

Chat and interact with your data using Azure AI Studio... and more!

Lorenzo Barbieri
Principal Consultant @ SoftwareOne

Vectors and Embeddings

The background features a vibrant rainbow gradient from yellow to purple. It is populated with various 3D-rendered geometric objects: a large, multi-colored sphere in the upper left; a large, transparent, multi-faceted prism-like structure on the right; a smaller, textured sphere on the far right; and a red, faceted object at the bottom center. The overall aesthetic is futuristic and abstract.

What is a vector?

Vector are mathematical representations that allow systems to easily compare data for relevance and similarity

Create an embedding

"A red tent"



OpenAI CLIP
In
Azure AI Studio

<https://openai.com/index/clip>

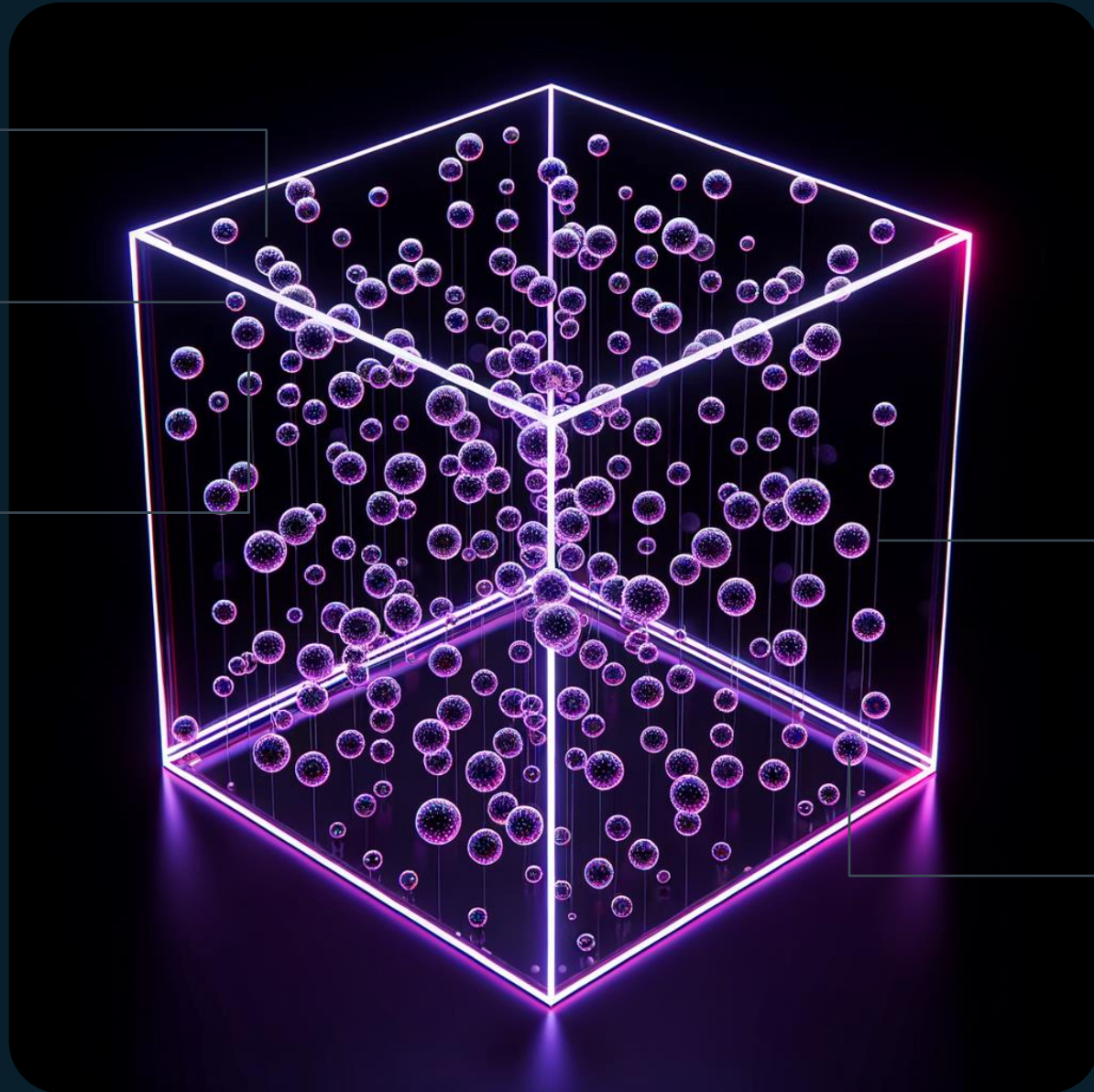
[0.031198505, 0.094205155,
0.044205155, 0.019588541,
0.019588541, 0.019588541 ...]

[0.031198505, 0.094205155,
0.044205155, 0.019588541,
0.019588541, 0.019588541 ...]

