



AZURE OPENAI SERVICES

Dall-e and Conclusions



DALL-E

DALL-E

VALUE PROP FOR DALL-E

Generate original and diverse visual content that matches the user's input, without requiring any pre-defined templates or categories.

It can combine different concepts in novel and creative ways e.g a giraffe wearing a suit.

It can be used for art, design, entertainment, education and research

DALL-E

DALL-E 1 & 2

DALL-E 1

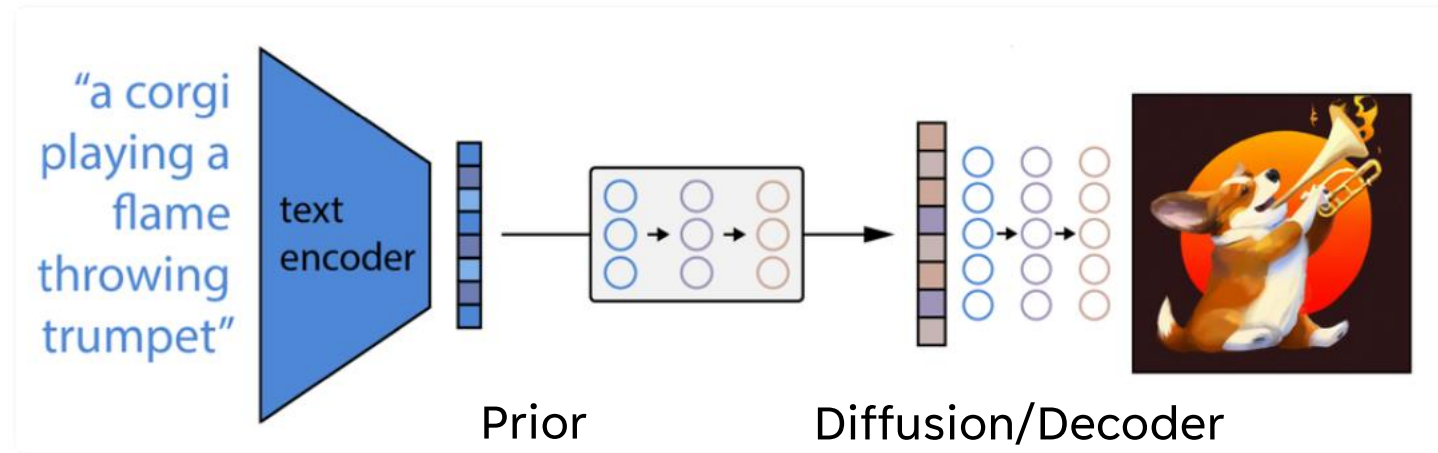
- Dall-E 1 generates realistic visuals and art from simple text
- The resolution of the images produced by Dall-e is 1024 pixels

Dall-E 2

- Dall-E 1 generates realistic visuals and art from simple text
- The resolution of the images produced by Dall-e is 1024 pixels
- DALL-E 2 discovers the link between visuals and the language that describes them. It employs a technique known as “diffusion”, which begins with a pattern of random dots and gradually changes that pattern to resembles a picture when it recognizes particular characteristics of that image
- The resolution of the images produced by Dall-E 2 is 4x better than its

DALL-E

HOW DALL-E WORKS?



**Contrastive Language-Image
Pre-training (CLIP)**

GLIDE - *invert* the image
encoding process in order
to stochastically decode
CLIP image embeddings.

DALL-E

DALL-E 2 LIMITATIONS

It Can't generate violent content, explicit images, political content, and other sensitive images.
(shooting stars, blood diamond)

Dall-E 2 can do scenes with generic backgrounds (a city, bookshelves in a library, a landscape) but even then, if that's not the main focus of the image then the fine details tend to get pretty scrambled

Position of objects relative to which other objects is also not good

DALL-E 2 can't count the number of objects if they are main focus, maximum it can handle 2 or 3 objects

