



IT Problems solved. Consultancy expertise at Contractor prices. Solutions guaranteed.

Microsoft Intelligent Data Platform

Power AI innovation, no matter where your data resides

Databases • Analytics • AI • Governance



Azure SQL Database



Azure Cosmos DB



Azure Database for PostgreSQL



Azure Database for MySQL



Microsoft Fabric



Azure Databricks



Azure AI



Microsoft Purview



snowflake®



MongoDB®

ORACLE

Redis

Get the most from your data

Achieve your ambitions

Build confidently

Build your way



Consultancy at contractor prices



A low risk choice



Consulting as a Subscription



Agile implementation



Flexible Consulting Framework



"Speed to Value" is our USP



Greater control over suppliers



Tracked communication



Decrease in lead times



Control over the process



Greater transparency



Current workflow status



Large models demand extensive resources, expertise, and precision

To realize the potential of large models, we must overcome the barriers



Managing infrastructure

Managing infrastructure and environment dependencies for models can be cumbersome.



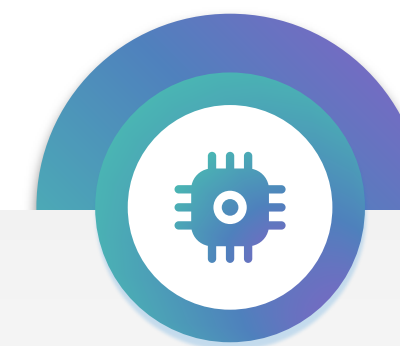
Training and fine-tuning large models

Training and fine-tuning large models is often a compute and time-intensive task.



Finding and comparing models

Finding and comparing the right models for you based on your own criteria isn't the most straightforward.



Using AI responsibly

Sophisticated content moderation tools are necessary to ensure AI solutions are ethical and effective.



Consultancy at contractor prices



A low risk choice



Consulting as a Subscription



Agile implementation



Flexible Consulting Framework



"Speed to Value" is our USP



Greater control over suppliers



Tracked communication



Decrease in lead times



Control over the process



Greater transparency



Current workflow status



Azure AI offers the best model selection

Gold
Microsoft Partner

LLMs and SLMs

Flagship LLMs

- GPT-4
- Mistral Large
- Llama3 70B, 405B
- Command R+
- AI21

Small Language Models

- Phi3
- Mistral OSS models
- Llama3 8B

Modalities, tasks and tools

Multi-modal

- GPT-4o, Phi3-vision

Image generation

- Dalle3, Stability AI

Embedding models

- Ada, Cohere
- Function calling & JSON support

Regional and domain specific

- Core42 JAIS Arabic language LLM
- Mistral Large is focused on European languages
- Nixtla TimeGEN-1 - Timeseries forecasting

Open and propriety

Premium models first on Azure: OpenAI, Mistral Large, Cohere Command R+

100s of Open models from HuggingFace

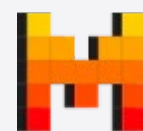
Open models from Meta, Databricks and Snowflake, Nvidia



Microsoft Phi



Azure OpenAI



Mistral AI



Meta AI



Databricks



Cohere



Hugging Face



NVIDIA



Deci AI



Nixtla



G42



Snowflake



Consultancy at contractor prices



A low risk choice



Consulting as a Subscription



Agile implementation



Flexible Consulting Framework



"Speed to Value" is our USP



Greater control over suppliers



Tracked communication



Decrease in lead times



Control over the process



Greater transparency



Current workflow status

ERP Project control staffing & systems:

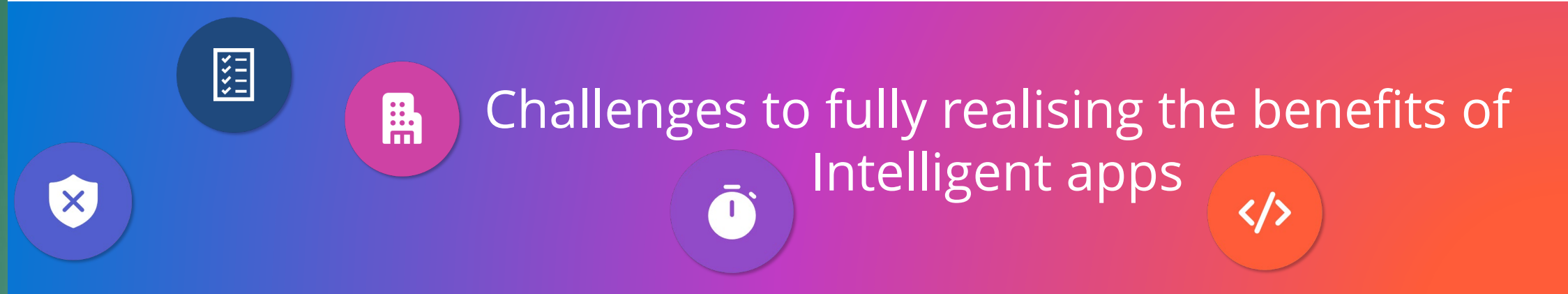
Connect people, content,
data and systems



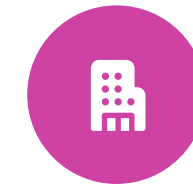
- Your Project & IT problems solved
- Flexible managed staffing model
- Less risk: seasoned Project consultants
- Consulting expertise at Contractor prices
- Bespoke integration to match your needs
- Solutions Guaranteed