Dog

Kitten

Cat



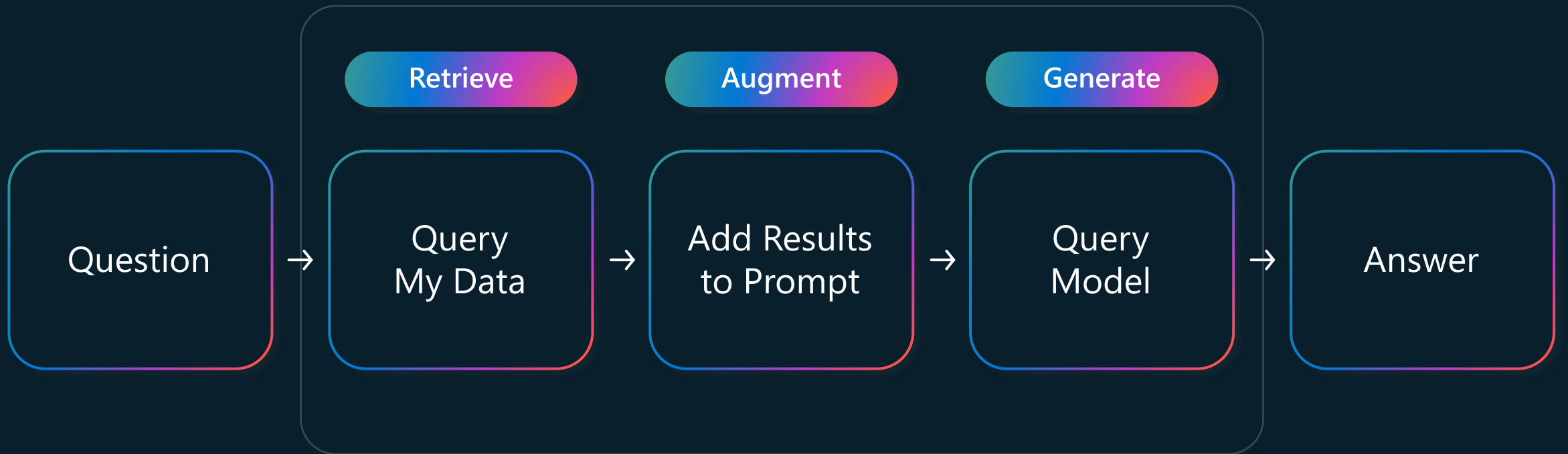
Boat

The Moon

Retrieval and Generation



Retrieval Augmented Generation



Augmenting the prompt

System

Task

You are an AI agent for the Contoso Trek outdoor products retailer. As the agent, you answer questions briefly, succinctly, and in a personable manner using markdown and even add some personal flair with appropriate emojis.

Documents

The following documentation should be used in the response. The response should specifically include the product id.

TrailWalker hiking Shoes

The Adventurer Pro Backpack is designed to provide comfort, durability, and ample storage space for your outdoor adventures. Familiarize yourself with the key features of the backpack [...]

TrekHiker Walking Booths

The Adventurer Pro Backpack is designed to provide comfort, durability, and ample storage space for your outdoor adventures. Familiarize yourself with the key features of the backpack [...]

User

I need warm waterproof shoes to go on a hike.



Assistant

Sure, I'd be happy to help! 😊 Based on the available documentation, I can recommend two choices from the Contoso Trek catalog.

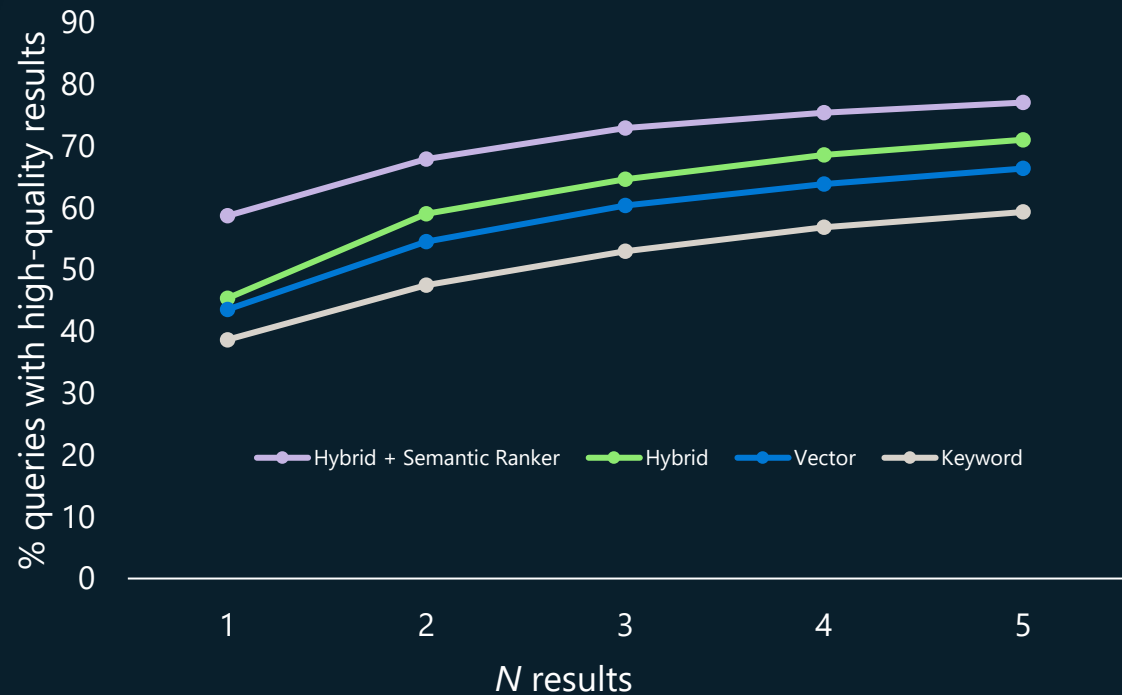
1.Product:
TrailWalker Hiking Shoes ID:
36244753 **Brand**
: TrekReady
[...]