a duck crossing finish line



A boy with seven cats



DALL-E

DALL-E 2 LIMITATIONS

Dall-E can't write the correct spellings in the image

Dall-E can't draw realistic human faces

BCIT written on blue wall





OVERALL USE CASES

REAL WORLD SCENARIOS

CONTACT CENTER ANALYTICS

Post-call analysis of transcript

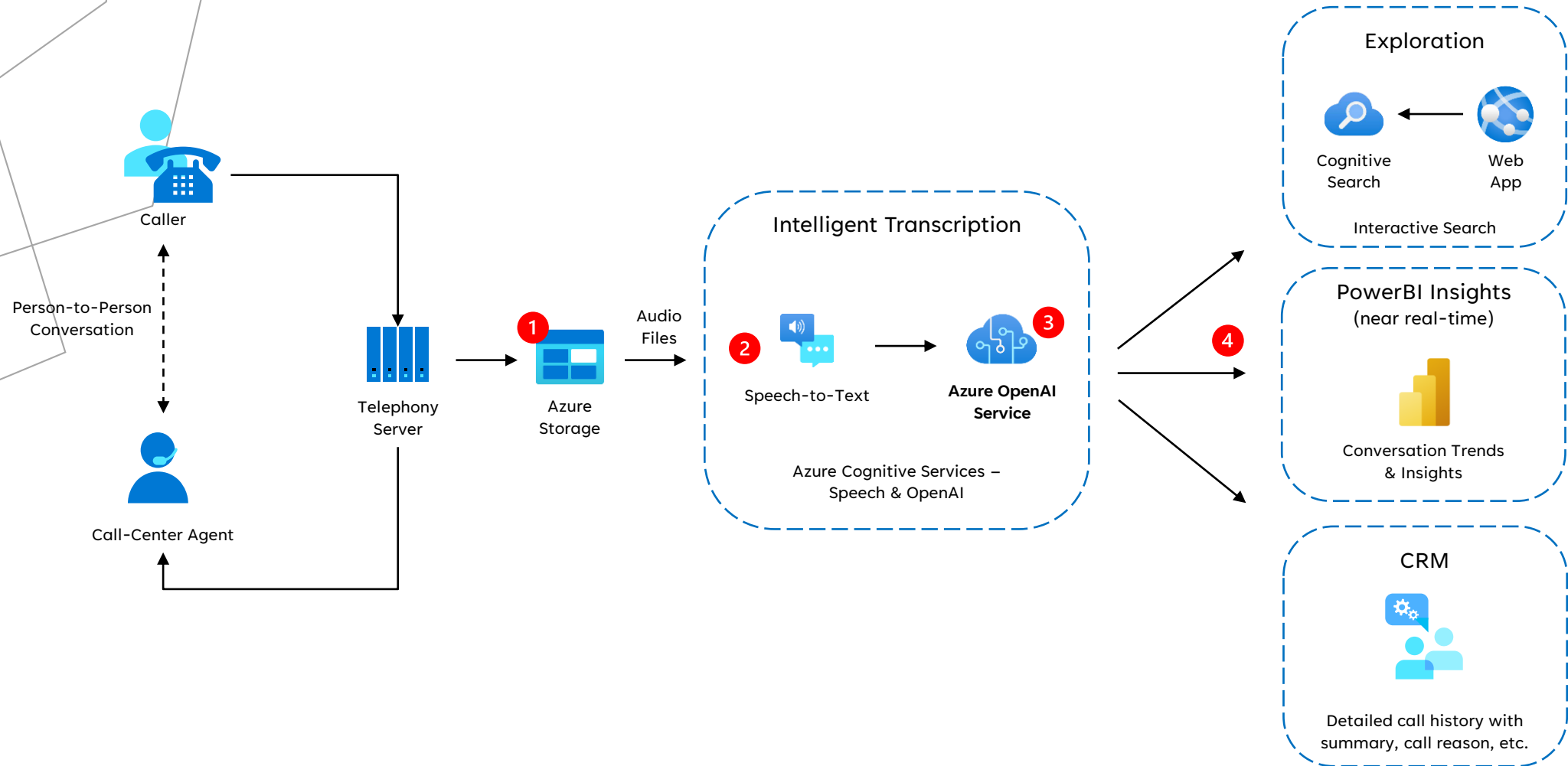
- Summarization (including outcome)
- Sentiment Analysis
- Entity extraction (customer identification, reason for call, campaign/product mentions)
- Automatic call auditing
- Search through call contents
- Reporting/dashboard

Support for individual call analysis or volume call processing.

Can be expanded for real-time in-call analysis scenarios.

REAL WORLD SCENARIOS

CONTACT CENTER ANALYTICS



REAL WORLD SCENARIOS

ENTERPRISE SEARCH Q&A

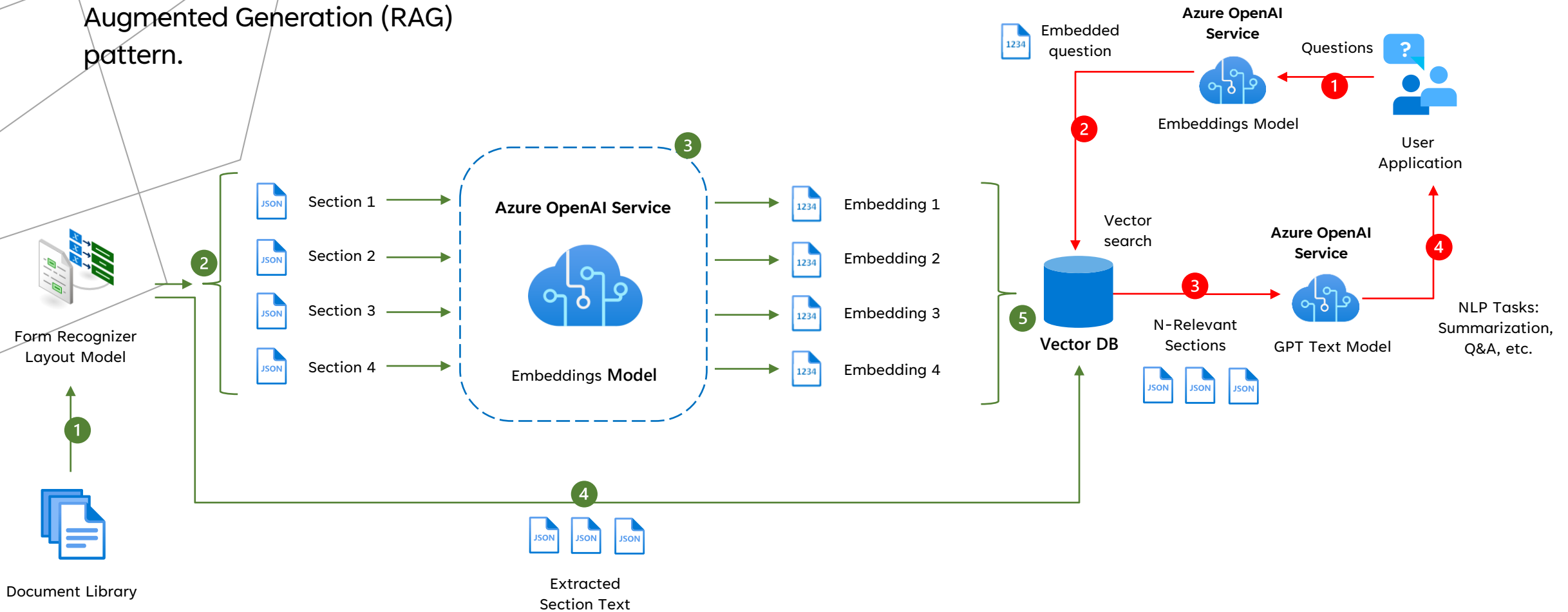
Natural language interface to surface internal enterprise data

