

# Azure and Open AI:

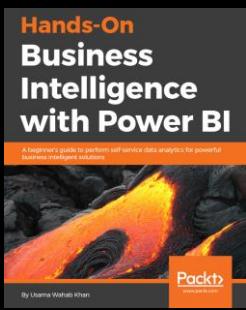
## Partners in transforming the world with AI

Fundamental





Microsoft®  
Most Valuable Professional  
**Reconnect**



# Usama Wahab Khan

Father, data Scientist, Developer/Nerd, Traveler



Twitter : @usamawahabkhan  
LinkedIn : Usamawahabkhan





# Today's Opportunity – Your Opportunity

# Agenda

Topic	
<b>Introduction to AI</b>	
Current State of AI	
Types of AI	
Popular AI Tools	
<b>Azure OpenAI Service Overview</b>	
<b>Introduction to AOAI Models</b>	
<b>Responsible AI</b>	
<b>Introduction to OpenAI</b>	<b>Module # 1</b>
<b>Why OAI vs. AOAI</b>	
<b>Demo</b>	
<b>Deep Drive Architecture for Azure OpenAI</b>	
Different Azure Open AI Architecture	
Exploring Text Completion in AOAI	
Overview of ChatGPT (Large Conversational Foundation Model)	
Overview of GPT-3 / GPT-4 (Large Language Foundation model)	
Overview of DALL-E (Large Image to Text Foundation Model)	
Overview of Codex (Large Code Foundation Model)	
Contact Center Analytics using Speech API & Azure OpenAI Service	
AI-Powered Q&A over Enterprise Data Sources	
<b>Available Azure OpenAI Accelerators</b>	<b>Module # 2</b>
<b>Overview and Technical deep dive of GPT -3/4 models</b>	
Model naming convention	
Understanding Tokens and Probabilities	
Key concepts: Natural Language Understanding and Generation examples (Zero-Shot, One-Shot, Few-Shot)	
ChatGPT Prompt Examples	<b>Module # 3</b>
Model Adaptation (Fine-Tuning)	
Embeddings	
Prompt Engineering Guide and Techniques	

# Introduction to AI



# Agenda

- **Introduction to AI**
  - Current State of AI
  - Types of AI
  - Popular AI Tools
  - **Azure OpenAI Service Overview**
- **Responsible AI**
- **Why OAI vs. AOAI**
- **Uses and Capabilities**
- **Demo**



Assigning human-like qualities to digital experiences



Perceives its environment



Mimics cognitive functions

# What is AI?

010100  
010101  
010100  
Learns from example in volumes of data

Program that writes itself based on examples

Classifies, recommends, predicts, groups, segments

Weak AI

Separate cognitive functions, seeing, natural language, vision

AI

ML

Strong AI

Combining weak AI with a consciousness or "mind"

# Artificial Intelligence, What is it?

## What is AI?

Simply put, AI is the creation of software that imitates human behaviors and capabilities. Key workloads include:

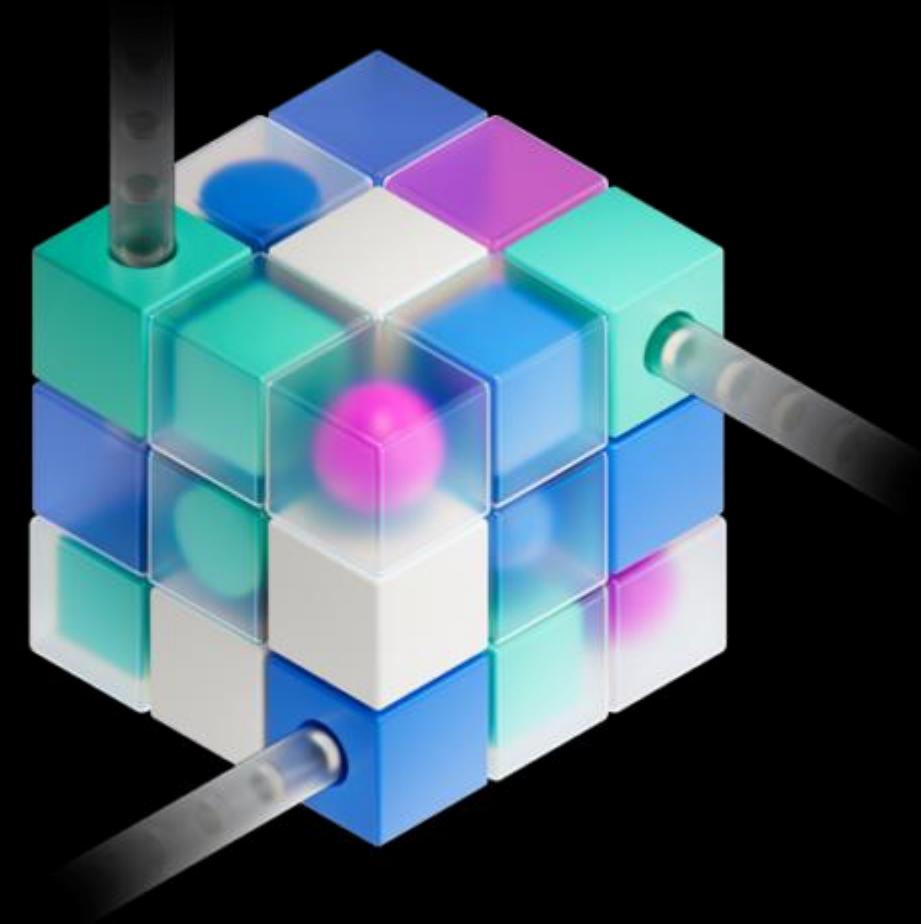
**Machine learning** - This is often the foundation for an AI system, and is the way we "teach" a computer model to make predictions and draw conclusions from data.