Azure AI Search

Powerful information retrieval
for generative AI applications

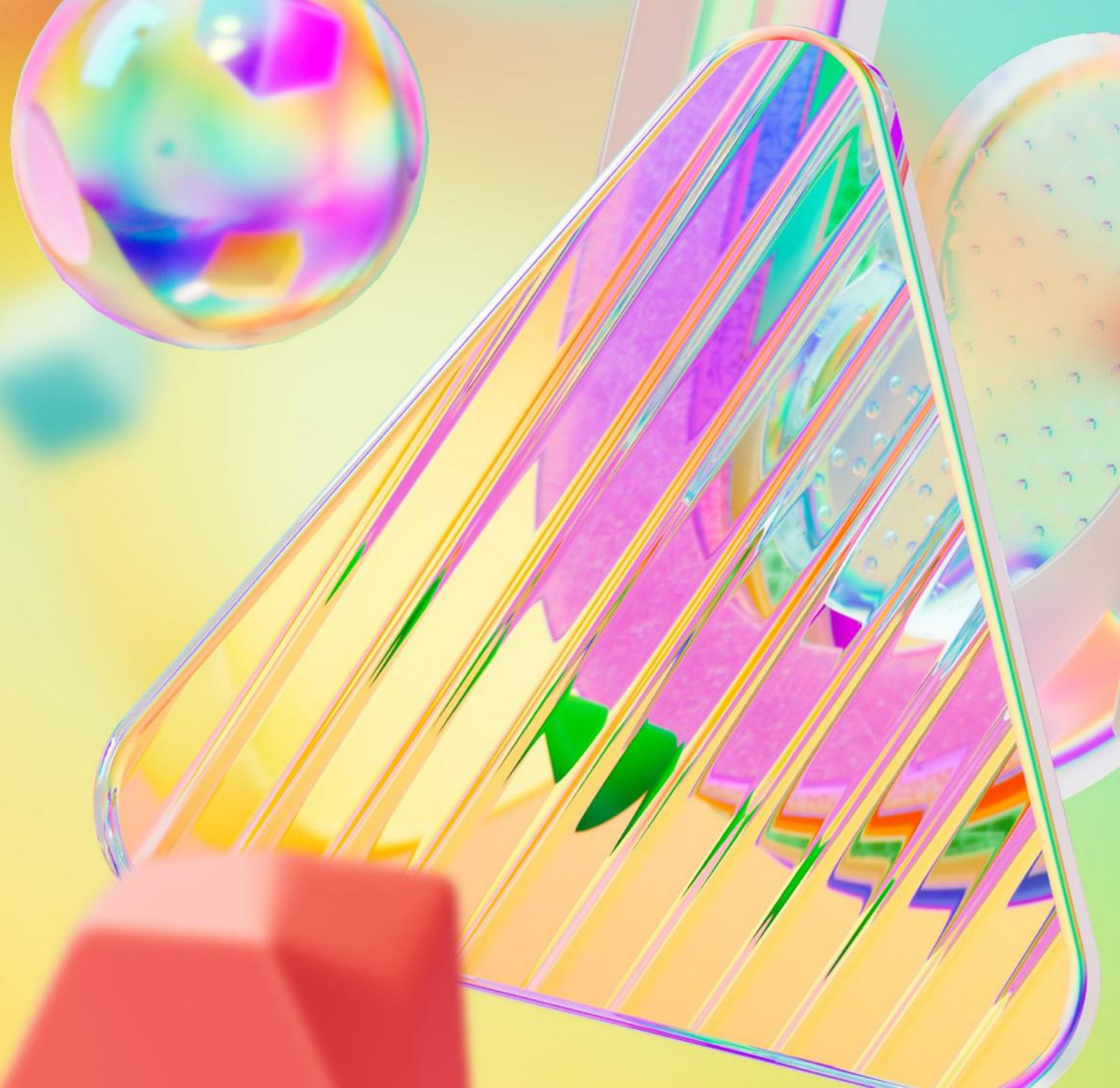
Azure AI Search: vectors + keywords + re-ranking



Percentage of queries where high-quality chunks are found in the top 1 to 5 results, compared across search configurations. All retrieval modes used the same set of customer query/document benchmark. Document chunks were 512 tokens with 25% overlap. Vector and hybrid retrieval used Azure OpenAI Ada-002 embeddings.

[Azure AI Search: Outperforming vector search with hybrid retrieval and ranking capabilities – Microsoft Community Hub](#)

Creating a custom Copilot

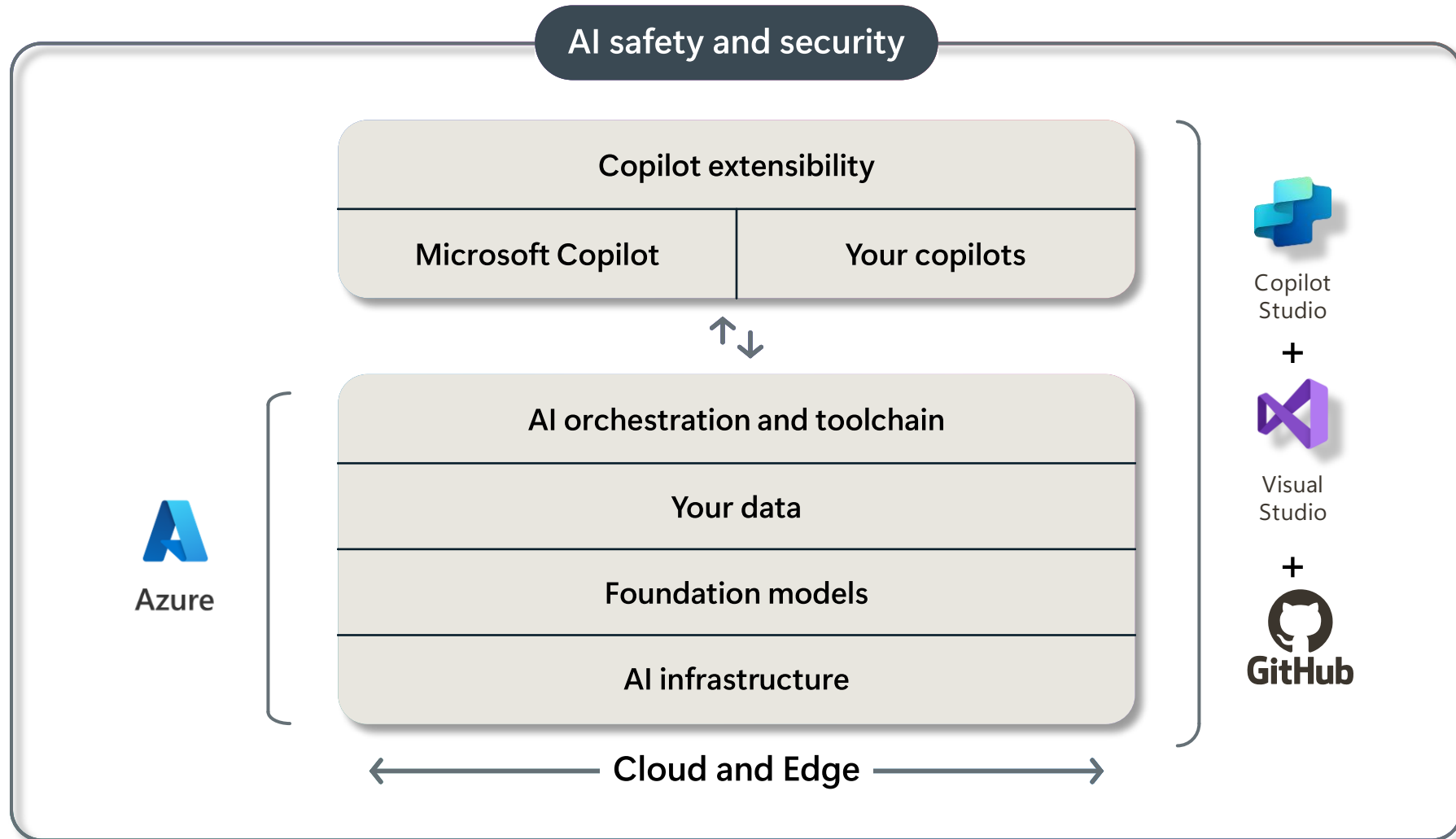


What is a Copilot?