- Leverage internal knowledge from multiple data sources
- Combine results in a human-like natural language answer
- Use citations to explain where information came from and allow user to access source
- Support follow up questions

REAL WORLD SCENARIOS

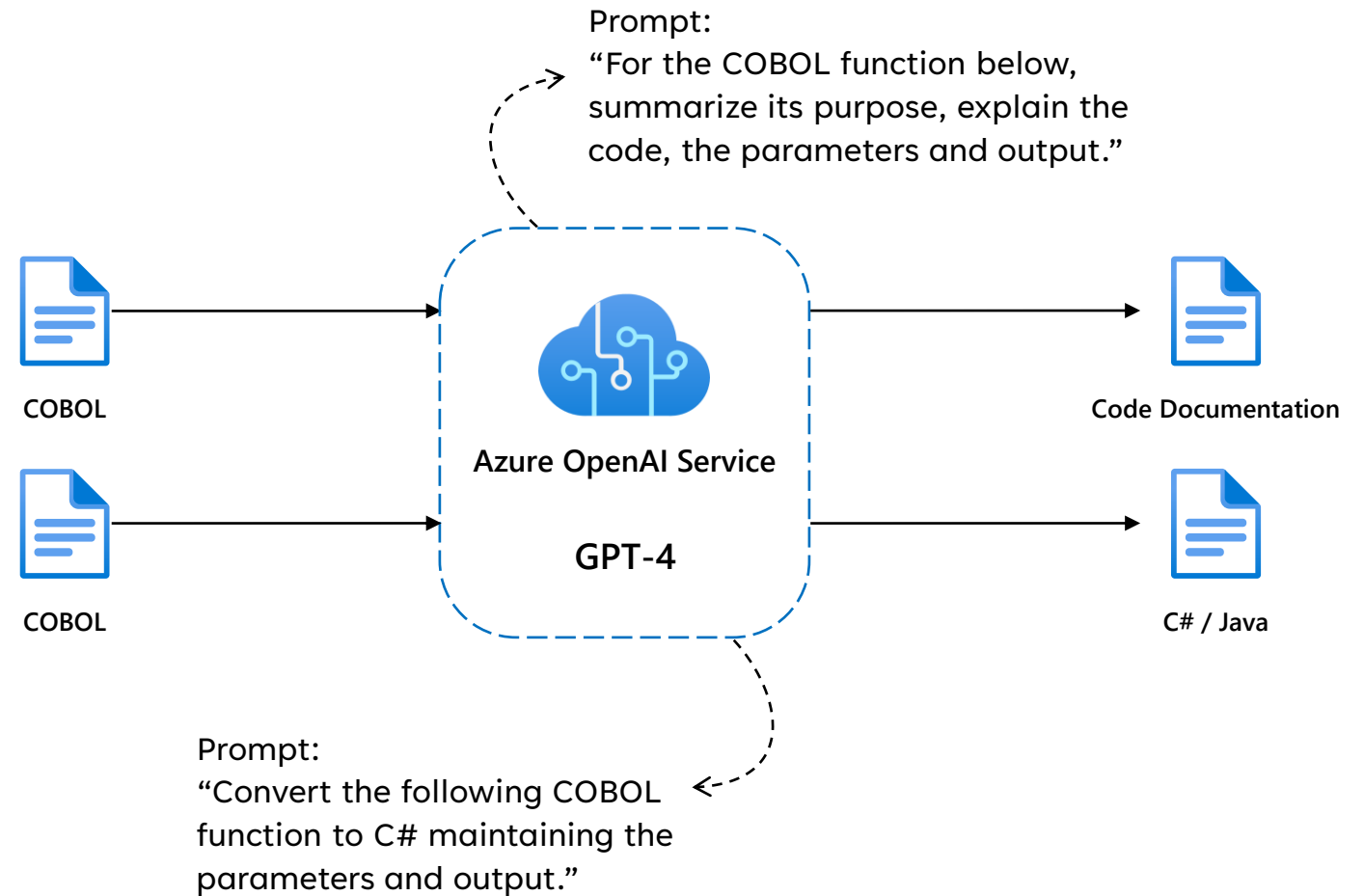
ENTERPRISE SEARCH Q&A

This is also known as Retrieval Augmented Generation (RAG) pattern.