Solve the Challenge
of Everyday processes

+44 (0) 207 101 0695 | big-dada.co.uk



Shortage of skilled developers



Legacy app and data architectures



Long release cycles



Growing security threats



Aligning use cases to business outcomes



Azure AI application platform



Developer services

Build New or Modernize

Platform engineering and AI-assisted operations

API and Integration services

Container services

Platform-as-a-Service

Serverless

Data platform

AI platform

Partner Tools and Services

Security, Governance and Responsible AI

AI-assisted development

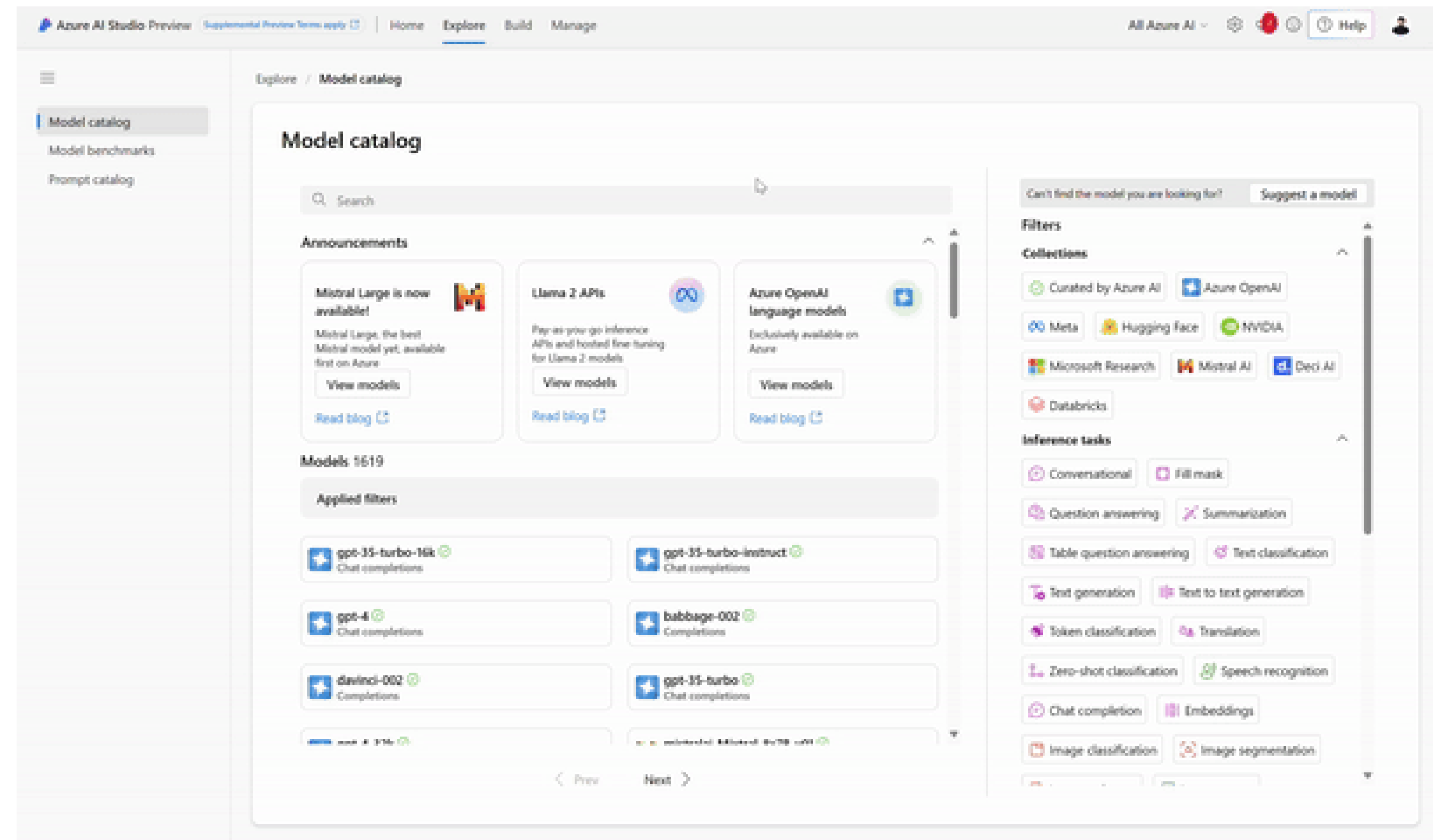
Any App, architecture and language

Overview of Model Catalog capabilities

Catalog featuring the best foundation model collections

- Popular OSS models handpicked and optimized by Azure AI.
- Curated by Microsoft, Meta, and Hugging Face, Cohere, Mistral, and more to offer thousands of foundation models for inference.
- Featuring Azure OpenAI models.