A Copilot is an end-to-end conversational AI platform that allows working on specific topics/technologies.



Copilot stack



Azure AI

Best-in-class AI foundation models



Azure AI Services

Pre-trained, turnkey solutions for intelligent applications



Azure Machine Learning

Full-lifecycle tools for designing and managing AI models



Responsible AI Tooling

Build and manage apps that are trustworthy by design



Azure AI Studio

A comprehensive platform to develop and deploy custom copilots

Azure AI Studio



Azure AI Studio

Generally available

ai.azure.com



The screenshot shows the Azure AI Studio Model catalog interface. The top navigation bar includes the Azure AI Studio logo, the current project name 'Contoso-0034', and user information 'Contoso Responsible R&D' and 'ContosoPharmaceutical-0034'. The left sidebar contains a navigation menu with sections: 'To all projects', 'Current project', 'Project overview', 'Settings', 'Get started' (with 'Model catalog' selected), 'Model benchmarks', 'Prompt catalog', 'Project playground' (with 'Chat', 'Completion', 'Assistants', 'Images'), 'Tools' (with 'Code', 'Prompt flow', 'Tracing', 'Evaluation', 'Fine-tuning'), and 'Components' (with 'Data', 'Indexes', 'Deployments', 'Content filters'). The main content area is titled 'Find the right model to build your custom AI solution' and features 'Announcements' for 'Phi-3 is now available', 'Build the future of AI with Meta Llama 3', and 'Cohere models are here'. Below the announcements are filter options for 'All filters', 'Publishers', 'Deployment type', 'Tasks', and 'Licenses'. A search bar is present, and the model list shows 1,234 models. The visible models include: gpt-4-32k (Text generation), llama-7b (Text generation), mistralai-Mixtral-8x7B-Instruct (Chat completion), databricks-dolly (Text generation), Deci-DeciLM-7B (Text generation), gpt-35-turbo (Text generation), llama-12b (Text generation), microsoft-phi-2 (Text generation), llama-finetune-12b (Text generation), gpt-4 (Text generation), tiuae-falcon-40b (Text generation), gpt-4 (Text generation), gpt-4-32k (Text generation), llama-7b (Text generation), mistralai-Mixtral-8x7B-Instruct (Chat completion), databricks-dolly (Text generation), Deci-DeciLM-7B (Text generation), gpt-35-turbo (Text generation), llama-12b (Text generation), microsoft-phi-2 (Text generation), llama-finetune-12b (Text generation), gpt-4 (Text generation), tiuae-falcon-40b (Text generation), and gpt-4 (Text generation). A footer note states 'Image may not reflect actual user interface.'

New in Azure AI Studio

Azure AI Studio general availability

Model catalog & benchmarking
generally available

Assistants API

Azure OpenAI Service: GPT 4 Turbo
with Vision generally available

Prompt flow: Debugging
& tracing in preview

New MaaS partners and
models in catalog in preview

Quick-start templates integrated with Azure
Developer (AZD) CLI

HiddenLayer partnership

Integration with AI Toolkit
for Visual Studio Code

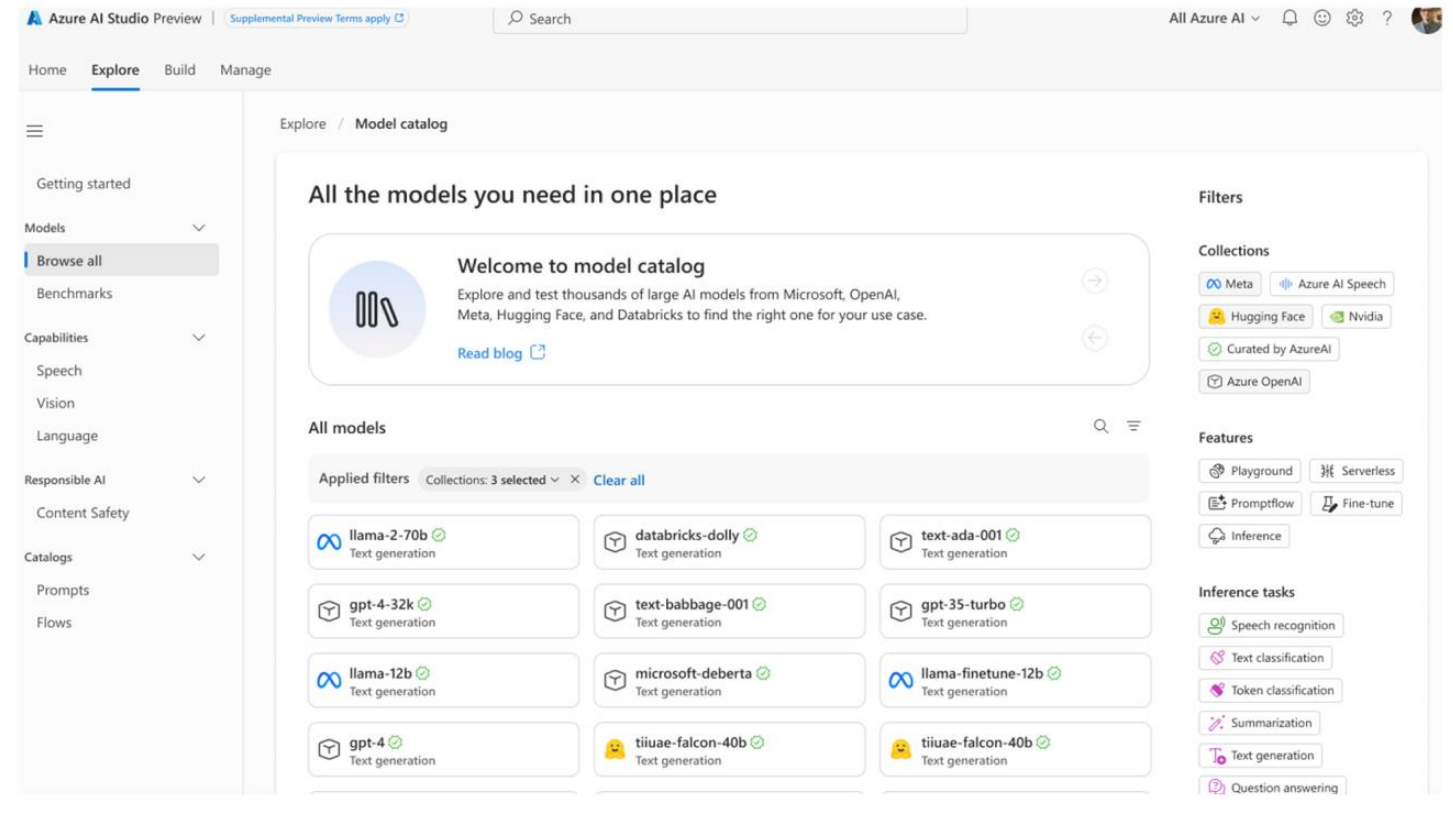
and more!