Anomaly detection - The capability to automatically detect errors or unusual activity in a system.

**Computer vision** - The capability of software to interpret the world visually through cameras, video, and images.

Natural language processing - The capability for a computer to interpret written or spoken language, and respond in kind.

**Knowledge mining** - The capability to extract information from large volumes of often unstructured data to create a searchable knowledge store.



# Types of AI

While there are various forms of AI as it's a broad concept, we can divide it into the following three categories based on AI's capabilities:

**Weak AI(Narrow intelligence)**, which is also referred to as Narrow AI, focuses on one task. There is no self-awareness or genuine intelligence in case of a weak AI.

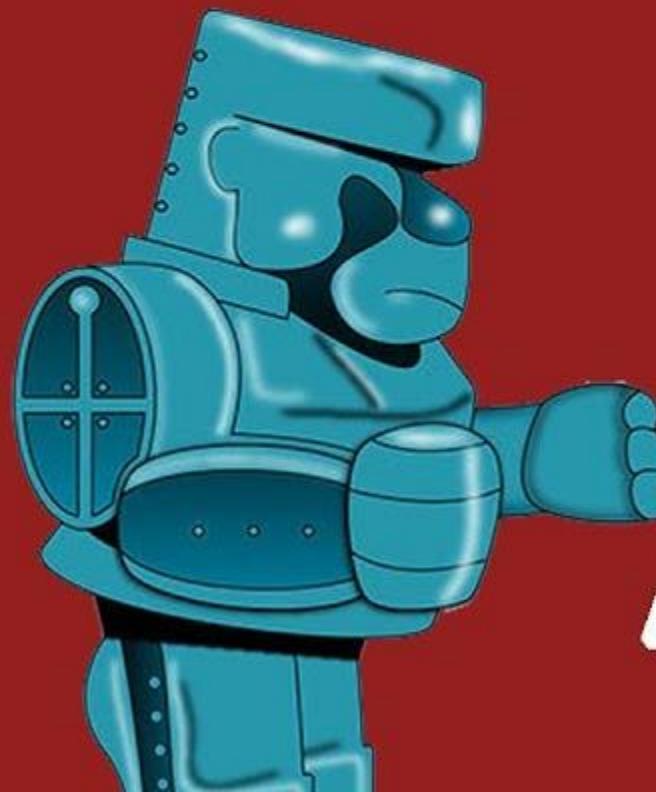
iOS Siri is a good example of a weak AI combining several weak AI techniques to function. It can do a lot of things for the user, and you'll see how "narrow" it exactly is when you try having conversations with the virtual assistant.

**Strong AI (General intelligence )**, which is also referred to as True AI, is a computer that is as smart as the human brain. This sort of AI will be able to perform all tasks that a human could do. There is a lot of research going on in this field, but we still have much to do. You should be imagining Matrix or I, Robot here.

**Artificial superintelligence** , Artificial superintelligence (ASI) is a form of AI that is capable of surpassing human intelligence by manifesting cognitive skills and developing thinking skills of its own.