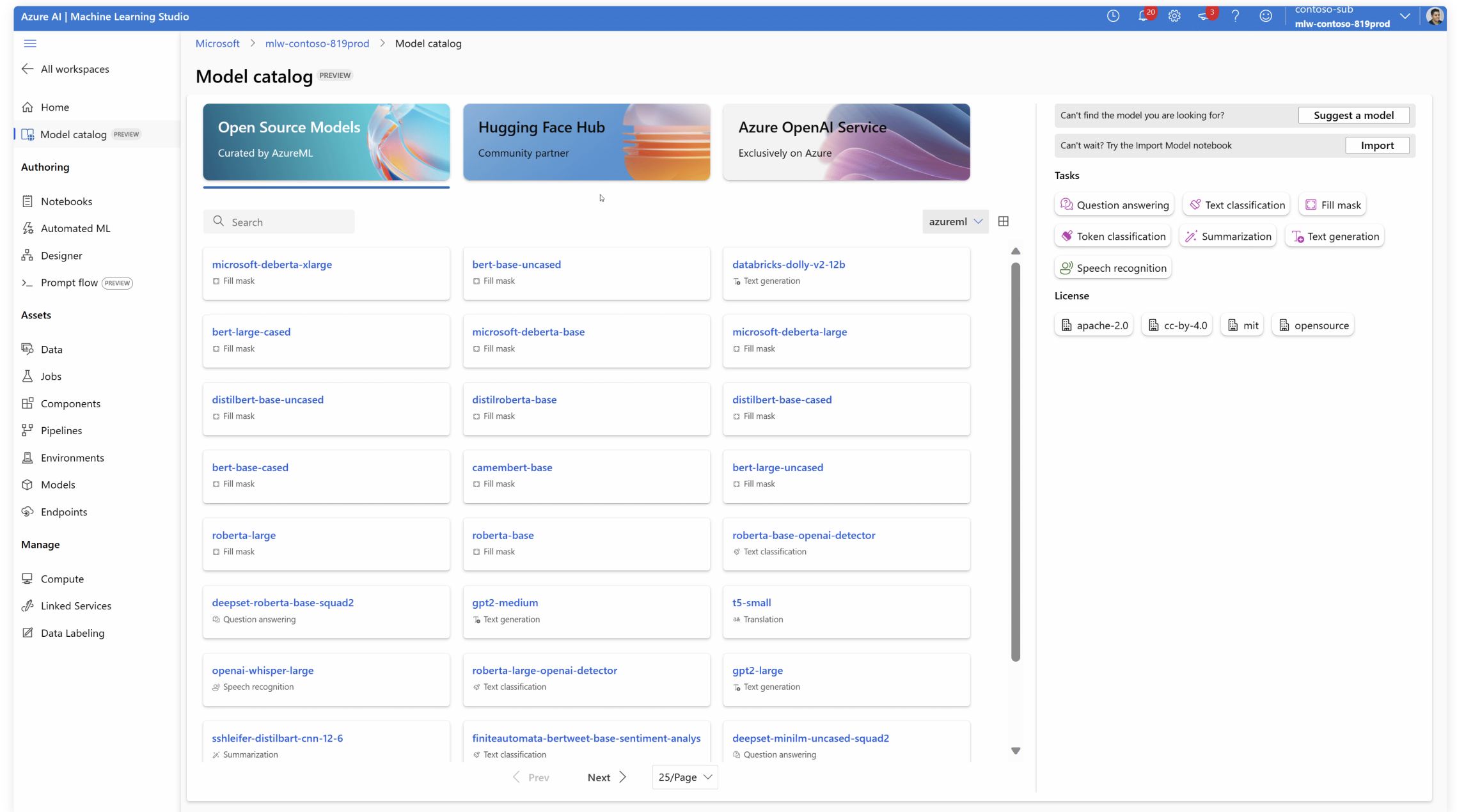
Users can access the Model Catalog in Azure AI Studio and Azure Machine Learning.





Model cards and playground

- Explore models by tasks.
- Model summary, link to the original model card, samples for inference, evaluation, and fine-tuning.
- Playground to try sample queries.