Azure AI Studio: choice of models

1400+ models from Azure OpenAI Service, Hugging Face, Meta and more.

Open-source models ready with curated fine-tuning pipelines.

Compare models by task using open datasets



Azure AI Studio: prompt flow

The screenshot displays the Azure AI Studio interface for configuring a prompt flow. The top navigation bar includes 'Build', 'BRK203-AI-Hub / ContosoOutdoor / Flows / contoso-chat', and 'Runtime'. The main workspace is divided into several panels:

- Flow Configuration:** Shows the 'contoso-chat' flow with a 'GPT4Vision' connection and a 'Prompt' section containing a system message and a task description.
- Graph:** A visual representation of the flow's execution path, showing nodes for 'ImageEmbedding', 'ImageSearch', 'GPT4Vision', 'CreateTextEmbedding', 'TextSearch', 'GPT4Turbo', and 'ReturnOutput'.
- Inputs/Outputs:** A table defining the flow's data flow:

Name	Type	Default value	Chat input	Chat history
chat_history	list			
question	list	Please input		
customerid	int	Please input		

Name	Reference	Chat output
answer	\$(ReturnResult.output)	
- Code Editor:** Shows the Python code for the 'CreateImageEmbedding' step, including imports for 'clip-embedder' and 'myconn'.

Orchestrate existing APIs, data stores, and machine learning models

Test, evaluate, and refine model behaviors

Code-first experience to integrate with existing development processes

Demo: Azure AI Studio



Identify the application pattern

RAG

Function/API

Agent

	RAG	Function/API	Agent
SCENARIO <i>example</i>	One well-defined task (Q&A) Known outcomes (answers) <i>looks for answers in product documentation</i>	Multiple tasks Controlled outcomes (calls) <i>query an API to validate account number</i>	Complex task Open outcomes <i>chain multiple APIs to solve a problem</i>
MODEL(S)	Question answering (understanding/summarization)	Intent detection and planning	Multiple
ORCHESTRATION	Systematic workflow	Flexible Resilient <i>(code is the trigger)</i>	Dynamic Self-healing <i>(agent is the trigger)</i>

Agent

An AI tool to accomplish complex tasks. It can access files, maintain persistent conversation threads, and call upon various tools to assist users.

Azure AI Studio / All hubs / build-demo-hub / All projects / build-demo-project / Assistants playground

← To all projects

Current project
build-demo-project

Project overview
Settings

Get started

Model catalog
Model benchmarks
Prompt catalog

Project playground

Chat
Assistants PREVIEW
Images
Completion

Tools

Code PREVIEW
Prompt flow
Tracing
Evaluation PREVIEW
Fine-tuning PREVIEW

Components

Data
Indexes
Deployments
Content filters

Assistants playground

Connected resource: aoaidanielscwed... </> View Code

Assistant setup

+ New Save Open Delete

1 Deployment ⓘ

gpt-35-turbo-1106

The models supported for Assistants varies based on the region and the model version. [Learn more](#)

Assistant name* ⓘ

Contoso Assistant

Assistant id: asst_yyXy5rt91an2YG6fRlxI0G8F

2 Instructions ⓘ

You are a helpful assistant that helps the user potentially with the help of some functions.

If you are using multiple tools to solve a user's task, make sure to communicate information learned from one tool to the next tool.

First, make a plan of how you will use the tools to solve the user's task and communicated that plan to the user with the first response. Then execute the plan making sure to communicate the required information between tools since tools only see the information passed to them; They do not have access to the chat history.

If you think that tool use can be parallelized (e.g. to get weather data for multiple cities) make sure to use the multi-tool use parallel function to execute.

