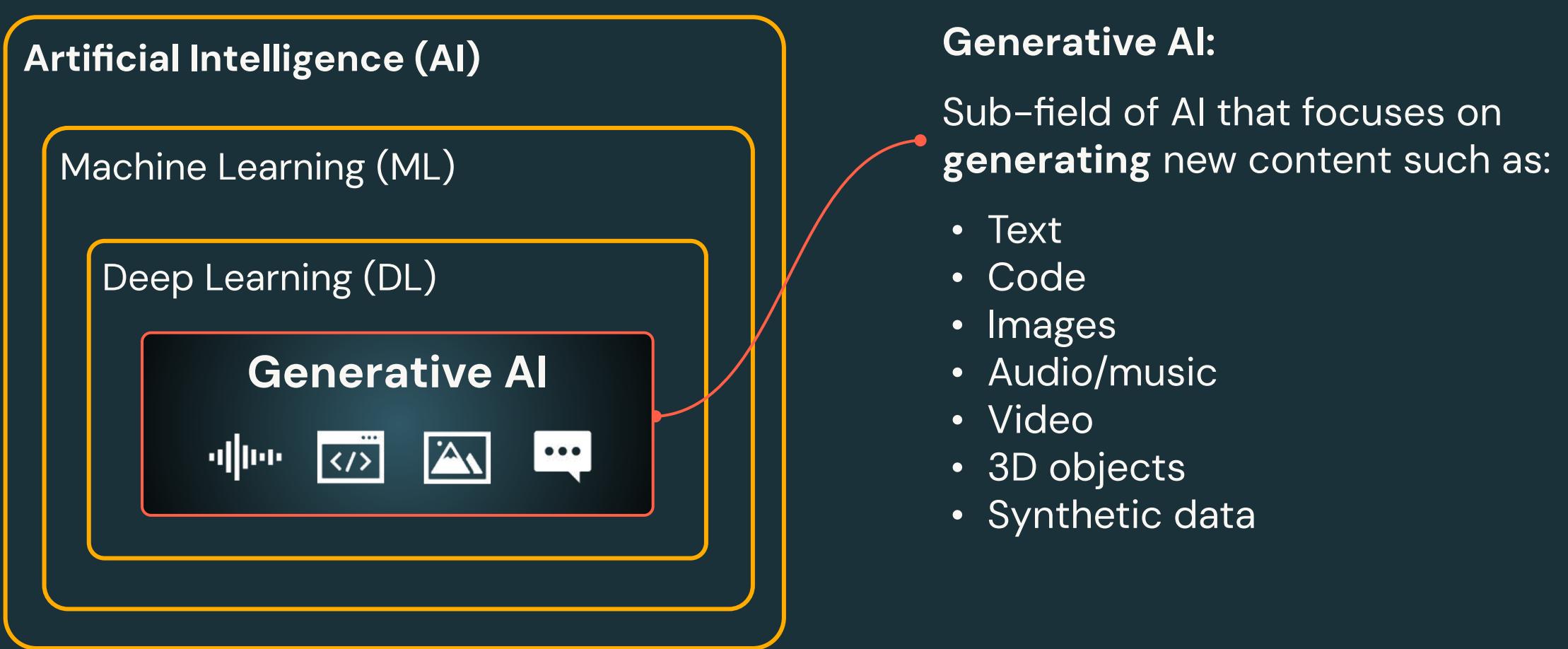
mlflowTM



What is Generative AI?



What is Generative AI?



Generative AI use cases

- Question and answers
- Content generation
- Content personalization
- Translation
- Language style transfer
- Story telling, creative writing
- Code generation, auto-completion

Translate English to Spanish:
What's the time?

What is the expected weather in Puerto Rico in December?

Summarize in 15 words
"During extra time, Messi then scored again to give Argentina a 3-2 lead. However, Mbappé scored another penalty to tie the game 3-3 with only minutes remaining, becoming the second man to score a hat-trick in a World Cup final. Argentina then won the ensuing penalty shoot-out 4-2 to win their third World Cup."

Generate a tagline for a fruit juice brand that focused on sustainable and organic farming.

LLM

Qué hora es?

In December, Puerto Rico experiences pleasant tropical weather with an average high temperature of 84°F (29°C) and an average low temperature of around 72°F (22°C).

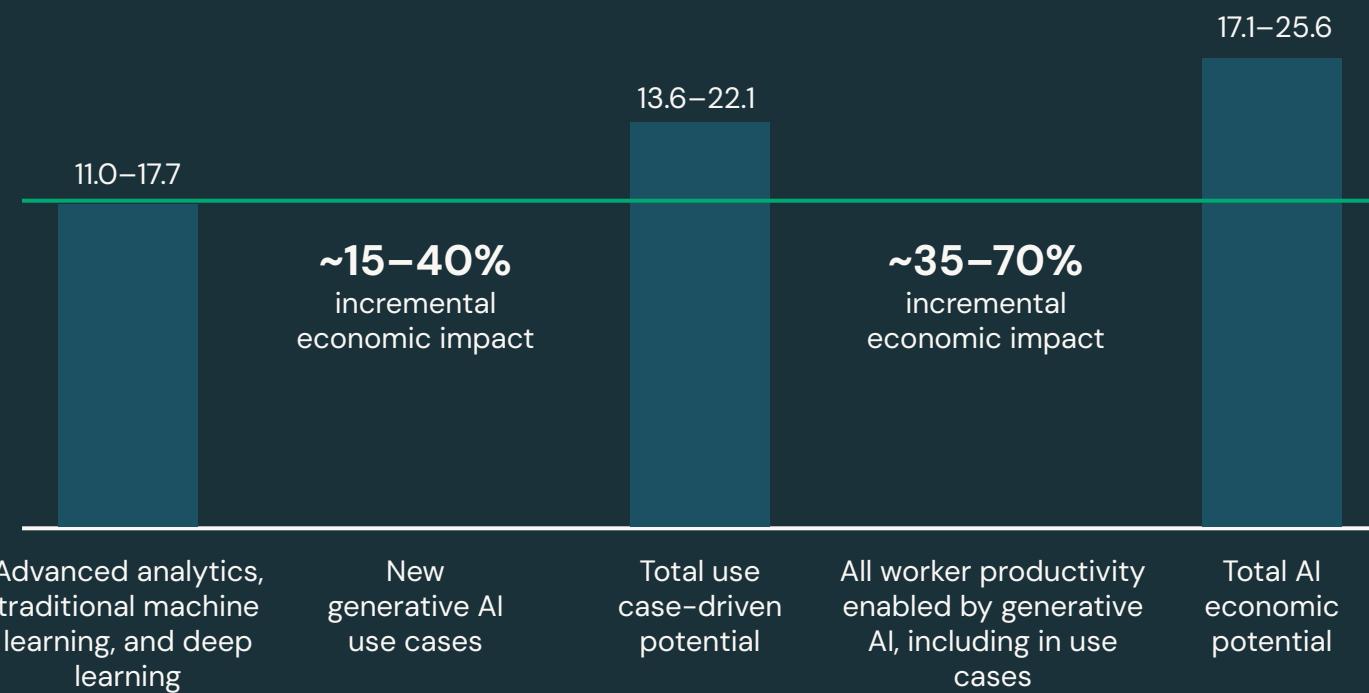
Mbappé scores a hat-trick but Argentina wins World Cup in a penalty shoot-out.

Purely natural, sustainably sourced. Taste the goodness of organic fruits with every sip.

Generative AI impact on global economy

Generative AI is expected to create significant value over and above what traditional AI and Analytics are expected to deliver

AI's potential impact on the global economy, \$trillion



McKinsey – [The economic potential of generative AI: The next productivity frontier](#)



The winners in every industry will be
data + AI companies



Gen AI will disrupt use cases across industries



Financial Services

Fraud monitoring and predictions



Healthcare & Life Sciences

Biomedical literature summarization & discovery



Comm, Media & Entertainment

Hyper-personalization for customer experience (CX)



Retail & Consumer Goods

Try before you buy with virtual fitting rooms



Manufacturing

Delightful, personalized customer experiences



Public Sector

Analysis of open-source Intelligence

Automating compliance data gathering

Clinical trial optimization

Enhancing customer support and self-service

Optimizing demand prediction and inventory

Increasing productivity and efficiency in operations

Modernizing legacy code bases

Accelerate underwriting and claims processing in insurance

Health insurance claim processing

Intelligent content creation and curation

Generate innovative product designs

Prescriptive and proactive field service

Regulatory compliance assistance



Challenge

The Rangers wanted to unlock the potential of massive, diverse data sets and combine it with the power of AI to help deliver best-in-league player performance and amazing fan experiences

Solution

Databricks unifies data from different sources—biomechanics, weather, consumer—and delivers AI-driven analytics that help boost player pitch and hit rates and prevent injuries. LLMs are in beta to streamline player ops, and personalize in-stadium experiences with chatbots