CODE GENERATION

- Generate comments and documentation for legacy code
- Explain complex algorithms in accessible natural language
- Translate code from one programming language to another
- Support for multiple programming languages
- **Human validation is still very important**





ADVANCED USE CASES

PROMPT CHAINING

Entity Extraction

Please **extract entities** from the following news article: 'The new iPhone model is set to be released next month. It has been highly anticipated by Apple fans and is expected to feature a larger screen and improved camera'

```
[{'text': 'iPhone', 'type':  
'PRODUCT'}, {'text': 'Apple',  
'type': 'ORGANIZATION'}]
```

Summarization

Please **summarize** the information about the product

```
[{'text': 'iPhone', 'type':  
'PRODUCT'}, {'text': 'Apple',  
'type': 'ORGANIZATION'}]
```

The iPhone is a line of smartphones designed and marketed by Apple Inc. These devices combine a computer, camera, and internet connectivity in a single device, and are known for their sleek design and user-friendly interface.

Sentiment Analysis

Please provide a **sentiment** for the following text:

The iPhone is a line of smartphones designed and marketed by Apple Inc. These devices combine a computer, camera, and internet connectivity in a single device, and are known for their sleek design and user-friendly interface.

Positive

ADVANCED USE CASES

DOCUMENT CHUNKING



Large Text
Beyond
Token
Limit



Chunks

Please summarize the following text: 'Climate change is a pressing issue that needs to be addressed immediately. The increasing levels of carbon dioxide in the atmosphere are causing the Earth's temperature to rise, which is leading to more frequent and severe natural disasters. Governments around the world need to take immediate action to reduce carbon emissions and invest in renewable energy sources. Failure to do so will have dire consequences for future generations'

Climate change is a pressing issue that needs to be addressed immediately. Governments around the world need to take immediate action to reduce carbon emissions and invest in renewable energy sources. Failure to do so will have dire consequences for future generations

Please summarize and extract topic: Global warming refers to the gradual increase in the overall temperature of the Earth's atmosphere, primarily caused by the burning of fossil fuels such as coal, oil, and natural gas. This burning releases carbon dioxide and other greenhouse gases into the atmosphere, which trap heat and cause the Earth's temperature to rise.

Climate change is a pressing issue that needs to be addressed immediately. Governments around the world need to take immediate action to reduce carbon emissions and invest in renewable energy sources. Failure to do so will have dire consequences for future generations

The text is discussing global warming, which is the gradual increase in the overall temperature of the Earth's atmosphere caused by the burning of fossil fuels. This burning releases greenhouse gases, such as carbon dioxide, into the atmosphere which trap heat and cause the temperature to rise. The text also highlights the urgency of the issue of climate change and the need for governments to take action to reduce carbon emissions and invest in renewable energy sources. Failure to address this issue will have negative consequences for future generations.

Topic: Climate change and Global warming.

Continue to
next prompt

MODEL REASONING – CHAIN-OF-THOUGHT

Before

What is the annual water demand of a single-family household containing four people who are at home an average of 200 days per year and use an average of 100 liters of water per day?

The annual water demand of a single-family household containing four people who are at home an average of 200 days per year and use an average of 100 liters of water per day is 8,000,000 liters per year.

After

What is the annual water demand of a single-family household containing four people who are at home an average of 200 days per year and use an average of 100 liters of water per day?

Let's think step by step and explain the calculation step by step.

Step 1: Calculate the total number of days in a year:

There are 365 days in a year.

Step 2: Calculate the total number of days the family is at home:

The family is at home an average of 200 days per year.

Step 3: Calculate the total water demand of the family:

The family uses an average of 100 liters of water per day. Therefore, the total water demand of the family is $100 \text{ liters} \times 200 \text{ days} = 20,000 \text{ liters per year}$.

Step 4: Calculate the annual water demand of the single-family household:

The annual water demand of the single-family household is 20,000 liters per year.

ORCHESTRATION

- Uses the LLM as a controller that, when given a task in natural language, plans the necessary steps to complete it and executes them.
- The steps can be invoking other models (expert models) or executing actions through plug-ins.