Consultancy at
contractor prices



A low risk
choice



Consulting as a
Subscription



Agile
implementation



Flexible Consulting
Framework



"Speed to Value"
is our USP



Greater control
over suppliers



Tracked
communication



Decrease in
lead times



Control over
the process



Greater
transparency



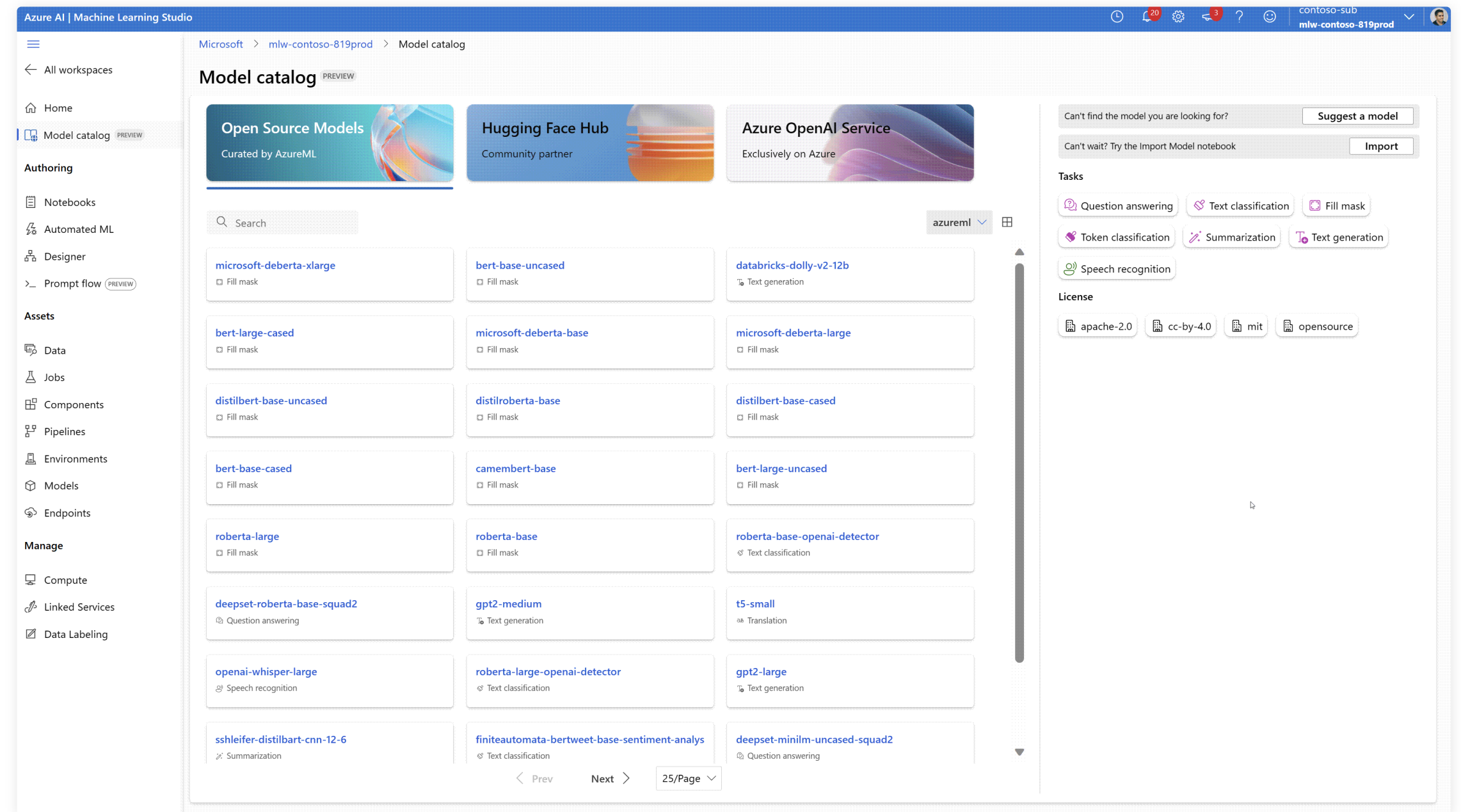
Current
workflow status



Deploy MaaP to managed endpoints

Managed Online Endpoints offer:

- Managed instances, no need to create or manage VMs/clusters.
- Traffic management for safe roll out: Split or shadow traffic across multiple model versions.
- Auto scale to several instances based on utilization metrics or schedule.
- Secure hosting with private endpoints secured in VENTs.
- Out-of-box monitoring and drift.



Consultancy at
contractor prices



A low risk
choice



Consulting as a
Subscription



Agile
implementation



Flexible Consulting
Framework



"Speed to Value"
is our USP



Greater control
over suppliers



Tracked
communication



Decrease in
lead times



Control over
the process



Greater
transparency



Current
workflow status



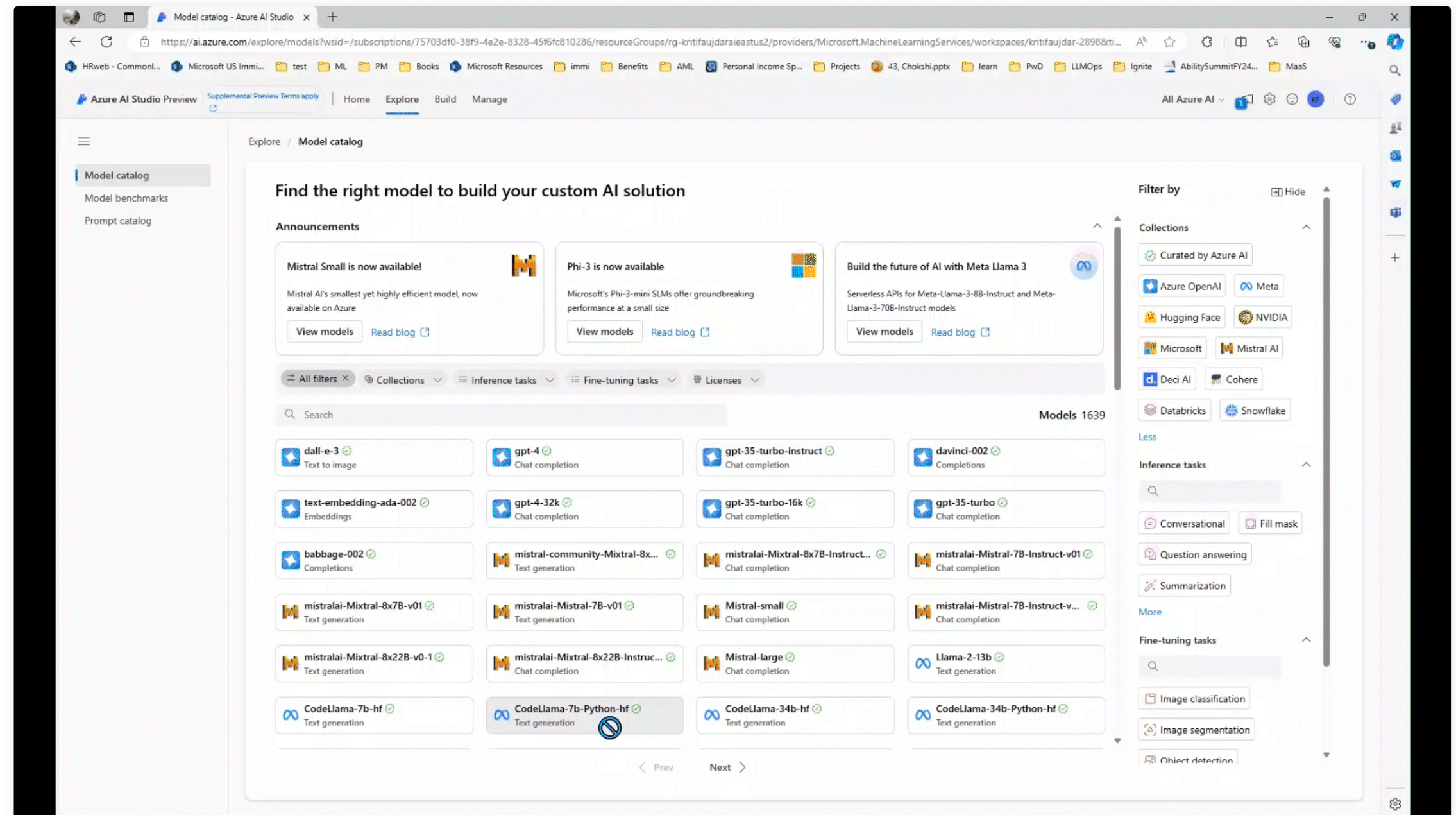
Deploy MaaS through pay-as-you-go inference APIs

PayGo inference APIs offer:

No need to host models in dedicated VMs.