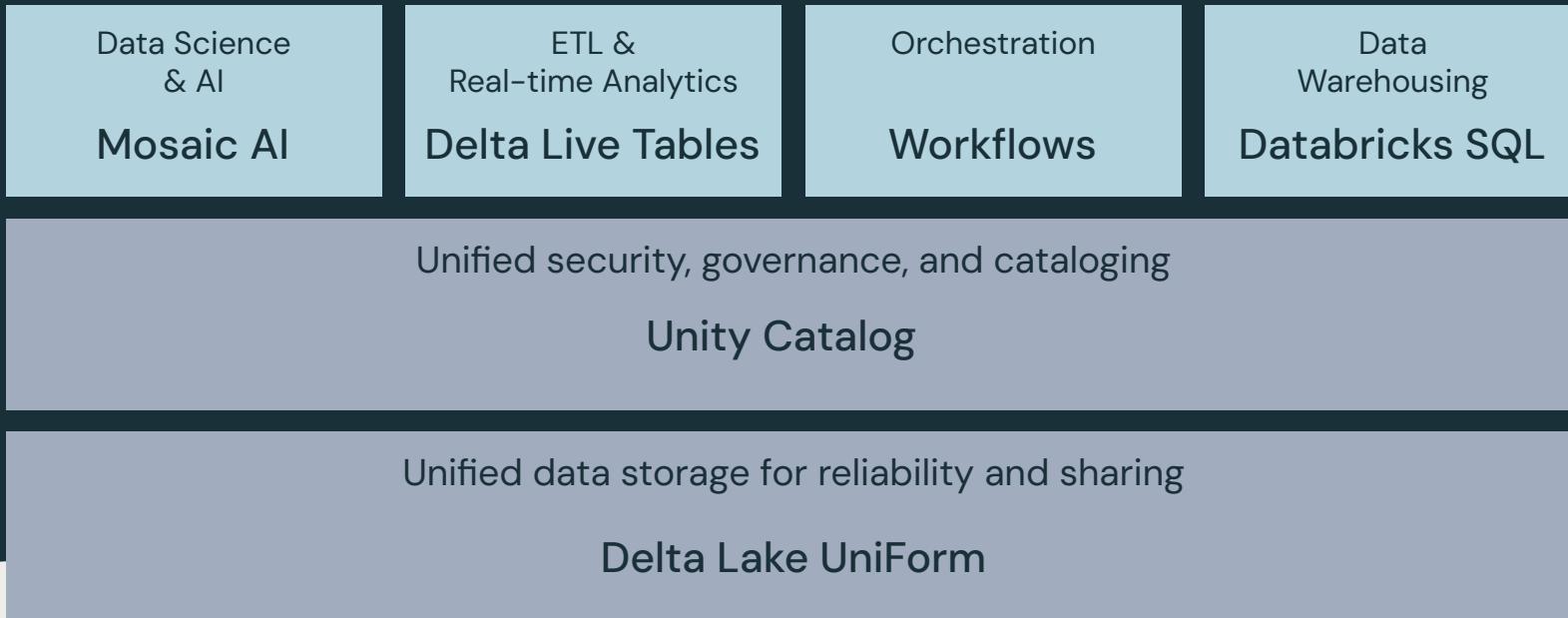
Impact

1st

ever World Series Championship

The Data Lakehouse

An open, unified foundation for all your data

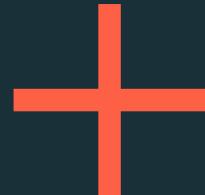


2020
Databricks pioneered
the lakehouse
architecture

Today
74% of global
enterprises have
adopted lakehouse
MIT Technology Review
Insights, 2023