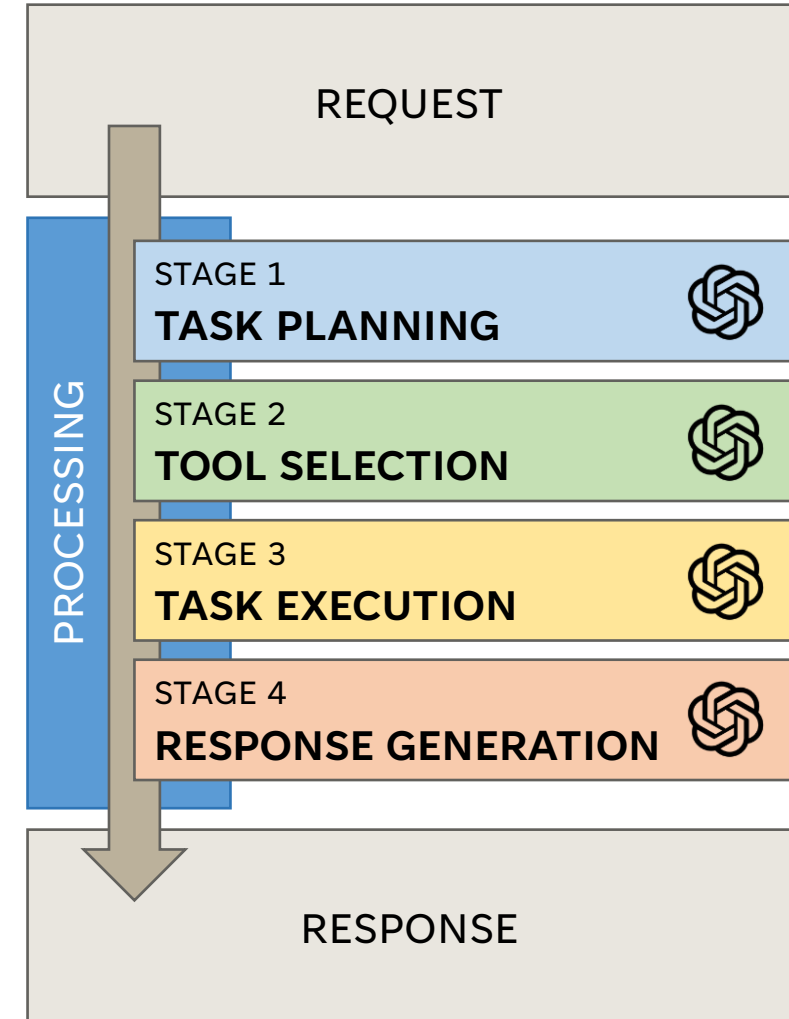
Applications

AutoGPT

<https://github.com/Significant-Gravitas/Auto-GPT>

HuggingGPT / Microsoft JARVIS

<https://github.com/microsoft/JARVIS>





TOOLS

TOOLS

LANGCHAIN

Open Source **framework** for developing applications powered by language models that are **data-aware** (can connect to data sources) and **agentic** (can interact with the environment).

Main Concepts

- **Modules** (modular abstractions) – components that handle common tasks such as prompts, indexes and memory.
- **Chains** – assembly of components put together to accomplish a use case.
- **Agents** – special type of chain that can decide which tools to use depending on the user input.
- **Tools** – functions that agents can use to interact with the world. Can be generic utilities, chains or other agents.

Features

- Huge ecosystem and community
- Support for multiple LLMs, document loaders, text splitters, retrievers and vector stores
- Offers several use cases ready to implement
- Support for Python and JavaScript

SEMANTIC KERNEL

Open Source **SDK**, lead by Microsoft, to integrate LLMs with conventional programming languages infusing applications with complex skills like prompt chaining, recursive reasoning, memory, semantic indexing and more.

Main Concepts

- **Kernel** – orchestrator of a user's Ask (goal).
Fulfills the goal using available skills, memories and connectors.
- **Planner** – component that breaks down the goal into steps based upon available resources.
- **Skill** – domain of expertise made available to the kernel as a function or group of functions.
- **Memory** – provides context for the Ask. Can be accessed as key-value pairs, local storage or semantic memory search.

Features

- Highly extensible
- Support for multiple models, skills and connectors
- Support for C#, Python and JavaScript