Billed based on input and output tokens used, making it easier to get started

Use models in the playground or within your favorite tools like prompt flow.





Evaluate models

- Benchmark model performance with your datasets.
- Compare metrics across evaluation jobs to identify models with best accuracy.
- Establish baseline performance to compare improvements with fine-tuning.

All workspaces

Home

Model catalogPREVIEW

Authoring

Assets

Manage

Notebooks

Automated ML

Designer

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

Compute

Linked Services

Data Labeling

Microsoft > model-evaluation-demo > Model catalog > azureml > deepset-roberta-base-squad2

deepset-roberta-base-squad2

Overview

Versions

Artifacts

Task: Question answering

Finetuning task: text-classification

Finetuning task: token-classification

Finetuning task: question-answering

Languages: en

License: cc-by-4.0

Refresh

Evaluate

Finetune

Deploy

Description

Inference samples

Finetuning samples

Model Evaluation

Sample inference ⓘ

Examples ▾

Query

What is my name?

Context

My name is John and I live in Seattle.

Test



Fine-tune models (MaaP)

Ready-to-use fine-tuning pipelines to get started quickly – no need to spend time installing frameworks/dependencies.

Optimizations to reduce fine-tuning resources and time.

Fine-tune using UI, Notebook (Python SDK) or CLI (YAML).

distilbert-base-uncased-finetuned-sst-2-english

Task: Text classification | Finetuning task: text-classification | Finetuning task: token-classification | Finetuning task: question-answering | Languages: en | License: apache-2.0

Refresh | Evaluate | Finetune | Deploy

Description

This is a fine-tuned version of DistilBERT-base-uncased, trained on SST-2, which reached 91.3 % accuracy on the dev set. Developed by Hugging Face, it's mainly intended to be used for topic classification and can be fine-tuned on downstream tasks, but it's important to keep in mind that it has certain biases, such as biased predictions for certain underrepresented populations and that it should not be used to create hostile or alienating environments for people. Additionally, the authors used the Stanford Sentiment Treebank (sst2) corpora for training the model. It's recommended to evaluate the risks of this model by thoroughly probing the bias evaluation datasets like WinoBias, WinoGender, Stereoset.

The above summary was generated using ChatGPT. Review the [original model card](#) for more information about the model, evaluation metrics, license, intended uses, limitations and bias before using it.

Inference samples

Inference type	Python sample (Notebook)	CLI with YAML
Real time	entailment-contradiction-online.ipynb	text-classification-online.sh
Batch	entailment-contradiction-batch.ipynb	coming soon

Finetuning samples

Task	Use case	Dataset	Python sample (Notebook)	CLI with YAML
Text Classification	Emotion Detection	Emotion	emotion-detection.ipynb	emotion-detection.sh
Token Classification	Named Entity Recognition	Conll2003	named-entity-recognition.ipynb	named-entity-recognition.sh
Question Answering	Extractive Q&A	SQUAD (Wikipedia)	extractive-qa.ipynb	extractive-qa.sh

Model Evaluation

Task	Use case	Dataset	Python sample (Notebook)	CLI with YAML
Text	Sentiment	SST-2	evaluate-model-sentiment.ipynb	evaluate-model-sentiment.sh

Sample inference

Query: Love the food here, we'll come back again.

Test

Looking for a model that can check if users are unhappy...



Customer story

“Having a **conversational interface** option gets us closer to our goals of **fully engaging the customer and personalizing their experience** by showing them the most relevant products at the most relevant time.”

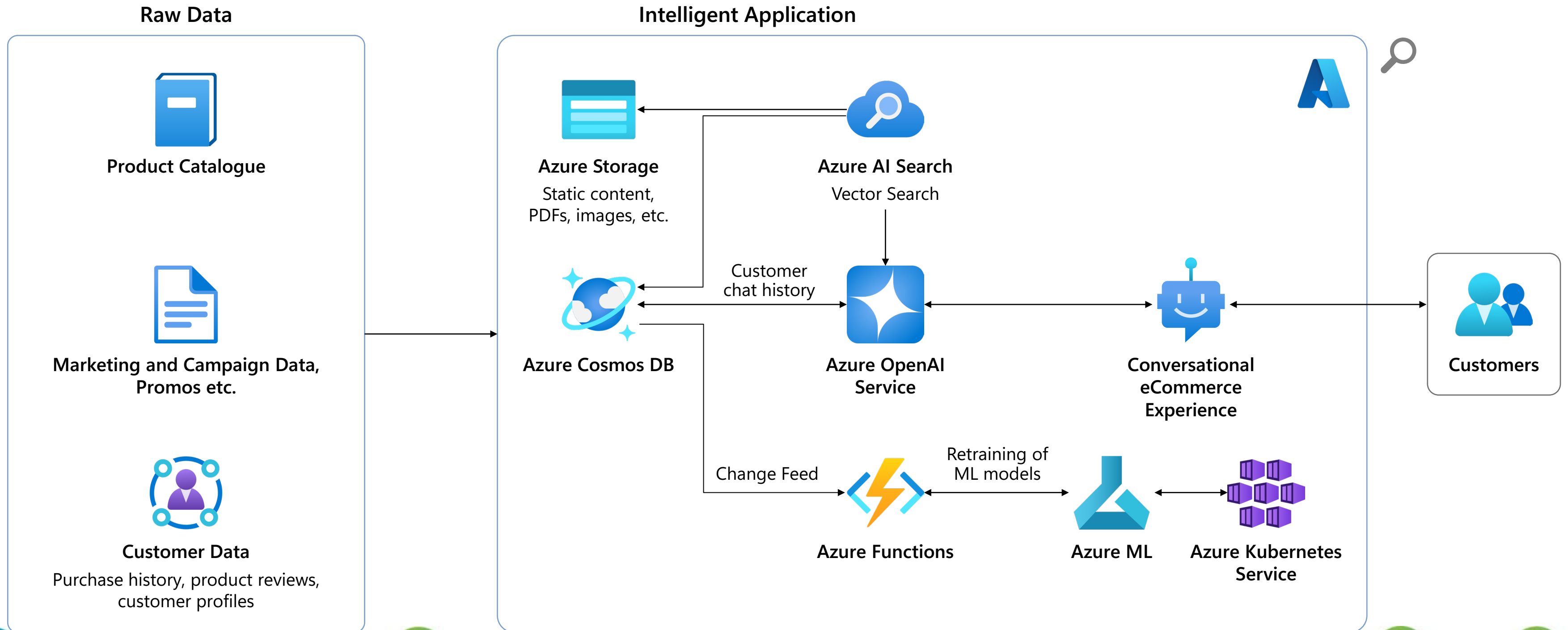
Cliff Cohen, Chief Technology Officer, Asos