# STRONG VS. WEAK

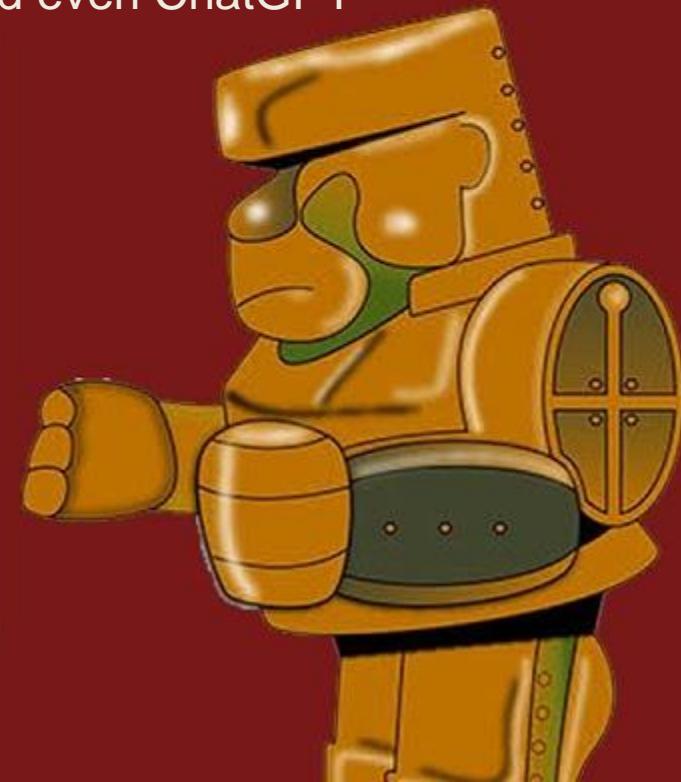
Strong AI is AI that acts exactly as a human would, think C-3PO, the Terminator or Commander Data. They exhibit emotions, real creativity, and can even have a sense of purpose.



# AI

Weak AI is AI that is confined to a narrow task, like when a system processes language into text or sorts all the pictures on your pc.

Examples of Weak AI include: Siri, Cortana, Bing, Netflix, and even ChatGPT



## Artificial Intelligence

### Machine Learning

### Deep Learning

### Generative AI



## Artificial Intelligence

the field of computer science that seeks to create intelligent machines that can replicate or exceed human intelligence



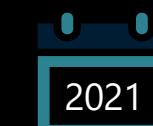
## Machine Learning

subset of AI that enables machines to learn from existing data and improve upon that data to make decisions or predictions



## Deep Learning

a machine learning technique in which layers of neural networks are used to process data and make decisions



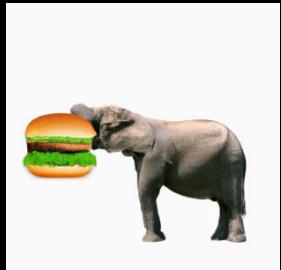
## Generative AI

Create new written, visual, and auditory content given prompts or existing data.

# Introduction to Generative AI

Generative AI refers to a type of artificial intelligence that has the ability to generate content that is, in many cases, indistinguishable from content created by humans. This AI can produce text, images, audio, or even video, often in response to a given input or prompt.

Generative AI operates by learning patterns and structures from large datasets and then using that knowledge to produce new content that fits within those learned patterns. It's a type of machine learning where the AI model learns to understand and mimic the characteristics of the data it has been trained on.



**Image or Video generation**

```
board = [' ' for x in range(9)]  
  
def print_board():  
    row1 = '| {} | {} | {} |'.format(board[0], board[1], board[2])  
    row2 = '| {} | {} | {} |'.format(board[3], board[4], board[5])  
    row3 = '| {} | {} | {} |'.format(board[6], board[7], board[8])  
  
    print()  
    print(row1)  
    print(row2)  
    print(row3)  
    print()
```