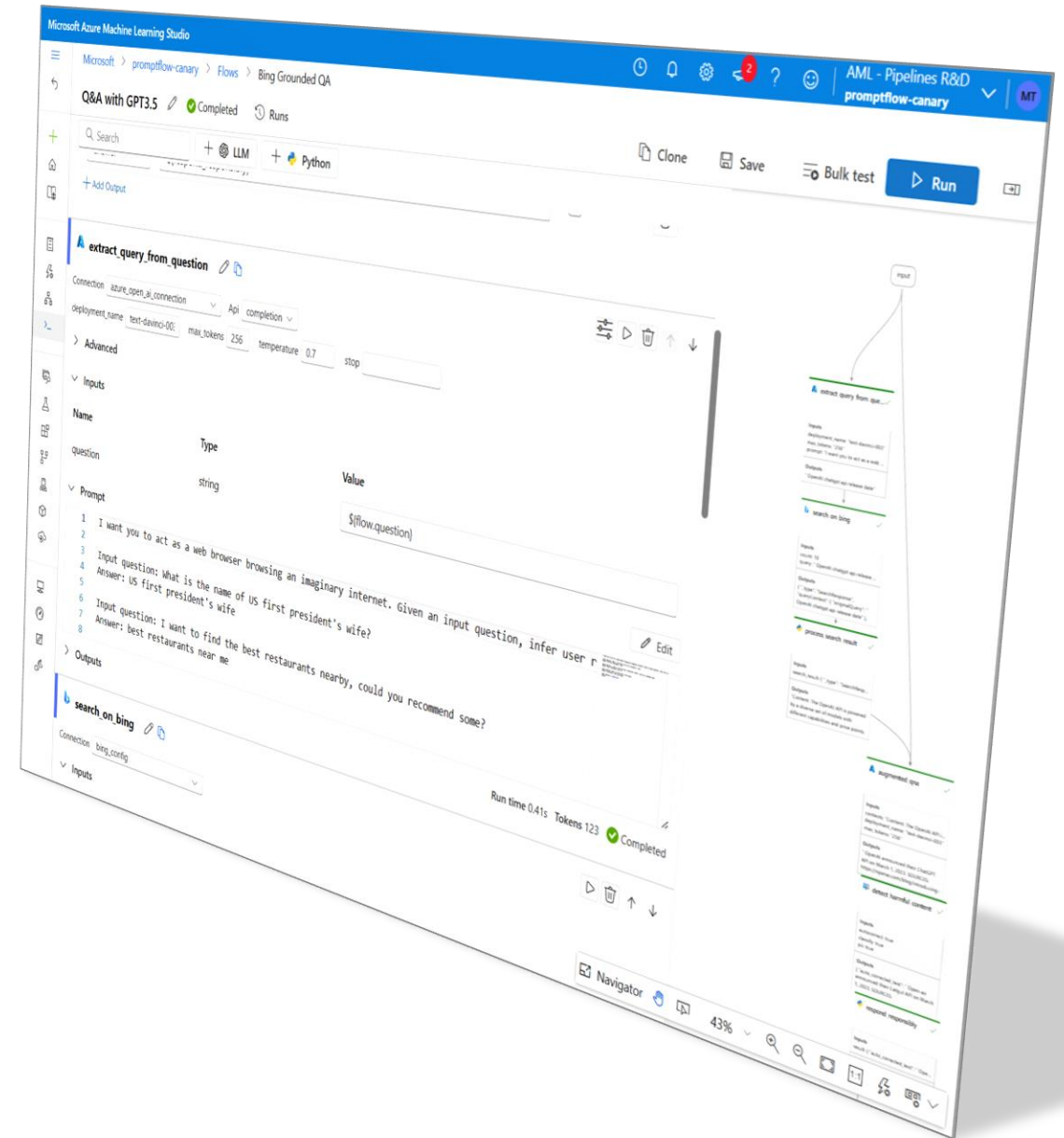
TOOLS

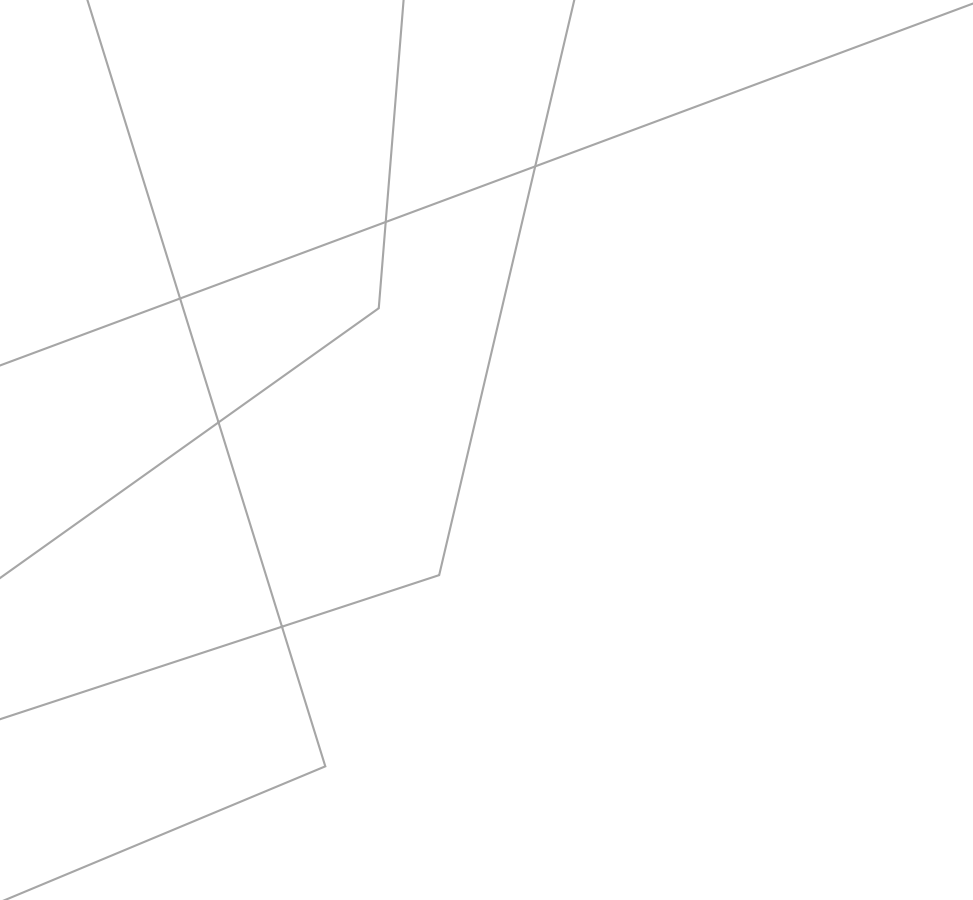
PROMPT FLOW

Prompt engineering tool included in Azure Machine Learning Studio.