Personalization and product recommendations

Deliver real-time recommendations and dynamic customer experiences

Use cases





Content generation and marketing



Use Case Description

- Generates text or image content using generative AI.
- Automates the creation of web or mobile content, such as product descriptions for an online catalog or visual campaign assets based on marketing narratives, accelerating time to market.
- Enable faster iteration and A/B testing to identify the best descriptions that resonate with customers.

Key considerations

- Integrate multiple data types (e.g. product information, reviews, user profiles etc..)
- Generative AI models that can generate text and images based on user profile and behavior
- Integration with marketing campaign or website content workflows



REINVENT CUSTOMER
EXPERIENCES

Use Case Examples

- Marketing campaign ideation and content drafting, automated translation
- Campaign or website message testing
- Automated summary of user reviews
- Marketplace product descriptions and image generation



Consultancy at
contractor prices



A low risk
choice



Consulting as a
Subscription



Agile
implementation



Flexible Consulting
Framework



"Speed to Value"
is our USP



Greater control
over suppliers



Tracked
communication



Decrease in
lead times



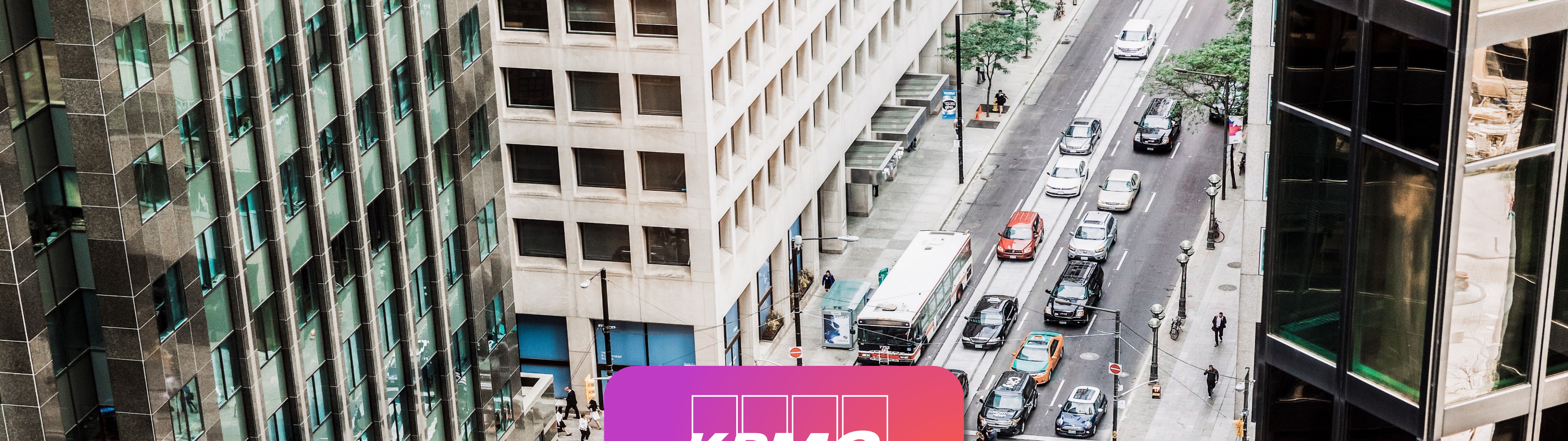
Control over
the process



Greater
transparency



Current
workflow status



KPMG

Customer story

“In the past, (this employee) might have had to liaise with HR or conduct a **time-consuming web search of multiple sources**...Now they simply ask KymChat to create a list of 10 role-specific interview questions, and they have their **interview prep finished in seconds**.”