Data Lakehouse

An open, unified foundation
for all your data



Generative AI

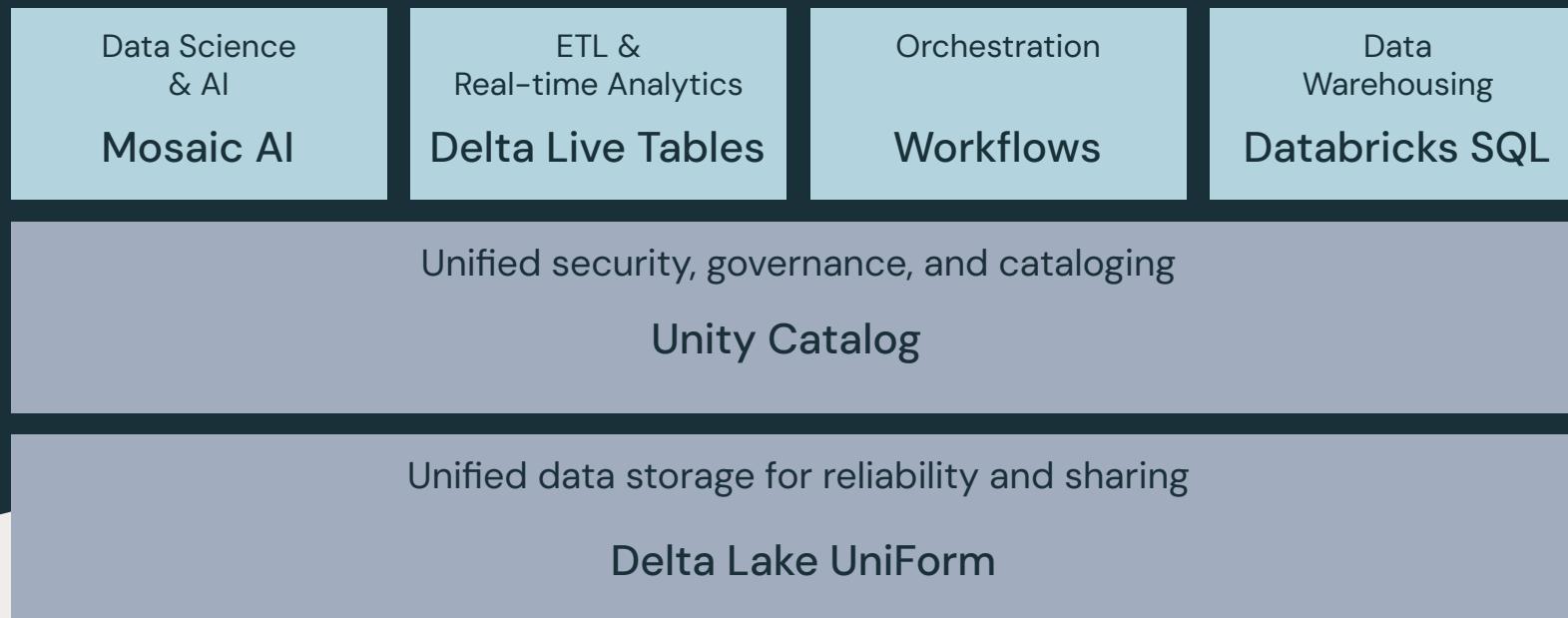
Easily scale and use data and AI



Data Intelligence Platform

Democratize data + AI across
your entire organization

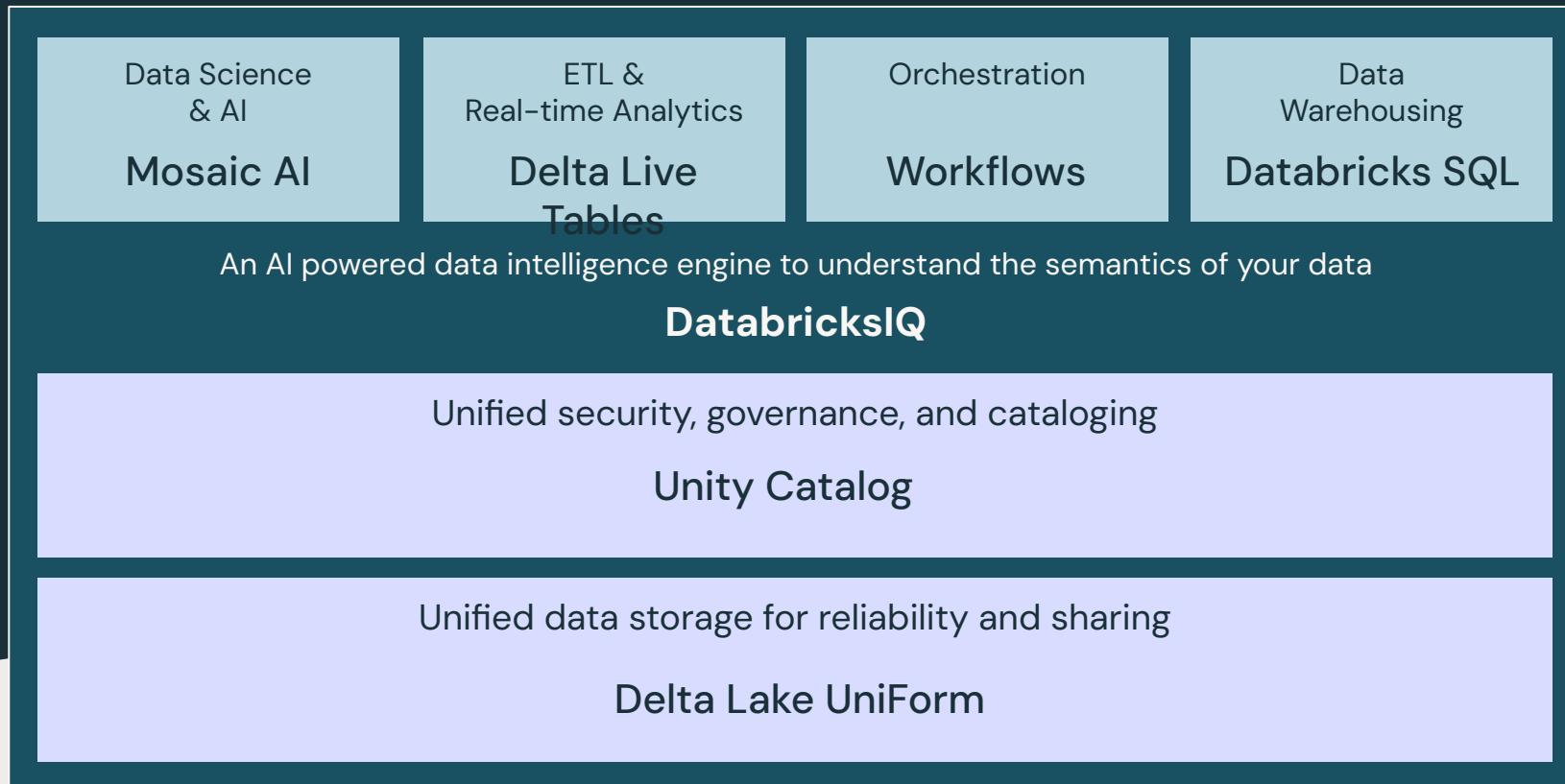
Databricks Data Intelligence Platform



Open Data Lake

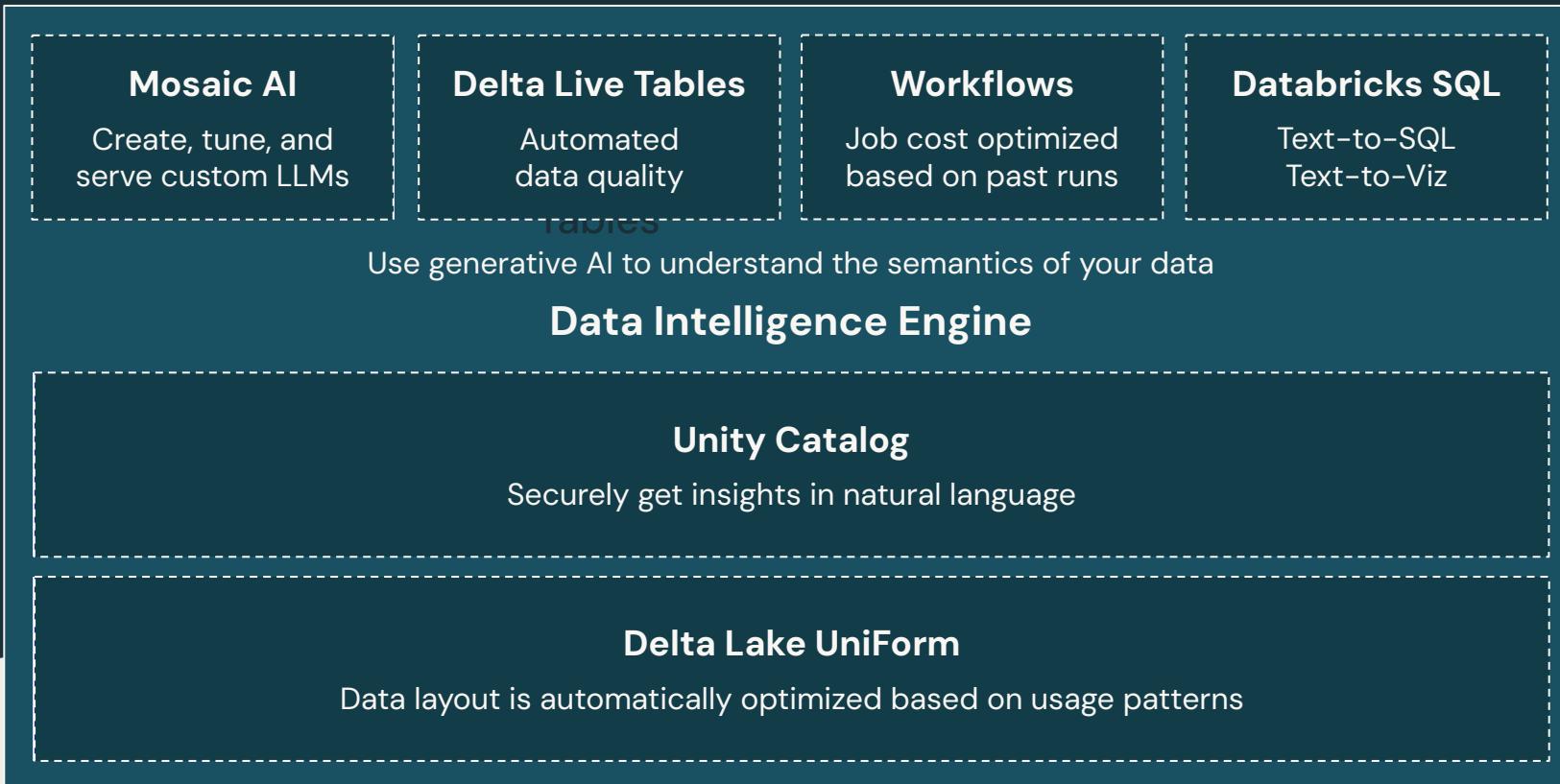
All Raw Data
(Logs, Texts, Audio, Video, Images)

Databricks Data Intelligence Platform



Open Data Lake
All Raw Data
(Logs, Texts, Audio, Video, Images)

Databricks Data Intelligence Platform



Open Data Lake

All Raw Data
(Logs, Texts, Audio, Video, Images)

Databricks Data Intelligence Platform

Mosaic AI

Gen AI

- Custom models
- Model serving
- RAG

End-to-end AI

- MLOps/MLflow
- AutoML
- Monitoring
- Governance

Mosaic AI

Create, tune, and serve custom LLMs

Delta Live Tables

Automated data quality

Workflows

Job cost optimized based on past runs

Databricks SQL

Text-to-SQL
Text-to-Viz

Use generative AI to understand the semantics of your data

Data Intelligence Engine

Unity Catalog

Securely get insights in natural language

Delta Lake UniForm

Data layout is automatically optimized based on usage patterns

Open Data Lake

All Raw Data
(Logs, Texts, Audio, Video, Images)

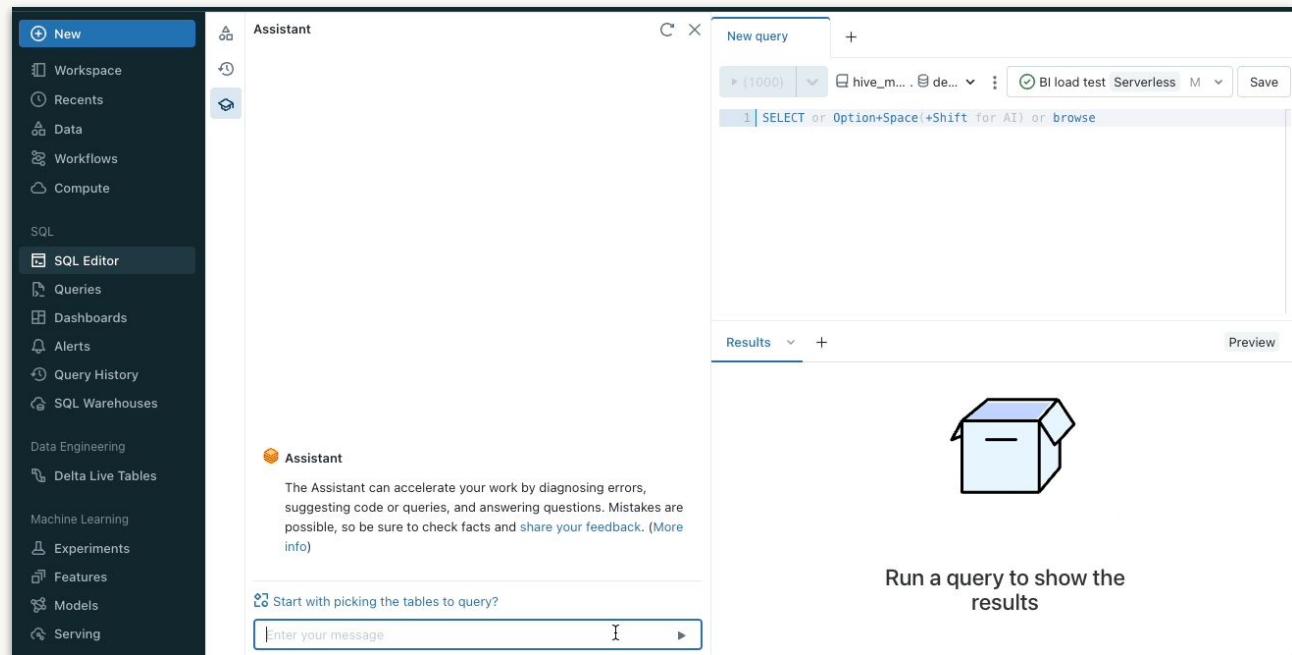
GenAI Features

Databricks, powered by AI

Example: Databricks Assistant

AI assistant with contextual understanding of your data—natively within Notebook, SQL editor and file editor.

- Generates and auto-completes code and queries
- Explains and fixes issues
- Integrates with Unity Catalog, offering contextual results relevant to your data assets



Models in Marketplace

Curated models – backed by optimized Model Serving



Simple: Grab the latest open source models in just a few clicks, without writing boilerplate code

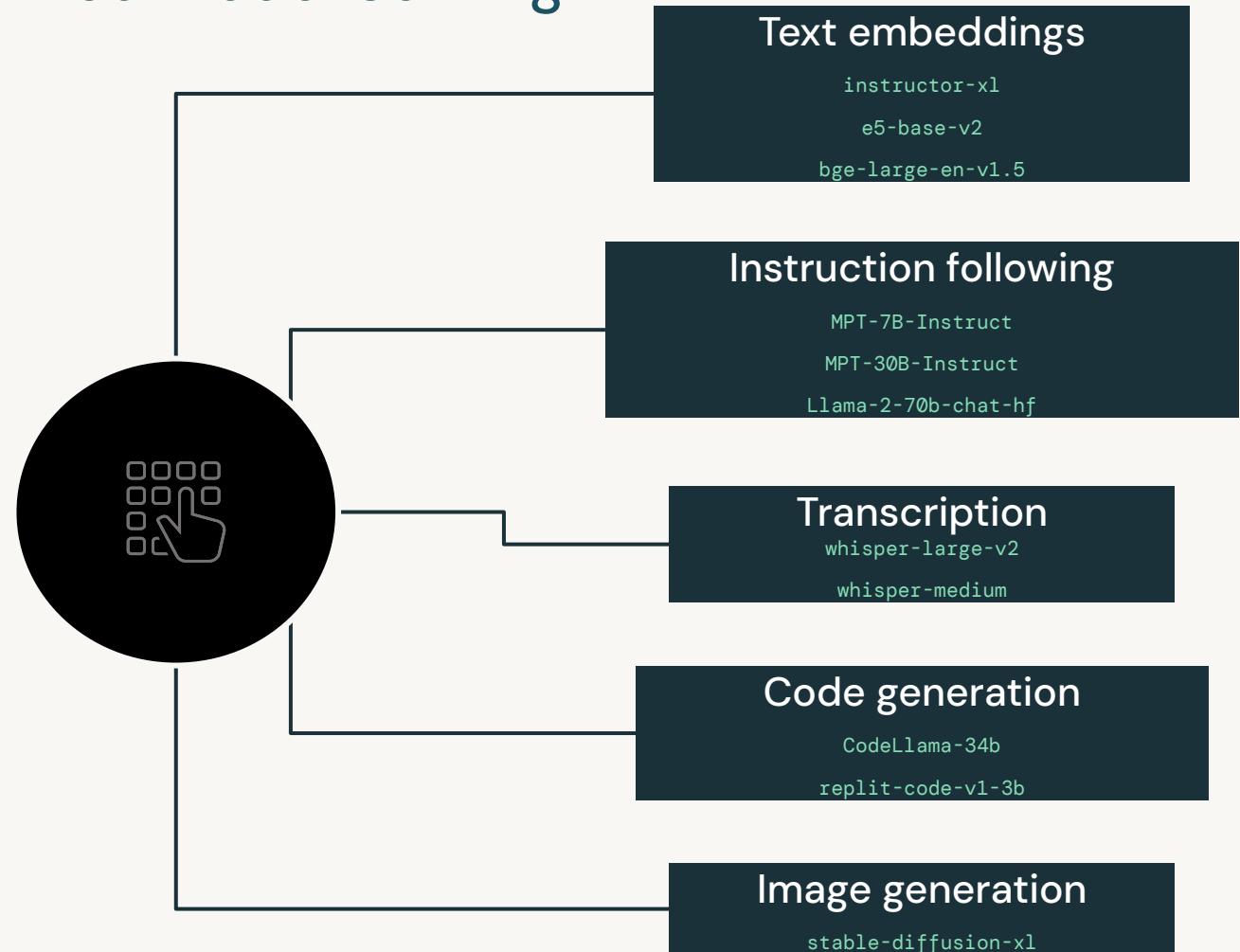


Optimized: Deploy to Model Serving and scale in a cost-effective and performant way, with automatic optimizations for common model architectures



Governed: Ensure compliance and governance by managing models from the marketplace in Unity Catalog

OSS model guidance: Research team findings are also [published here](#).