Main Features

- Create AI workflows that connect to various language models and data sources.
- One platform to build, tune, evaluate, deploy, and test AI workflows.
- Visual representation of prompt workflows using graphs. Each step is mapped to a notebook-like code cell.
- Evaluate the quality of AI workflows with rich set of pre-built metrics.
- Easy prompt tuning, comparison, tracking.

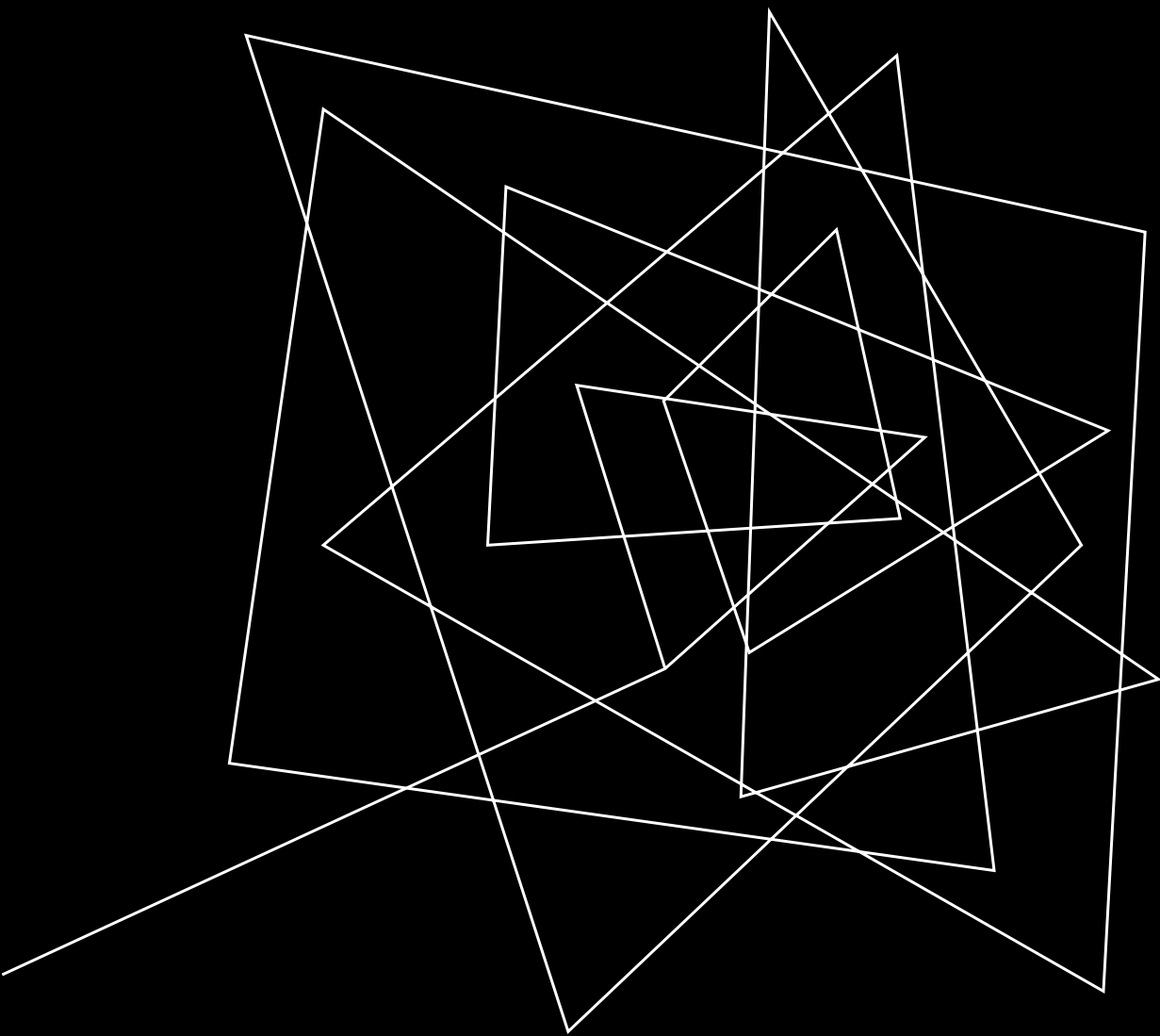




PROMPT FLOW

1. Connections
2. Workflows
3. Debugging

DEMO #5



CONCLUSIONS

CONCLUSIONS

DATA PRIVACY & SECURITY

Your data is **your** data

Data is stored encrypted in your Azure subscription

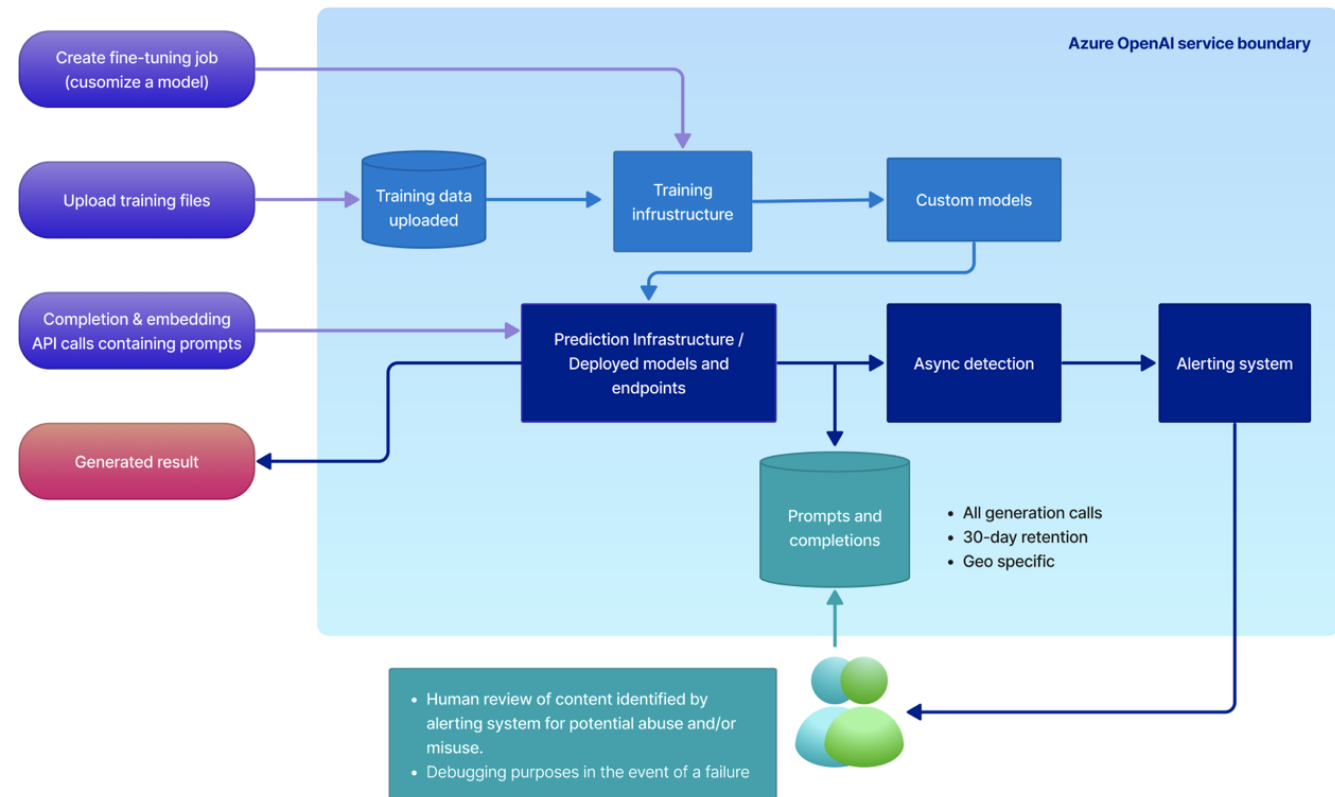
Your data from any fine-tuning is **not used** to train the foundation AI models

- Azure OpenAI Service provisioned in your Azure subscription
- Model fine tuning stays in your Azure subscription and never moves into the foundation AI models

Your data is **protected** by the most comprehensive enterprise compliance and security controls

- Encrypted with Customer Managed Keys
- Private Virtual Networks, Role Based Access Control
- Soc2, ISO, HIPPA, CSA STAR Compliant

Customer action



CONCLUSIONS

WHEN TO USE AZURE OPEN AI SERVICE



I need a general-purpose model that can handle multiple tasks.

e.g. translation + entity recognition + sentiment analysis



I need to generate high-quality content, while preserving data privacy and security

e.g. abstractive summarization, content writing, paraphrasing, code