Robert Finlayson, Senior Product Manager, KPMG Australia



Consultancy at
contractor prices



A low risk
choice



Consulting as a
Subscription



Agile
implementation



Flexible Consulting
Framework



"Speed to Value"
is our USP



Greater control
over suppliers



Tracked
communication



Decrease in
lead times



Control over
the process



Greater
transparency



Current
workflow status



Build your own copilot architecture

Gold
Microsoft Partner



Use cases



Legend:

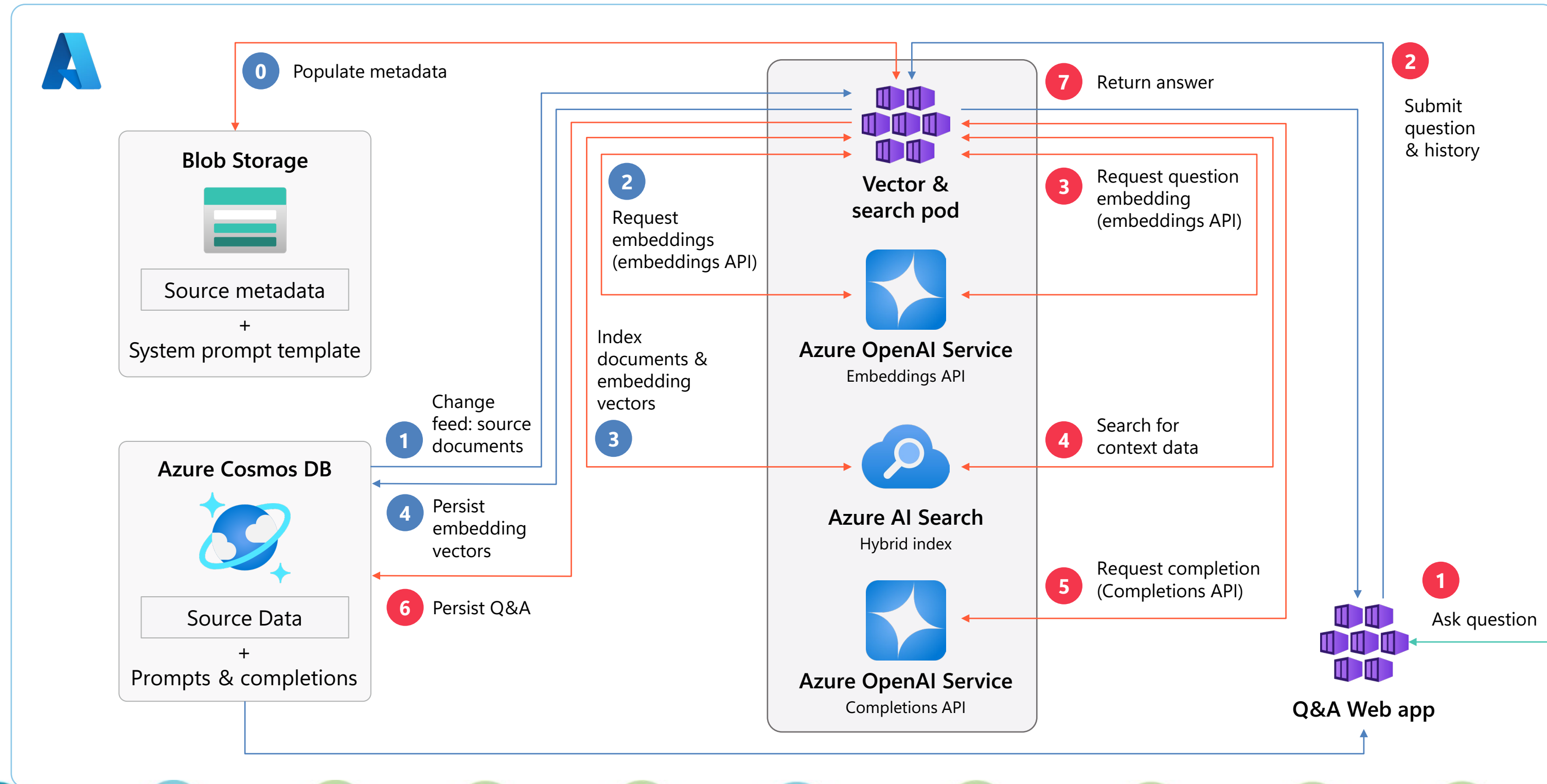
Semantic Kernel
orchestrated call

Other service calls

User interaction

Backend
data flow

Frontend
workflow



Chat users

Q&A Web app

1 Ask question

2 Submit question & history

7 Return answer

3 Request question embedding (embeddings API)

4 Search for context data

5 Request completion (Completions API)

2 Request embeddings (embeddings API)

3 Index documents & embedding vectors

1 Change feed: source documents

4 Persist embedding vectors

6 Persist Q&A

0 Populate metadata

Blob Storage

Source metadata
+
System prompt template

Azure Cosmos DB

Source Data
+
Prompts & completions

Vector & search pod

Azure OpenAI Service
Embeddings API

Azure AI Search
Hybrid index

Azure OpenAI Service
Completions API

Consultancy at contractor prices

A low risk choice

Consulting as a Subscription

Agile implementation

Flexible Consulting Framework

"Speed to Value" is our USP

Greater control over suppliers

Tracked communication

Decrease in lead times

Control over the process

Greater transparency

Current workflow status

GitHub repo



Transaction processing

Use cases



Use Case Description

- Automate data extraction, categorization, and validation of data from documents
- Transform unstructured data, such as claims, contracts, applications, invoices, and forms into structured digital information to facilitate downstream processes like insurance claim processing
- Leverage computer vision to visually inspect machinery and products for defects, preventative maintenance, and safety concerns
- Benefits: Process improvement, leading to greater productivity

Key considerations

- Manage industry-specific compliance standards and regulations governing transactional data
- Distributed database systems, designed for horizontal scaling, are increasingly used to handle large volumes of transactions.
 - AI and machine learning can automate routine tasks, and improve overall efficiency in transaction processing



REINVENT CUSTOMER
EXPERIENCES

Use Case Examples

- Order processing
- Buy online, pick up in store
- Banking payments
- Airline reservations
- Employee payroll records
- Patient appointment scheduling
- Quality assurance
- Equipment maintenance



Consultancy at
contractor prices



A low risk
choice



Consulting as a
Subscription



Agile
implementation



Flexible Consulting
Framework



"Speed to Value"
is our USP



Greater control
over suppliers



Tracked
communication



Decrease in
lead times



Control over
the process



Greater
transparency



Current
workflow status



Customer story

“Our first ambition at Swift is to **build a foundation model** for anomaly detection that underpins the **detection and prevention of fraud**. Our ultimate goal is collaborating with Microsoft and our community to start thinking about how we can stop fraud occurring in payments.”