Lakehouse Monitoring

Unified monitoring for reliable, insightful, and simple data-to-AI pipelines



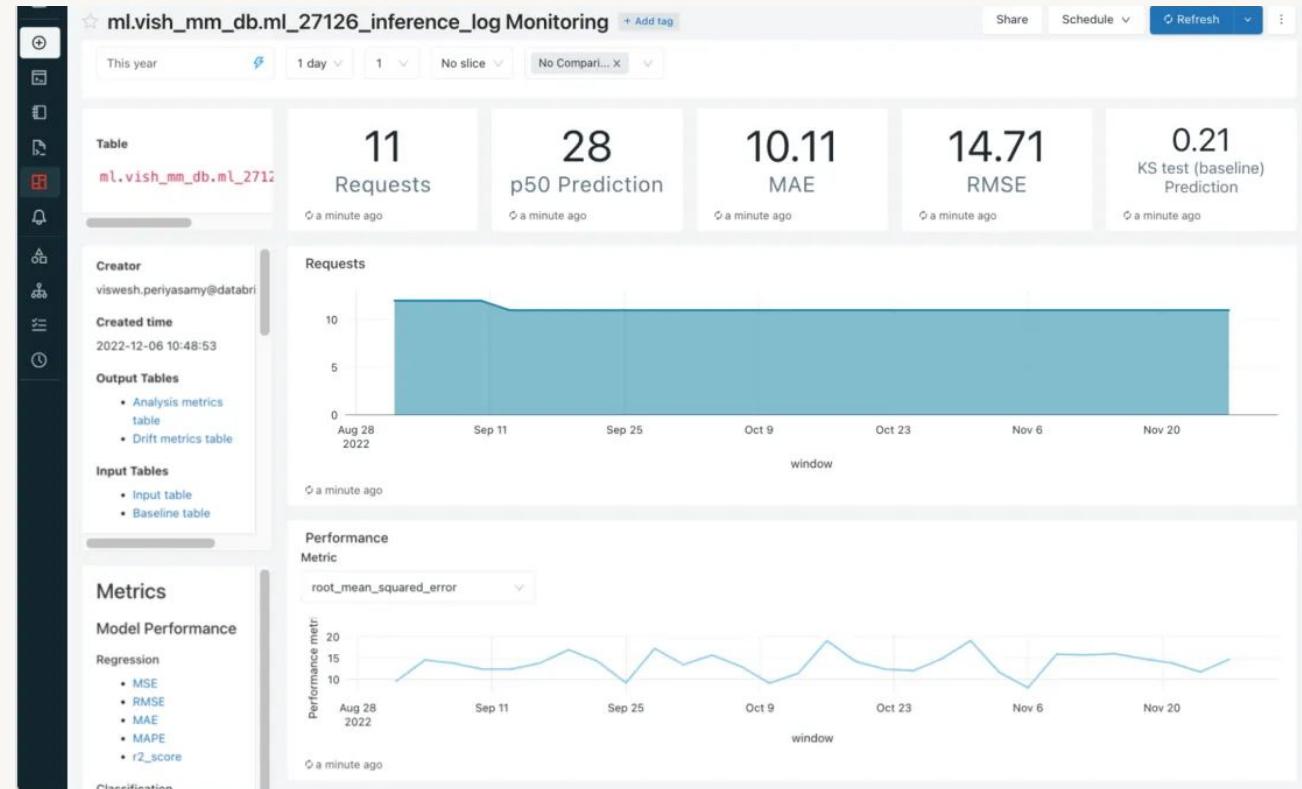
Simple: Log inference tables automatically, and generate metric tables and SQL dashboards.



Proactive: Automate alerts on table quality and custom metrics, and diagnose data or model issues.



Integrated: Track end-to-end lineage in the Unity Catalog for training data, feature tables, models, and inference logs, for simpler governance.



MLflow Deployments Server

Manage, govern, evaluate, and switch models easily. Previously called “AI Gateway”

All of Databricks Model Serving will provide these benefits.

Central Management and Governance



Routes let users query designated models using credentials managed by their organization.

Organizations can **rate-limit** routes to control costs.

Routes enable **logging** and **caching** of requests for better tracking and observability.

Unified APIs for AI Models and Providers



A **common API** for diverse models and providers means users don't have to familiarize themselves with vendor specific APIs and documentation.

Easily **upgrade to the latest and greatest LLM** without restructuring the rest of your code.

Enable Multiple Gen AI Use Cases



With simplified model access, **developers can focus on the end product** instead of infrastructure.

A common API helps future proof by providing the **flexibility to use the best models** as they are released.

Cost control and operational oversight lets you **responsibly scale AI and meet audit and compliance requirements**.

Vector Search

Create auto-updating vector indexes, managed by Unity Catalog

Choose your source table



id	text	col1	col2
1	The quick brown fox jumps ...		
2	How quickly daft...		
3	The five boxing wizards...		

Call a simple API to create a semantic search index

```
client.create_delta_sync_index(  
    endpoint_name="storage_endpoint",  
    index_name=f"{catalog}.{schema}.{index}",  
    source_table_name="ml.llm.spark_docs",  
    primary_key="id",  
    embedding_source_column="text",  
    ai_gateway_route_name="openai-embedding",  
    pipeline_type="CONTINUOUS")
```

Choose any embedding model

MLflow AI Gateway



- Ingestion pipelines managed for you
- Indexes managed by Unity Catalog
- Also, APIs for
 - Unmanaged embeddings
 - CRUD API upsert/delete

Call endpoint for real-time retrieval

```
result = index.similarity_search(  
    query_text="What is Spark Connect?",  
    columns=[ "id", "text", "link"],  
    filters={"doctype": "wiki"})
```

- Integrate with LangChain, LlamalIndex, etc.
- Scale out endpoints as needed

Documentation: [AWS](#), [Azure](#)



A Data Intelligence Platform enables you to truly democratize data and AI

➤ ACCESSIBLE WITH NATURAL LANGUAGE

Everyone can quickly and easily discover insights

➤ SIMPLE TO OPERATE

AI removes complexity and heavy lifting

➤ STRONG GOVERNANCE AND SECURITY

Train and serve custom GenAI applications that maintain privacy



Use Cases References

THOROGOOD®

Data & AI: The Right Way

Introduction



Andrew Kennedy

andrew.kennedy@thorogood.com
linkedin.com/in/aikennedy

- Chartered Statistician, Royal Statistical Society
- Databricks Certified ML Professional
- Microsoft Certified Azure AI Engineer Associate
- AWS Certified ML Specialist
- Google Cloud Certified Professional ML Engineer