3 Assistant tools

Functions ⓘ + Add function

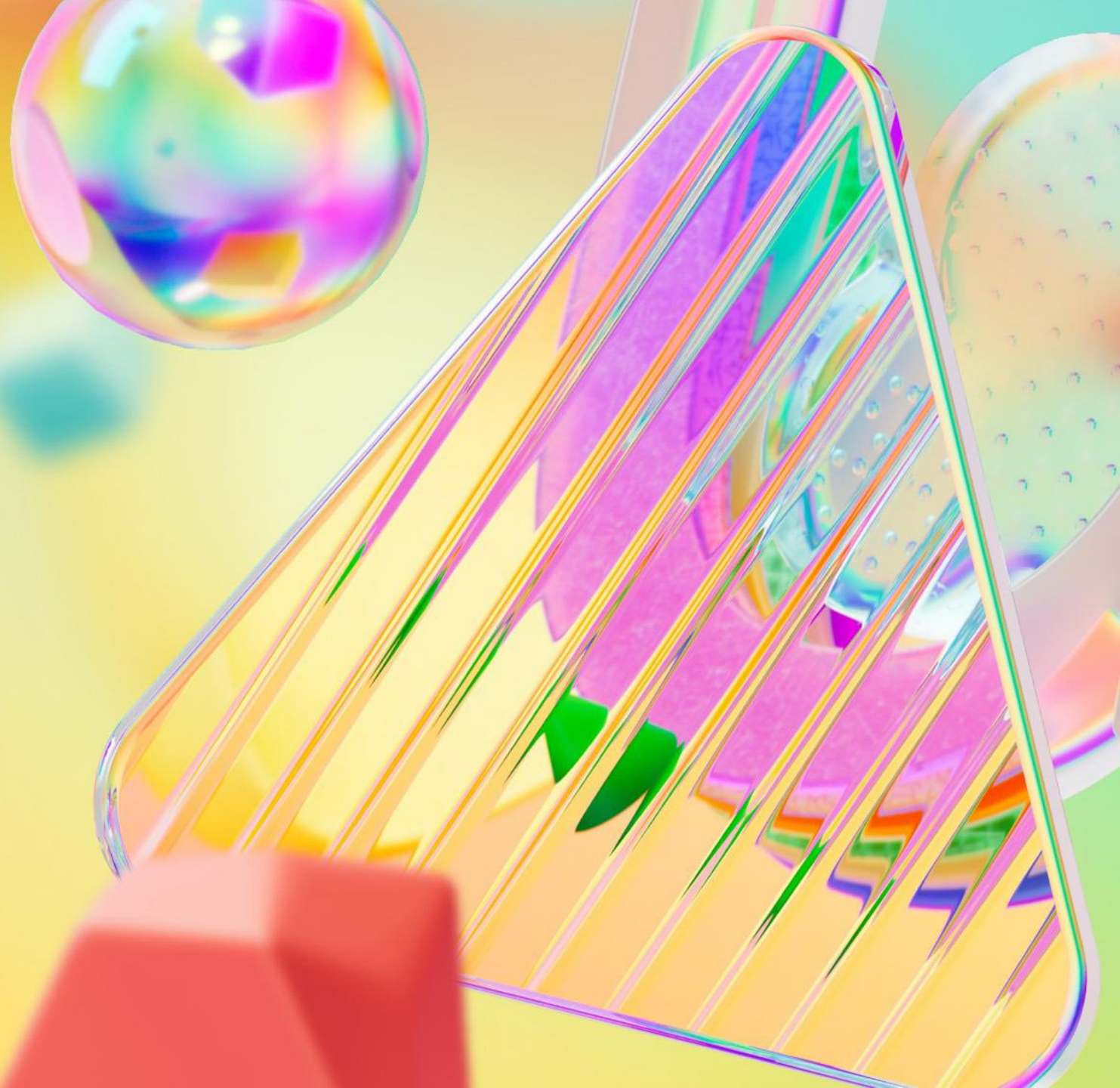
sales_data_insights

Code interpreter ⓘ

Files ⓘ + Add files

Upload files to use with the Assistant tools. You can upload up to 20 files, with a max size of 512 MB. [Learn more](#)

Copilot Studio





Copilot Studio

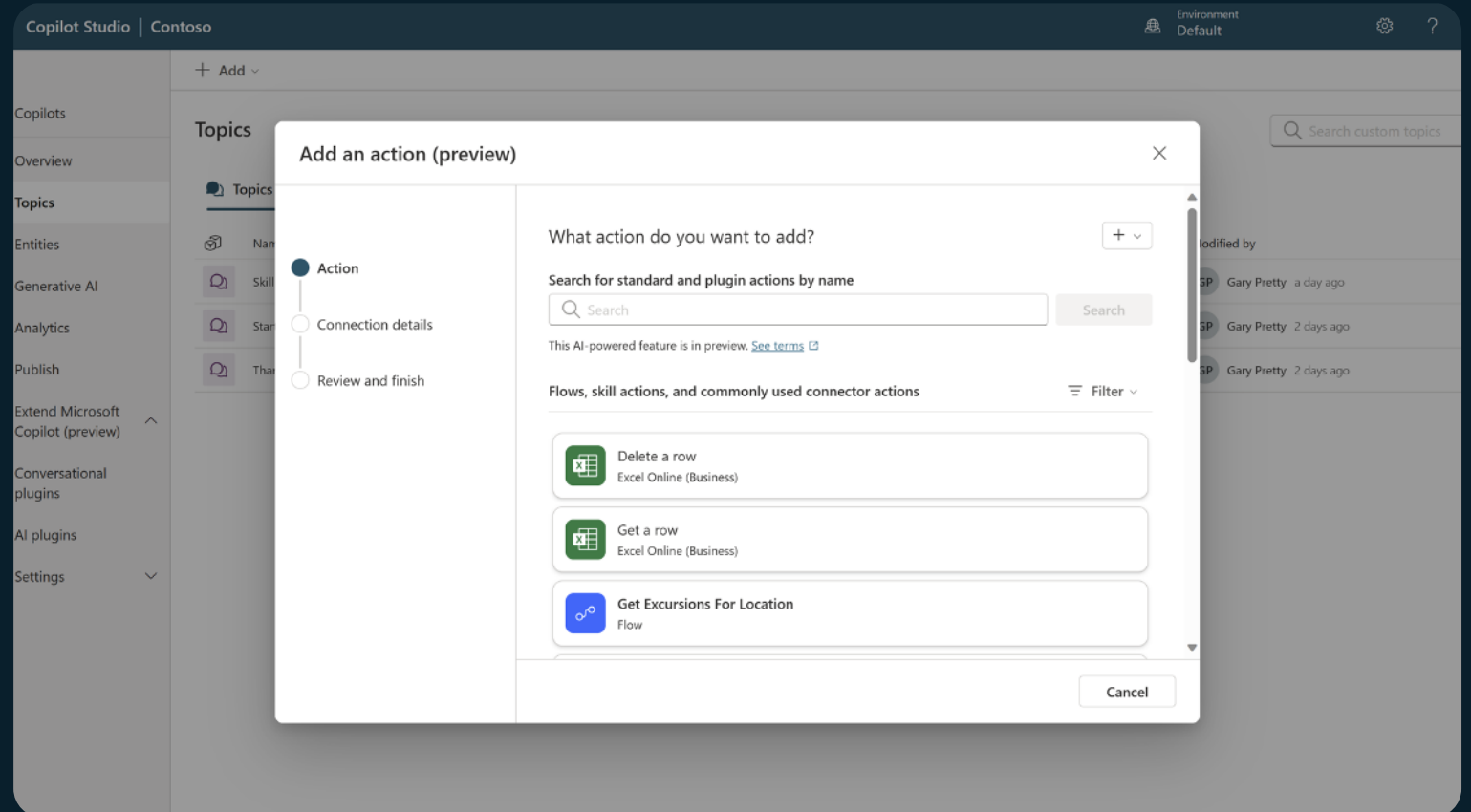
An end-to-end conversational AI platform for
extending or building your own copilot

Copilot Studio: 1500+ connectors – act!