**Code generation**

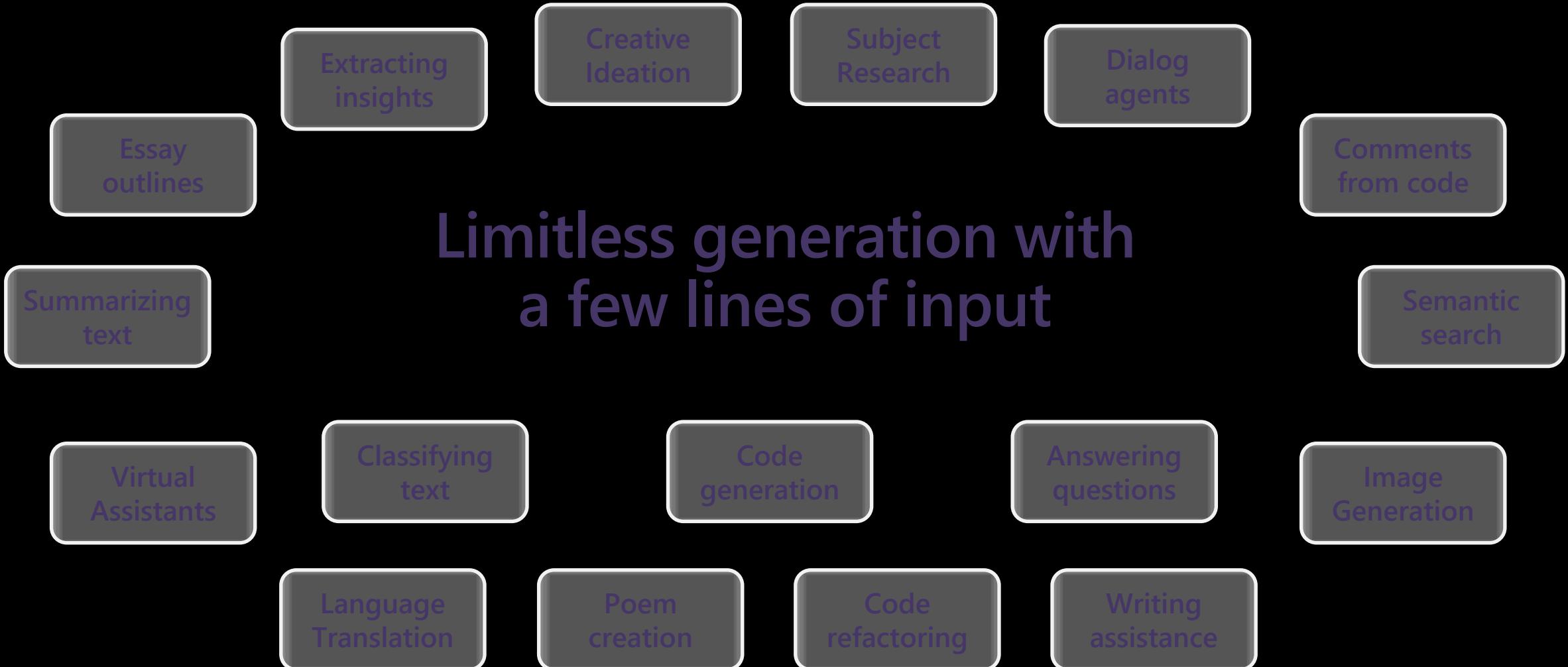
Artificial Intelligence

Machine Learning

Deep Learning

Generative AI

# Generative AI capabilities

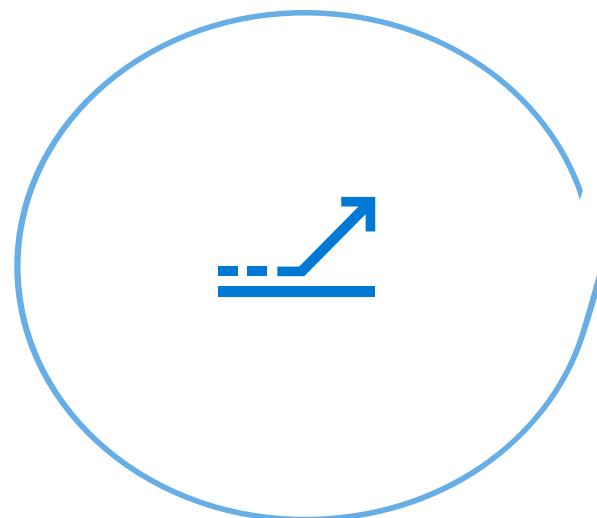


# Why AI? – Why Now?

***“AI is going to shape all of that we do.”***

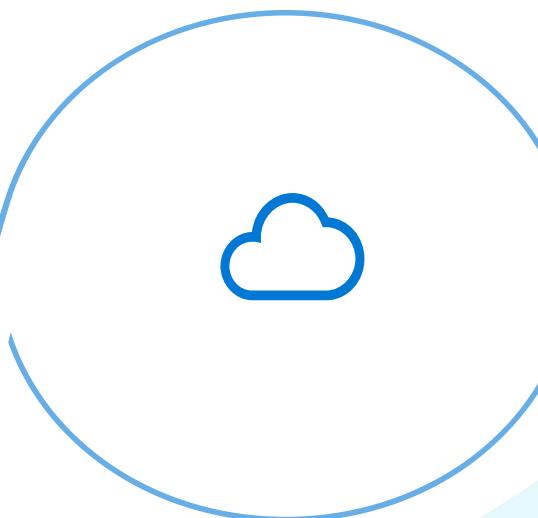
- Satya Nadella

\$1.2  
Trillion  
“Market for AI”  
by 2020  
—Forrester



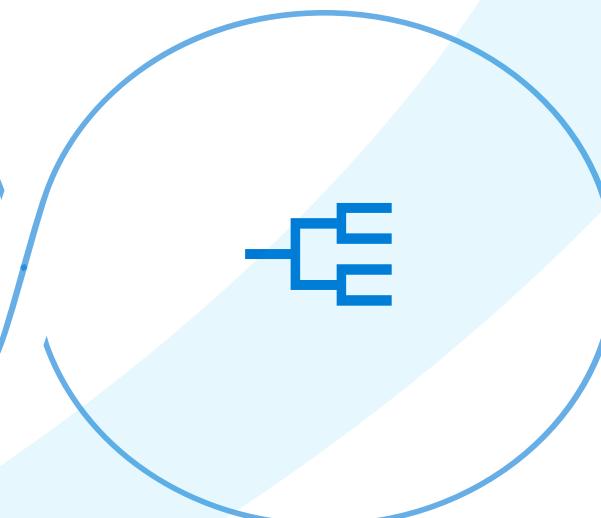
## Big Data

Incredible amounts of data are being generated every second, across every business and connected consumer device