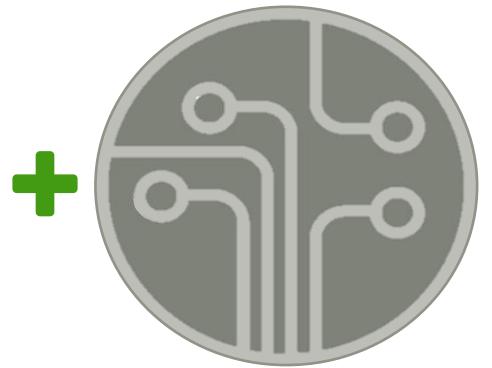
Thorogood overview

Who we are



analytics

data, modeling and
algorithm expertise
outclassing intuition



technology

understanding
possibilities to
maximize opportunities



business focus

locking onto
business goals to
deliver business
value

Data & AI Specialists

Independent, customer-oriented consultancy

Operating Globally

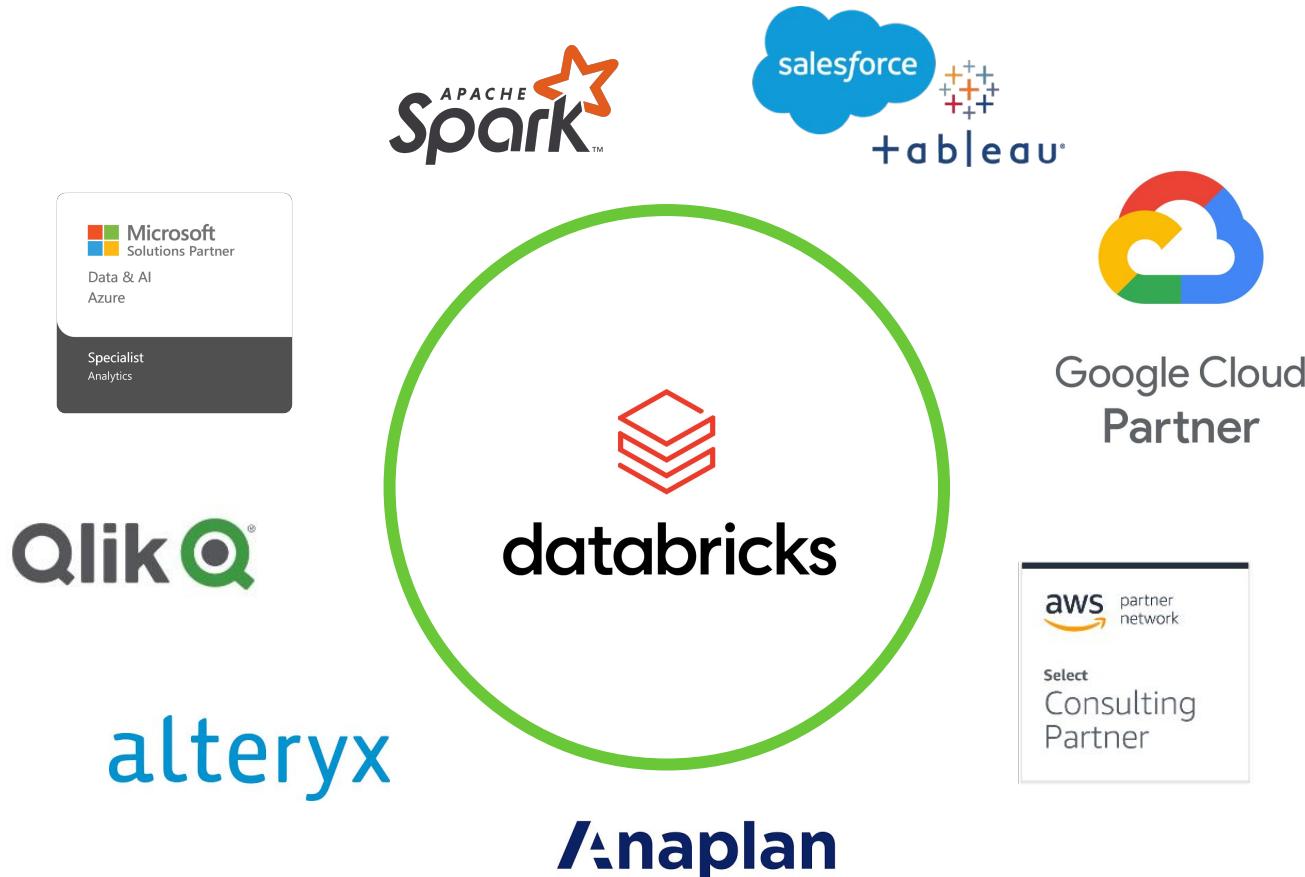
London • Singapore • São Paulo
Philadelphia • Boston • Bangalore

Strategies • Solutions • Services

Application Design & Development
Planning & Business Modeling
Strategies & Roadmaps
Data & Analytics Services
Training & Adoption
Maintenance & Support

About Thorogood

Some of the technologies we work with

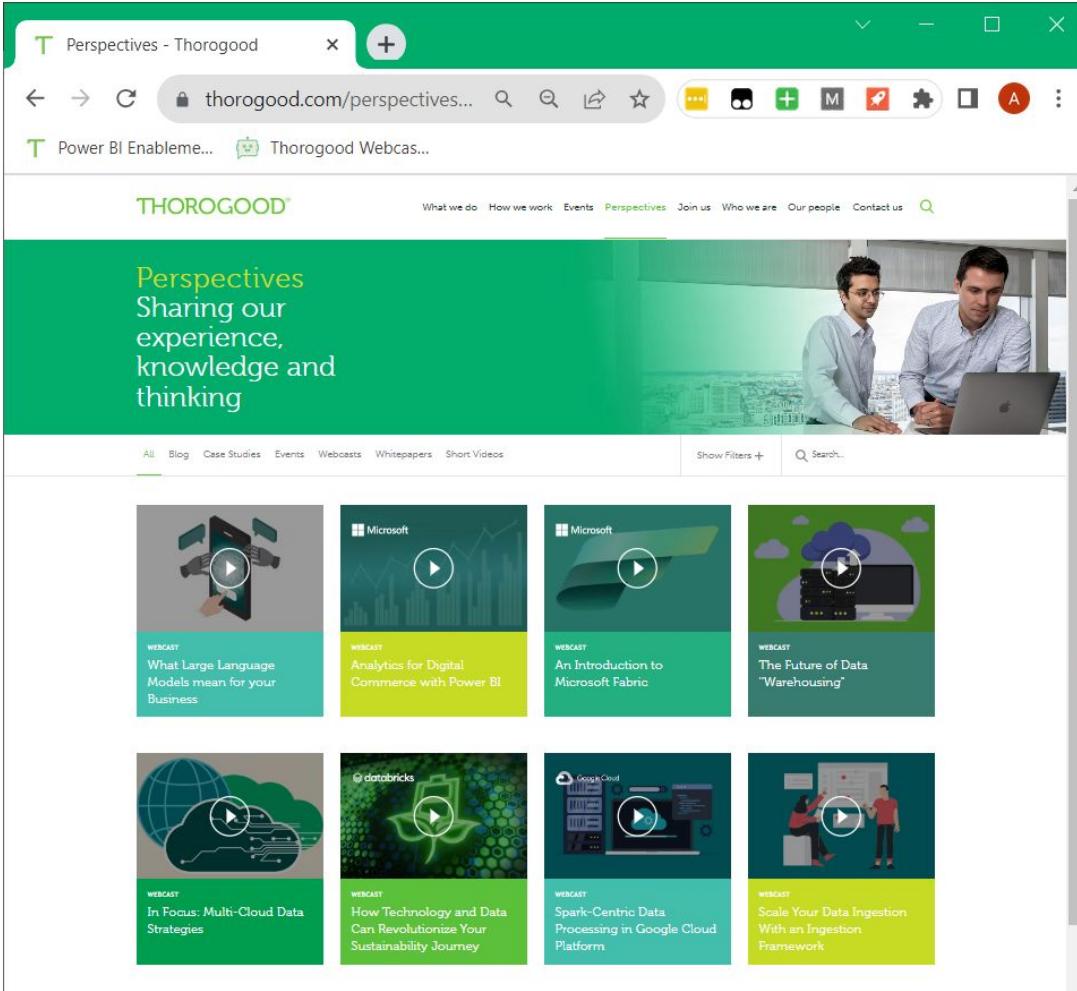


Making your content searchable

thorogood.com

Knowledge discovery and search

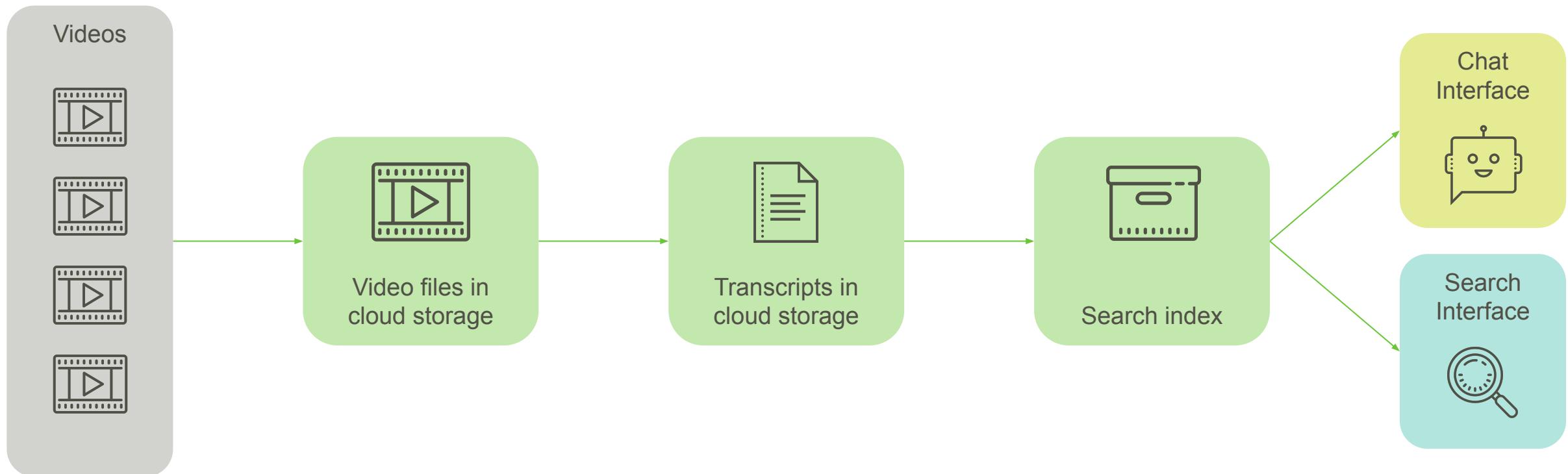
Making content searchable from Thorogood's webcast videos



How can we most easily find relevant content from videos?