Act on your data by using one of the 1500+ connectors

If your service doesn't have a connector, you can build your own!

Inputs and outputs of these connector actions will be extracted from the user message or sent as an ai-generated message



Copilot Studio: extend with Azure

AI Search

Azure OpenAI's models enable you to build powerful conversational experiences over your data with Copilot Studio.

Analytics

Build custom analytics by moving your data into extended data pipelines.

API Calls

Copilot Studio supports any API / HTTP calls to Azure services.

Language Services

Connect Conversational Language Understanding, including dialog triggering, interruptions, did-you-mean, and slot filling.



Bot Framework

Access Bot Framework Composer capabilities natively in Copilot Studio, call Azure Bot Framework skills from existing, and update and extend bots.

Application Lifecycle Management

Works with Azure DevOps for full ALM and solution management.

Telemetry

Use Azure app insights for custom telemetry on copilot usage.

Knowledge Base

Bring Azure Custom Question Answering into your copilot.

Copilot Studio: extend first-party Copilots

Customize Copilot for Microsoft 365

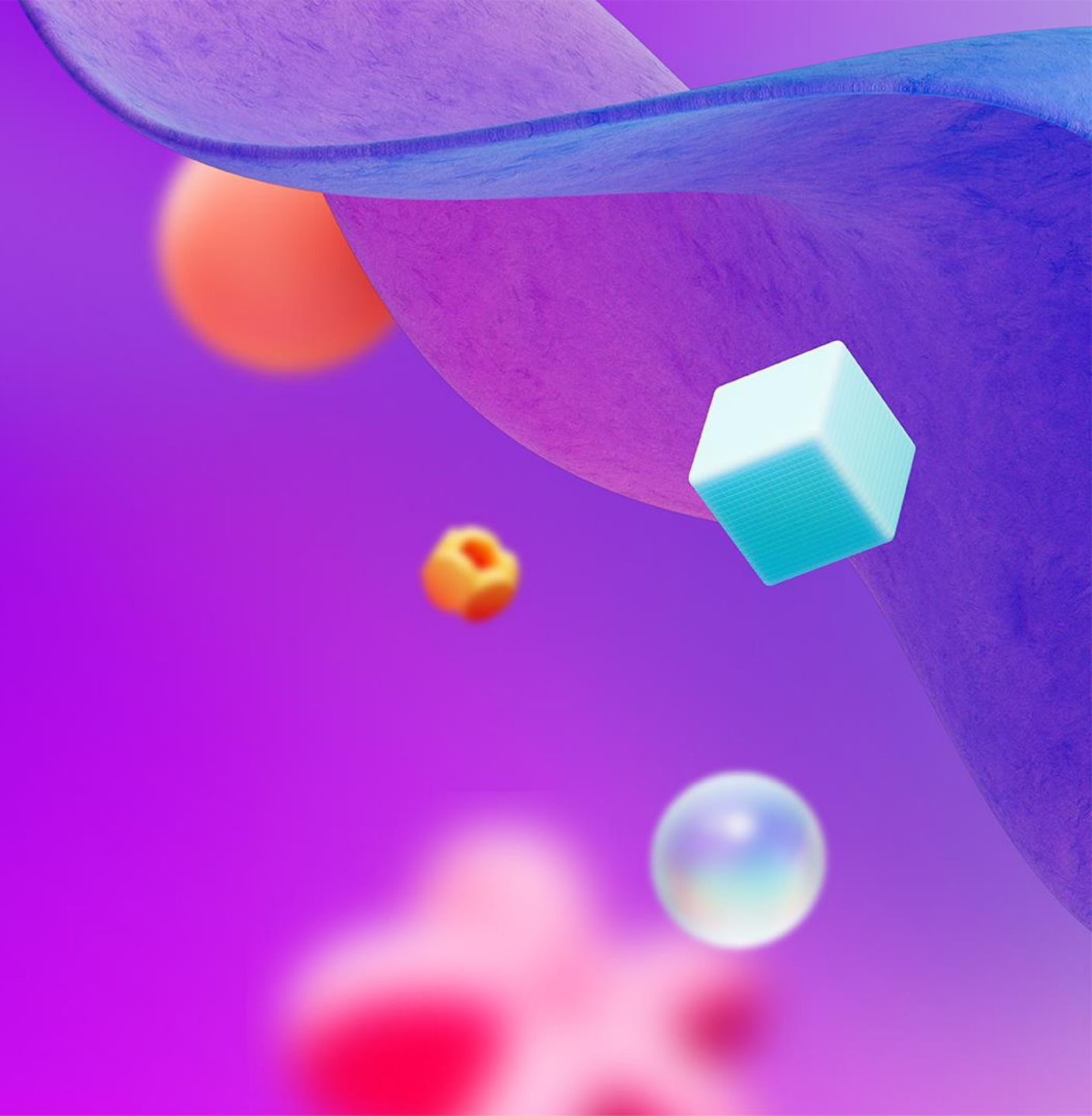
Customize Copilot to work with your unique business processes and data sources

Configure Copilot responses to specific questions

Build a Copilot extension



Demo: Copilot Studio



The real question is...

When to use what?

Build copilots **your** way.