Johan Bryssinck, AI/ML Product and Program Management Lead, Swift



Consultancy at contractor prices



A low risk choice



Consulting as a Subscription



Agile implementation



Flexible Consulting Framework



"Speed to Value" is our USP



Greater control over suppliers



Tracked communication



Decrease in lead times



Control over the process



Greater transparency



Current workflow status

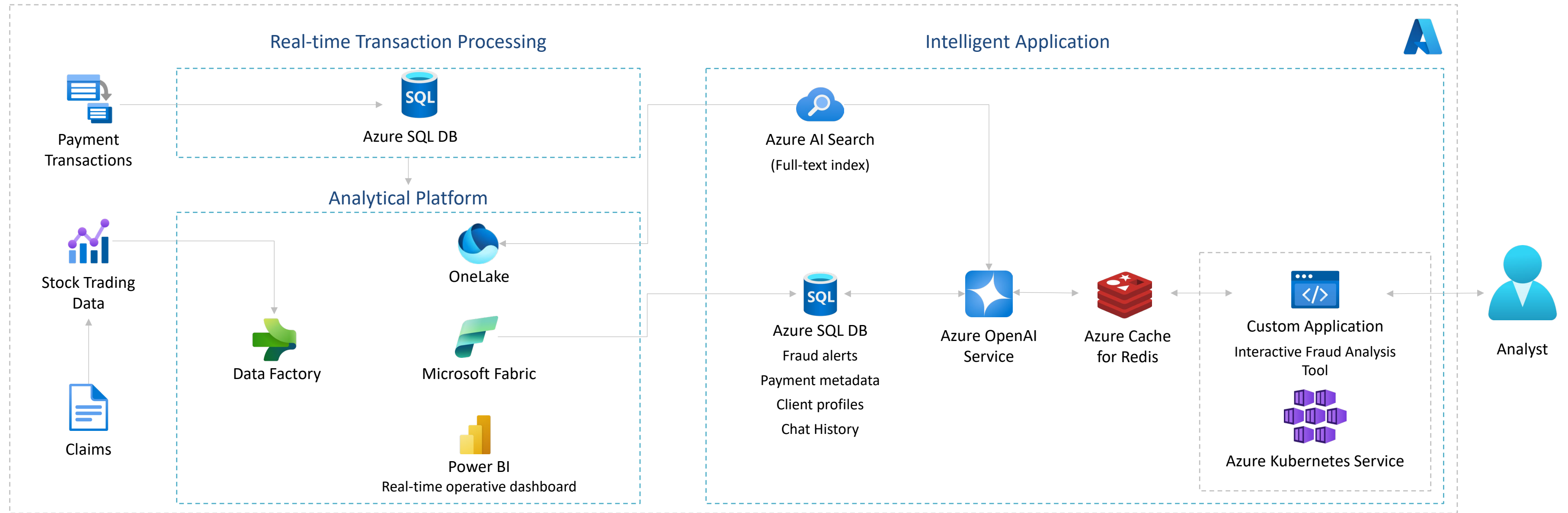


Sample architecture: Interactive fraud analysis tool

Gold
Microsoft Partner



Use cases



Fraud
Detection

This technical pattern enables the identification of fraudulent transactions, combining transaction and meta information with the context provided by LLMs.



Consultancy at
contractor prices



A low risk
choice



Consulting as a
Subscription



Agile
implementation



Flexible Consulting
Framework



"Speed to Value"
is our USP



Greater control
over suppliers



Tracked
communication



Decrease in
lead times



Control over
the process



Greater
transparency



Current
workflow status



Knowledge mining



Use Case Description

- Analyze conversations for keywords to **spot trends** and better **understand customers**
- Quickly find information to gain valuable **insights** and identify **patterns**
- Utilize information in downstream processes to improve **decision-making**

Key Considerations

- Capabilities to extract relevant information and make it available for further analysis
- Pre-trained machine learning for identifying patterns and making predictions



REINVENT CUSTOMER
EXPERIENCES

Use Case Examples

- Curated content summarization
- Knowledge extraction and organization
- Trend and sentiment analysis
- Automated report and research generation



Consultancy at contractor prices



A low risk choice



Consulting as a Subscription



Agile implementation



Flexible Consulting Framework



"Speed to Value" is our USP



Greater control over suppliers



Tracked communication



Decrease in lead times



Control over the process



Greater transparency



Current workflow status



Follow us for more tools, White Papers, templates and job openings

Phone +44(0)207 1010 695

Email info@big-dada.co.uk

Location 118 Pall Mall, London, SW1Y 5EA



Consultancy at
contractor prices



A low risk
choice



Consulting as a
Subscription



Agile
implementation



Flexible Consulting
Framework



"Speed to Value"
is our USP



Greater control
over suppliers



Tracked
communication



Decrease in
lead times



Control over
the process



Greater
transparency



Current
workflow status