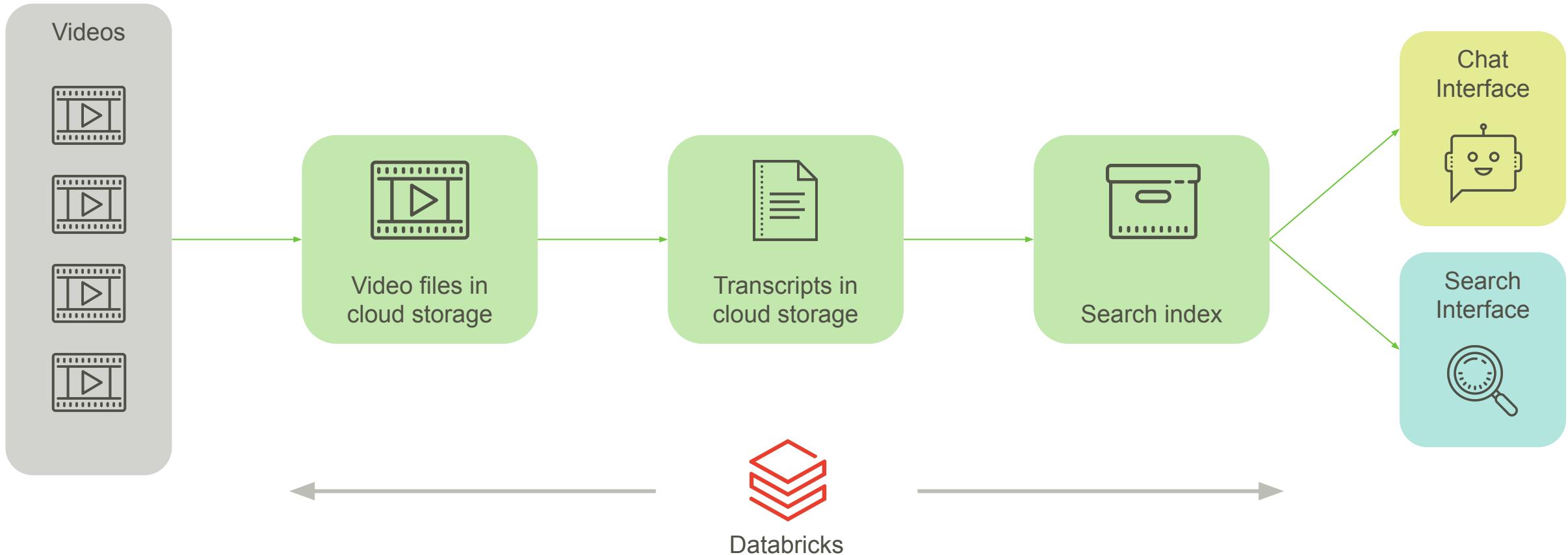
Knowledge discovery and search

Making content searchable from Thorogood's webcast videos



Knowledge discovery and search

Making content searchable from Thorogood's webcast videos

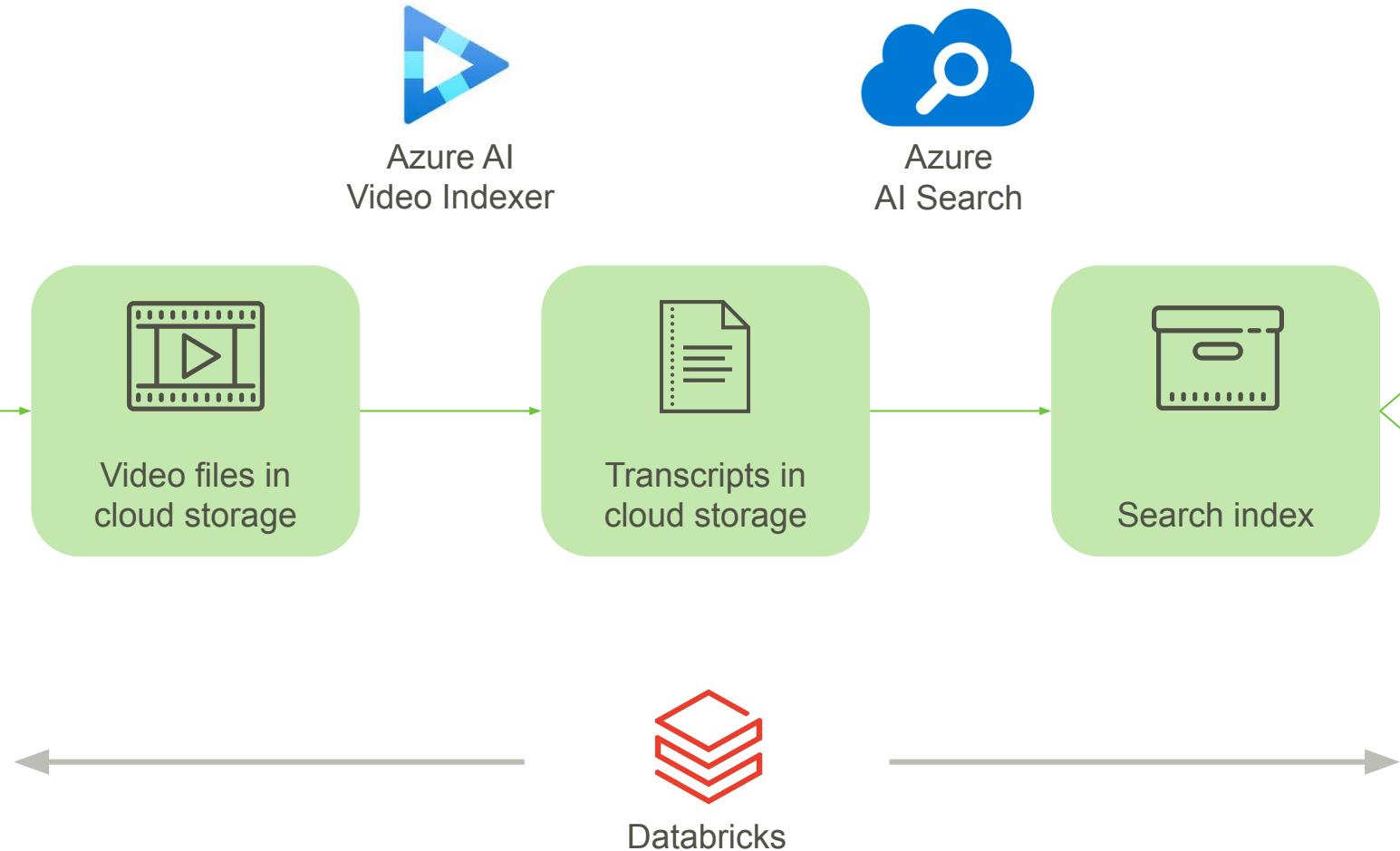
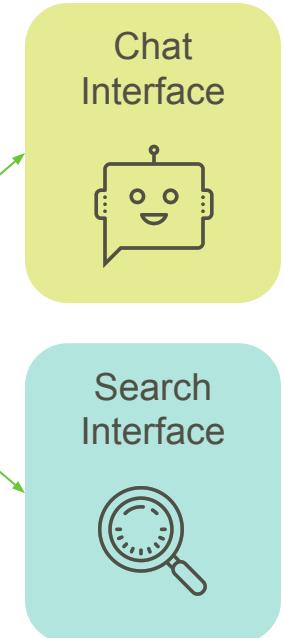
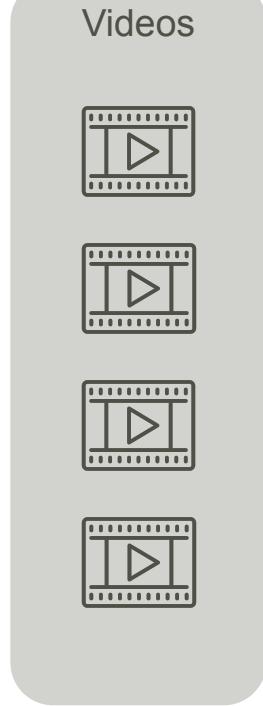