Microsoft Copilot Studio



Azure AI Studio



Azure AI Studio and Copilot Studio

Use the option that **fits** the scenario **the best**

Best of both worlds

Azure AI Studio

In Azure AI Studio you can **choose the best model** for the job

Fast and frictionless **prototyping for quick results**

Azure AI Studio offers **state of the art evaluation, safety** and testing tools for your LLM application

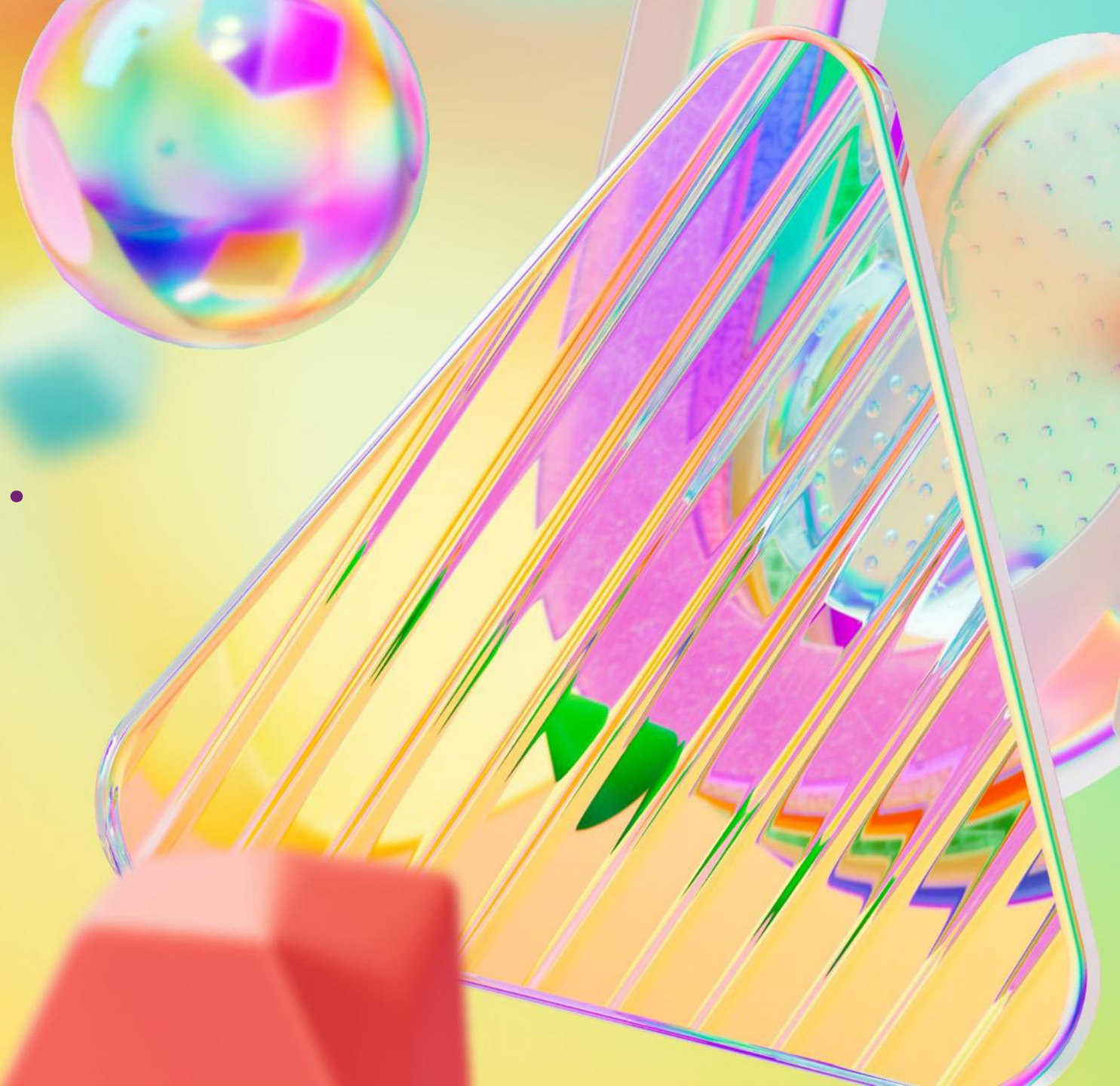
Copilot Studio

1400+ connectors available out-of-the-box + you can create your own connector

Channels like **Microsoft Teams, Facebook, and Slack** are easy to configure

Enterprise dialog management and conversational orchestration

Last but not least...
Responsible AI



Microsoft's responsible AI principles



Fairness



Reliability
& safety



Privacy &
security



Inclusiveness



Transparency

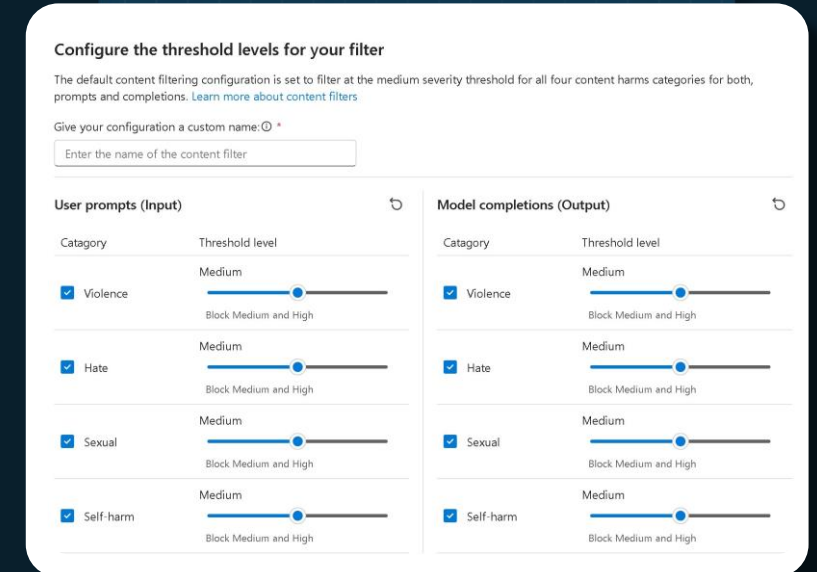


Accountability

Pioneering new engineering tools and practices



Evaluation and RAI dashboard



Content safety

Thank you!!!



Lorenzo Barbieri
lorenzo.barbieri@softwareone.com

Connect with me on LinkedIn



[LinkedIn.com/in/geniodelmale](https://www.linkedin.com/in/geniodelmale)