I need rapid prototyping and quick time to market for many use cases



I could use a model with little or no additional training



I want to explore solutions / use cases that have been described previously



Vision



Speech

Azure OpenAI Service



Language



Decision

Azure Cognitive Services

CONCLUSIONS

KEY TAKEAWAYS

- Infusing AI into applications has **huge potential**.
- Azure Open AI Service is an **enterprise-ready** offering on top of Open AI models, bringing Azure's security, compliance and Responsible AI principles to solutions.
- Azure Open AI Service is usually just a **piece of the puzzle**. A complete solution will require additional components.
- Large Language Models, such as GPT, can execute **different types of tasks with few or no training**.
- **Prompt engineering** is critical, but it is more of an art than a science. It can yield better results than Fine Tuning at a much lower cost.
- **Tools and frameworks** are evolving fast. Leverage them to develop more complex use cases.

CONCLUSIONS

RESOURCES

Azure Open AI Service documentation

<https://learn.microsoft.com/en-us/azure/cognitive-services/openai/>

Azure Open AI Samples

<https://github.com/Azure/azure-openai-samples>

Sample: ChatGPT with Enterprise Data

<https://github.com/Azure-Samples/azure-search-openai-demo/>

Sample: Azure Open AI Embeddings Q&A

<https://github.com/Azure-Samples/azure-open-ai-embeddings-qna>

Tool: LangChain

<https://docs.langchain.com/docs/>

Tool: Semantic Kernel

<https://github.com/microsoft/semantic-kernel>



THANK YOU