Knowledge discovery and search

Making content searchable from Thorogood's webcast videos

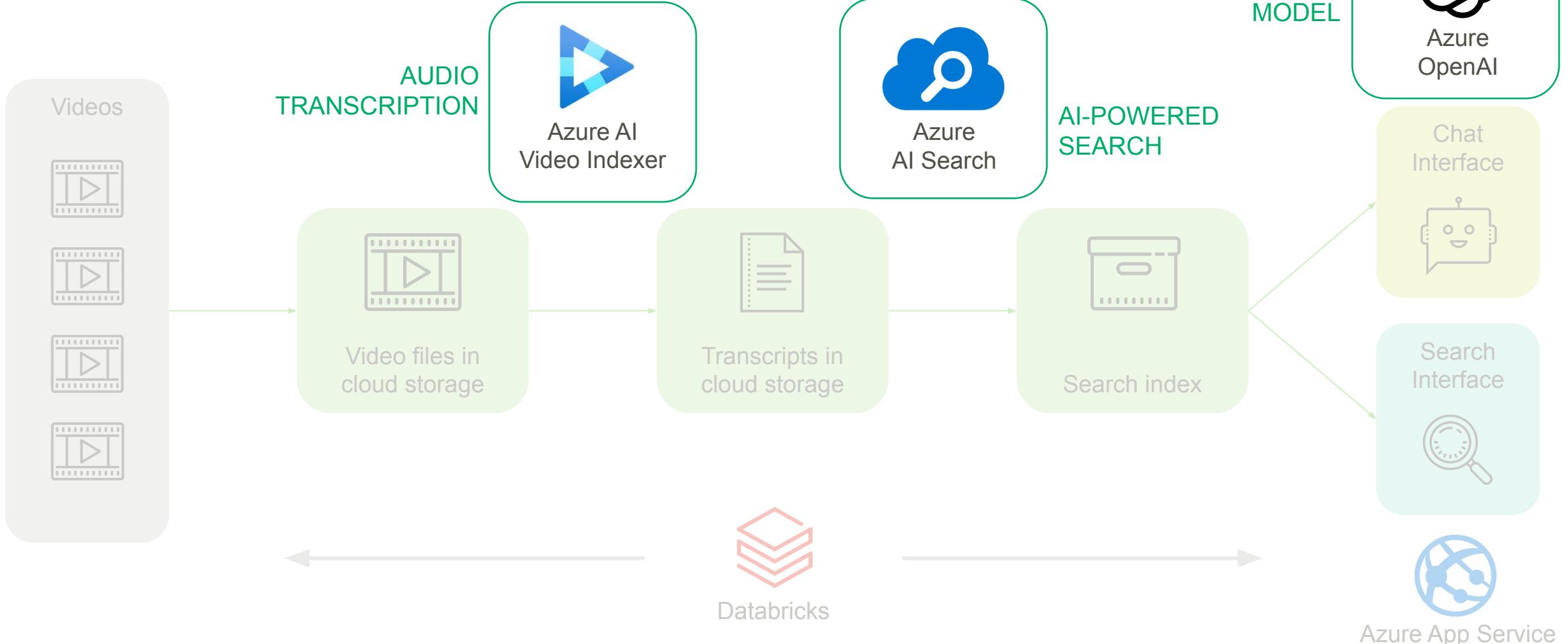


Azure
OpenAI



Knowledge discovery and search

Making content searchable from Thorogood's webcast videos



Demo

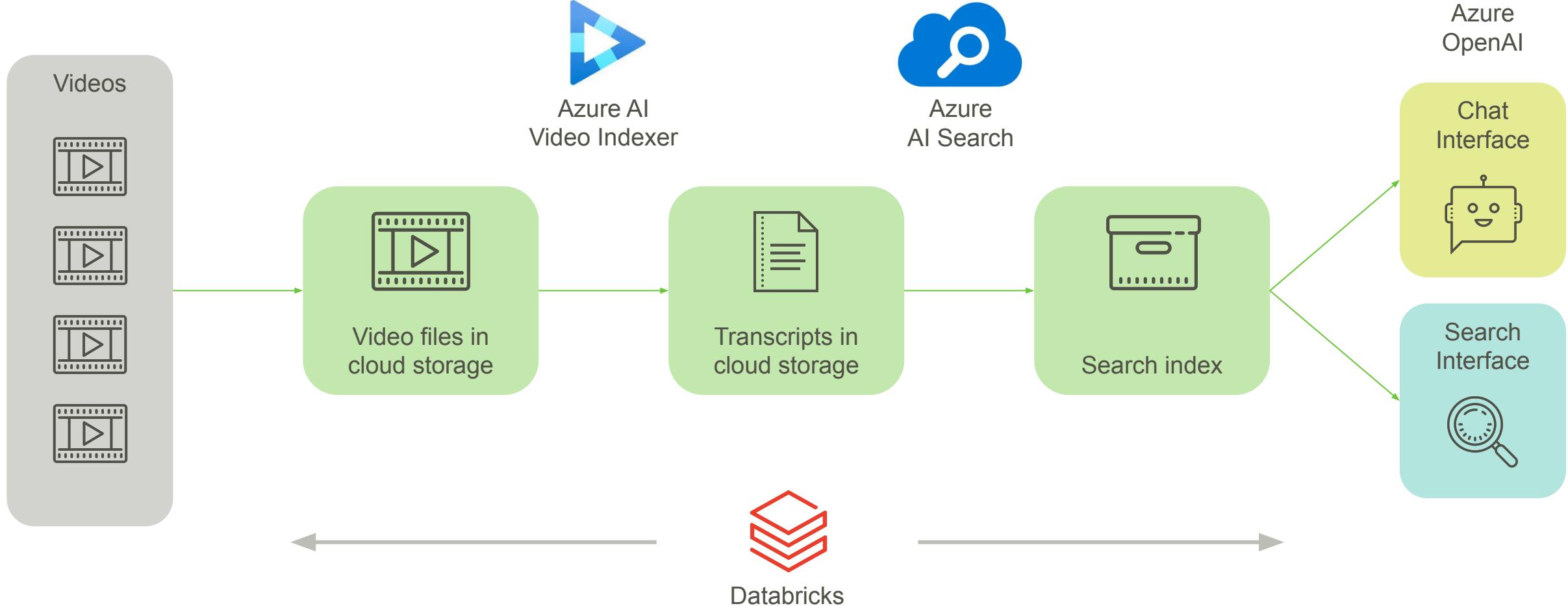
thorogood.com

Knowledge discovery and search

Making content searchable from Thorogood's webcast videos



Azure
OpenAI



Customer example

Making your content searchable

thorogood.com

Indexing and Querying Quarterly Reports

The Business Goals



Goals

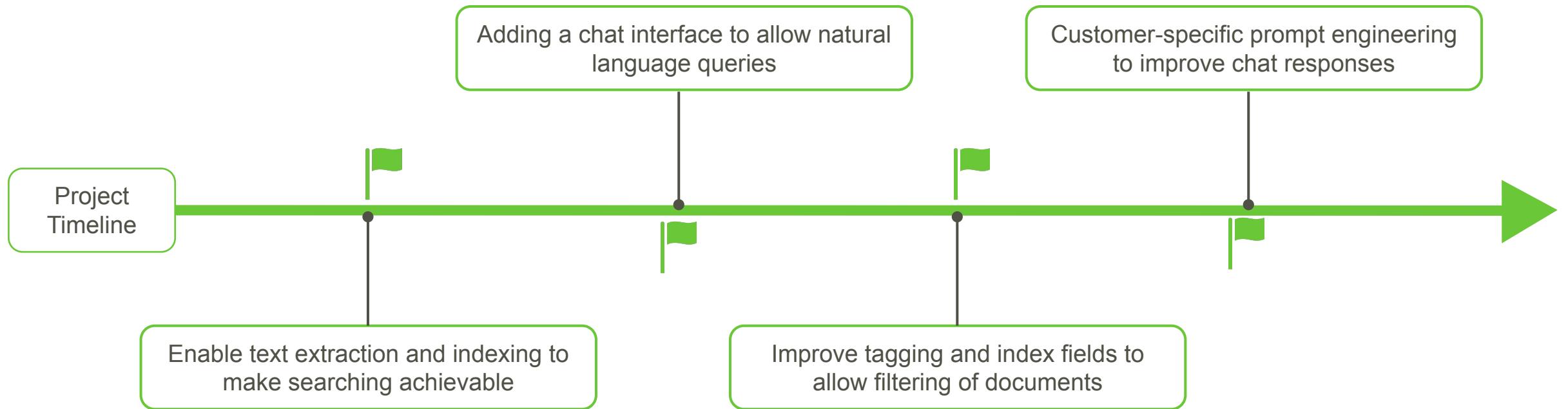
Consolidate reports for easy access

Add AI-powered insights to human interpretation

Amplify the investment team's time

Indexing and Querying Quarterly Reports

High level project milestones

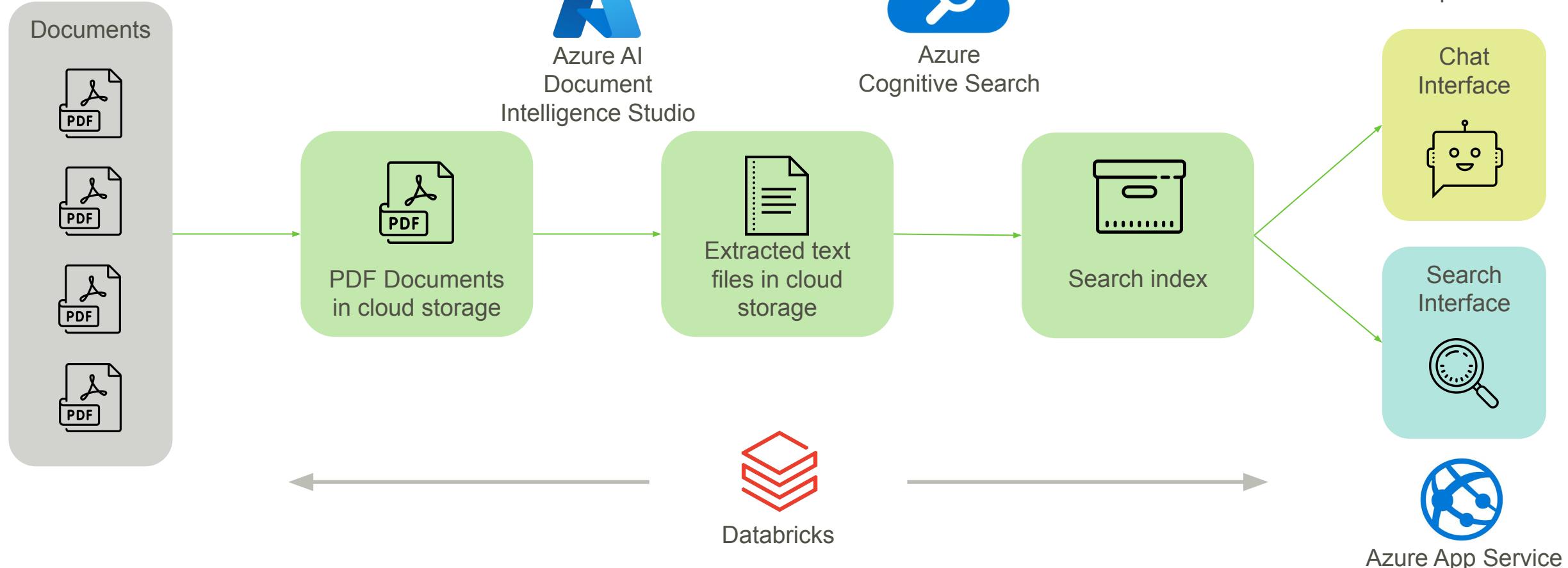


Indexing and Querying Quarterly Reports

Proposed architecture

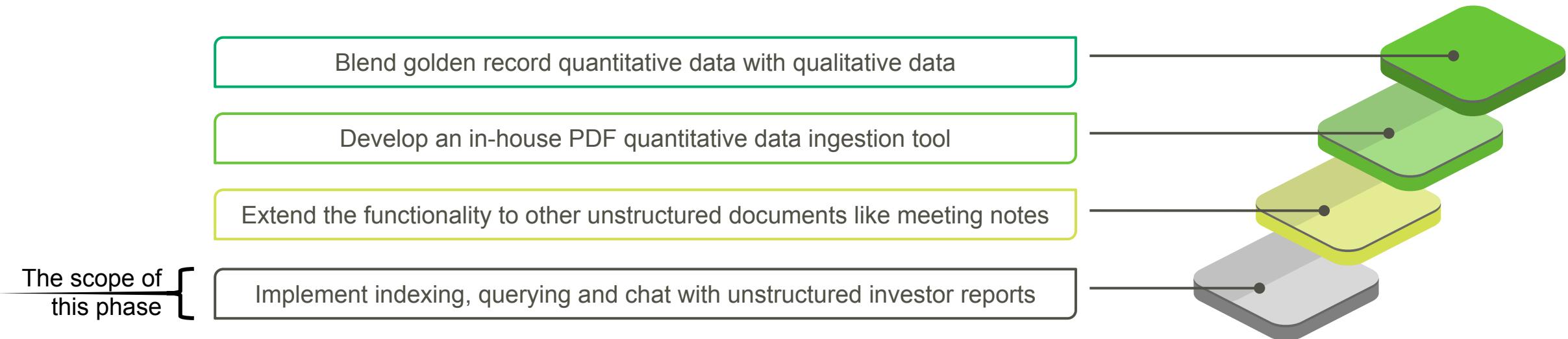


Azure
Open AI



Future vision for the solution

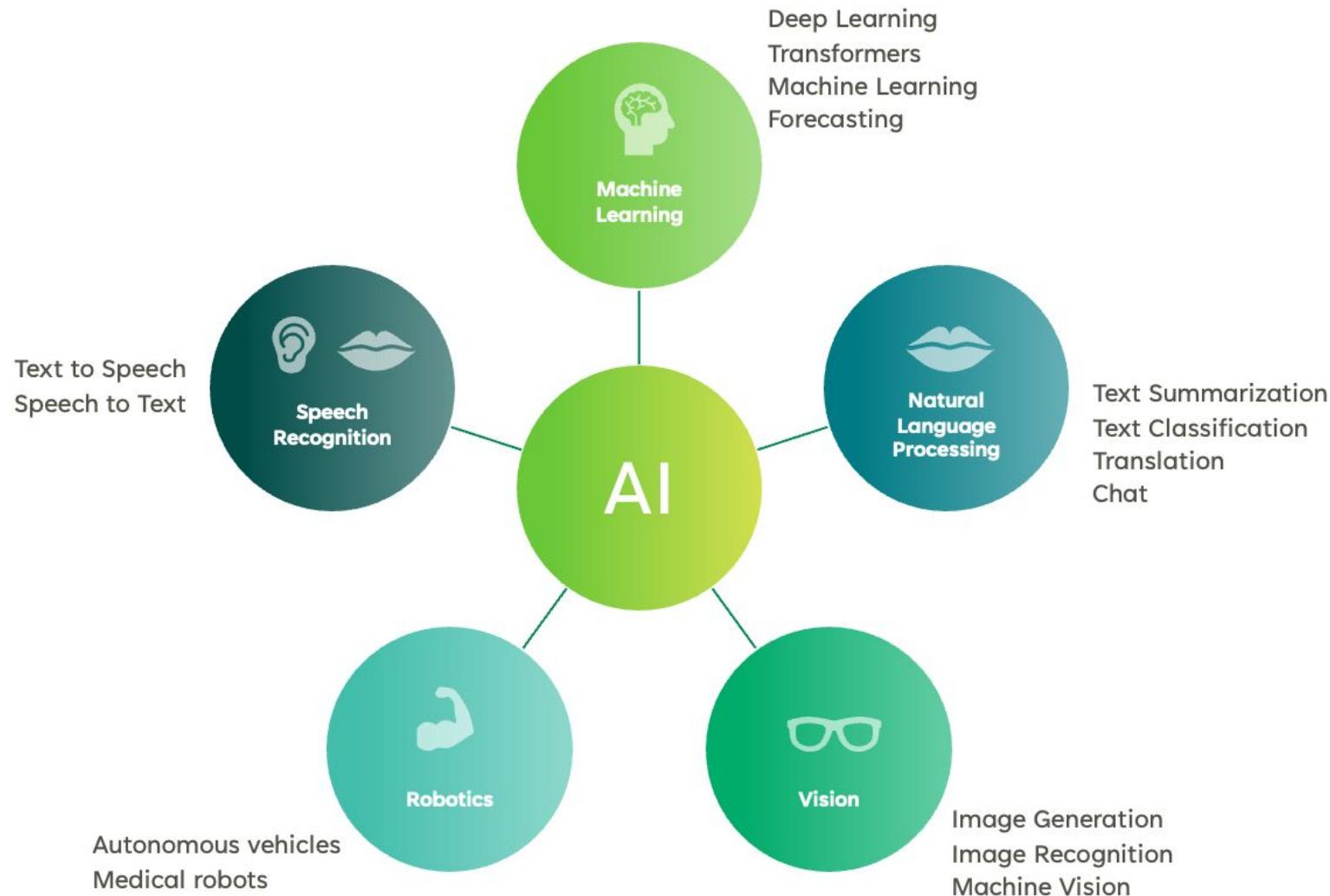
A stepping stones approach



Observations, suggestions and what next?

thorogood.com

Observation and Suggestions



- Focus on the business cases, and what will deliver value for your organization
- Think wider than just chat models
- Testing, monitoring and ongoing improvement is key
- AI can build on and augment your existing data

As you embark on a journey into the world of
Generative AI,

here's the immediate next step...

What's Next

Governing Your Data Estate in the Age of AI

Register now



Reshape the way you perceive and implement governance in your data and AI initiatives. This webinar is tailored for data enthusiasts, business leaders, and technology professionals seeking to strike the right balance between governance and innovation in GenAI use cases.

Data Intelligence Day 2024

Register now



Join us for an in-person half-day event and dive into this fundamental shift in data platforms. Whether you aim to democratize data and AI across your organization or you simply want to see how data intelligence platforms handle production-scale AI training, governance or ETL — you'll walk away with actionable strategies.

Find the next Data Intelligence Day near you

Complimentary one-hour consultation?

Sign up now



We would like to invite you to complimentary consultation with the **Thorogood** team. Please register your interest here; a member of our team will reach out shortly.



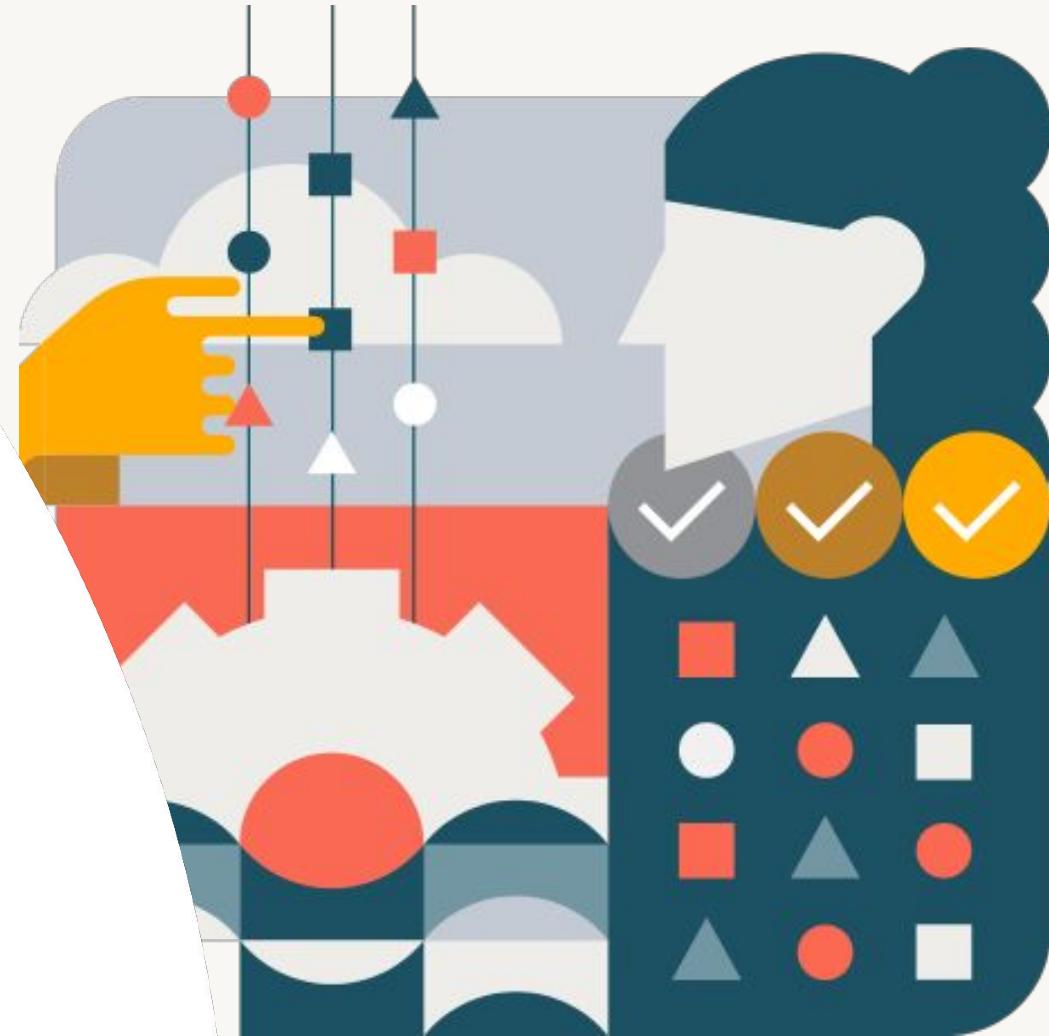
18 April | Live Webinar

Governing Your Data Estate in the Age of AI

In this session, you'll learn:

- How to achieve governance for your Data and AI with Databricks Unity Catalog
- How Unity Catalog gives data teams the power to create a single source of truth for their data and AI assets
- How to streamline processes with Lakehouse AI, from managing access permissions to sharing data and AI assets across regions, clouds, and data platforms
- Real-world examples of Unity Catalog (UC) governance models to illustrate effective governance strategies

[Register now](#)





DATA INTELLIGENCE DAY

Asia Pacific | In-person

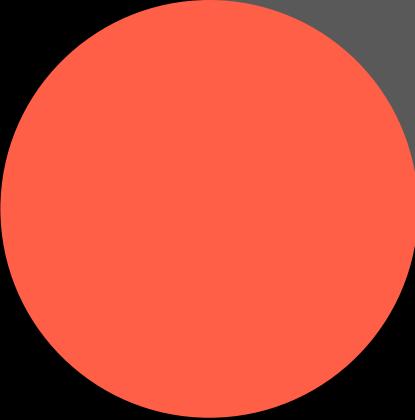
Data Intelligence Day 2024

As AI reshapes the world, the leaders in every industry will be those who use data and AI to power their organizations. Discover the new unified platform that delivers AI for your data, and data for your AI: the Databricks Data Intelligence Platform.

Join us for an in-person half-day event and dive into this fundamental shift in data platforms. Whether you aim to democratize data and AI across your organization or you simply want to see how data intelligence platforms handle production-scale AI training, governance or ETL — you'll walk away with actionable strategies.

Find the next Data Intelligence Day near you

Register now



Complete this feedback form and let us know how we can help shape your organization's GenAI journey.





databricks