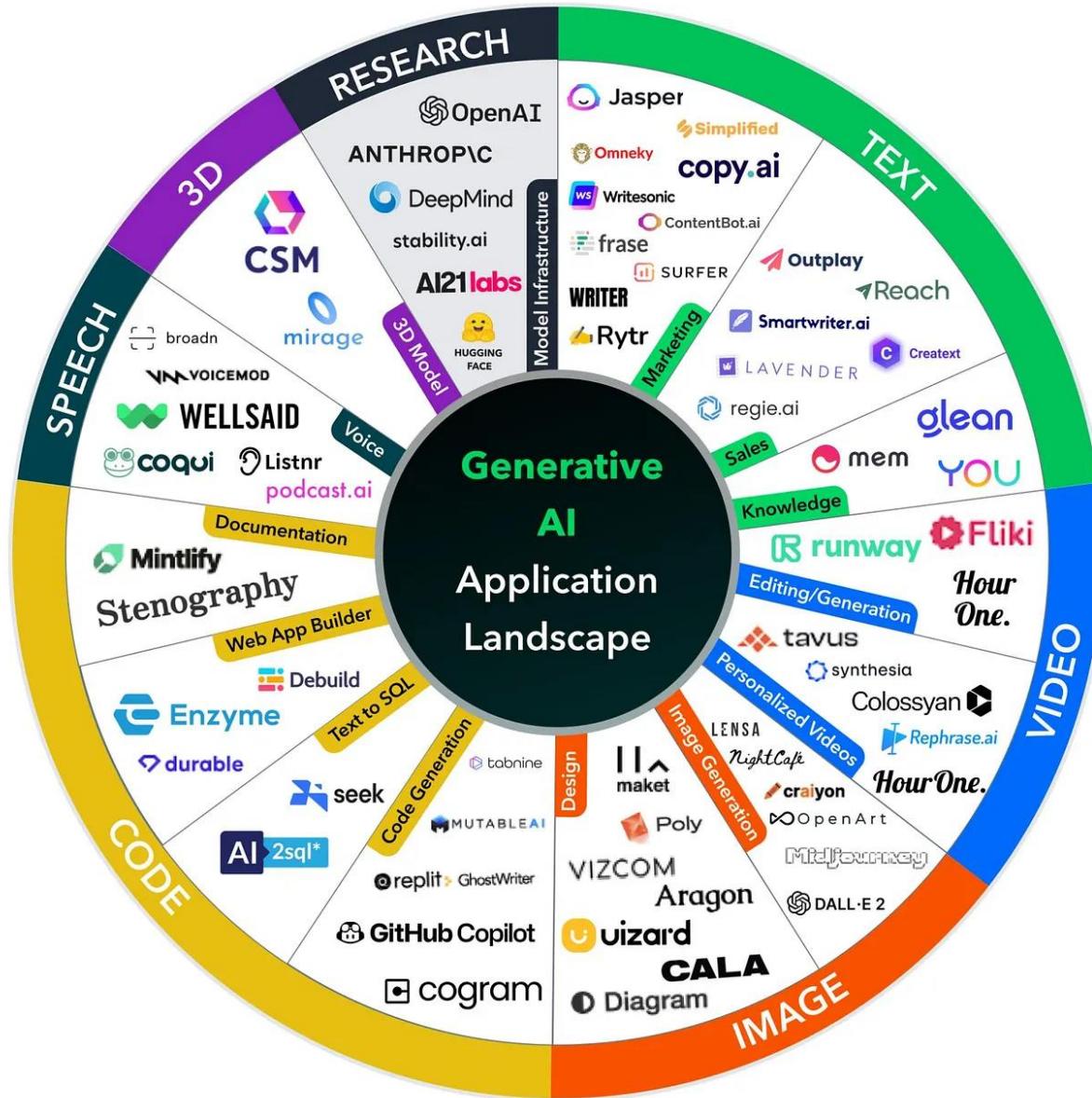
## Massive Compute Power

Availability of massive & specialized compute resources (especially on the cloud) to process big data with specialized algorithms



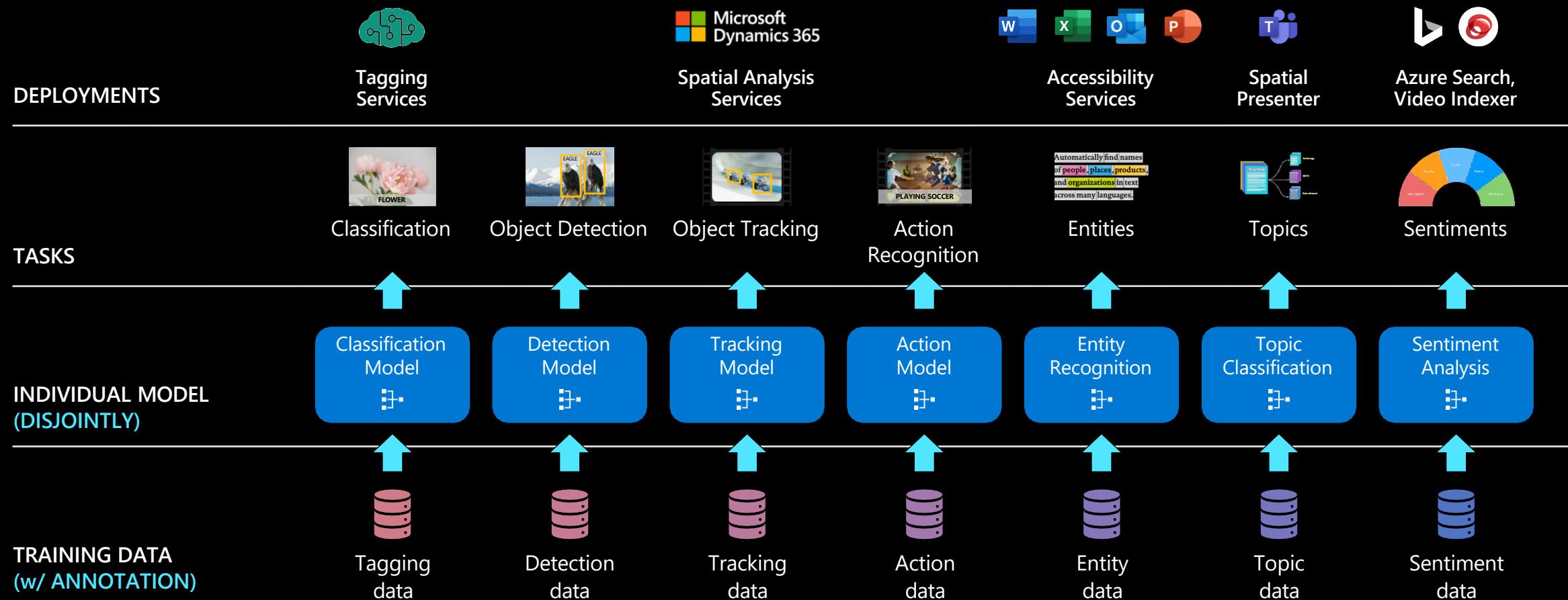
## Breakthrough Algorithms

Advancements in applying sophisticated mathematics & statistics has allowed algorithms tackle advanced tasks such as a neural network training or complex simulations



# Traditional model development

High cost and slow deployment—each service is trained disjointly



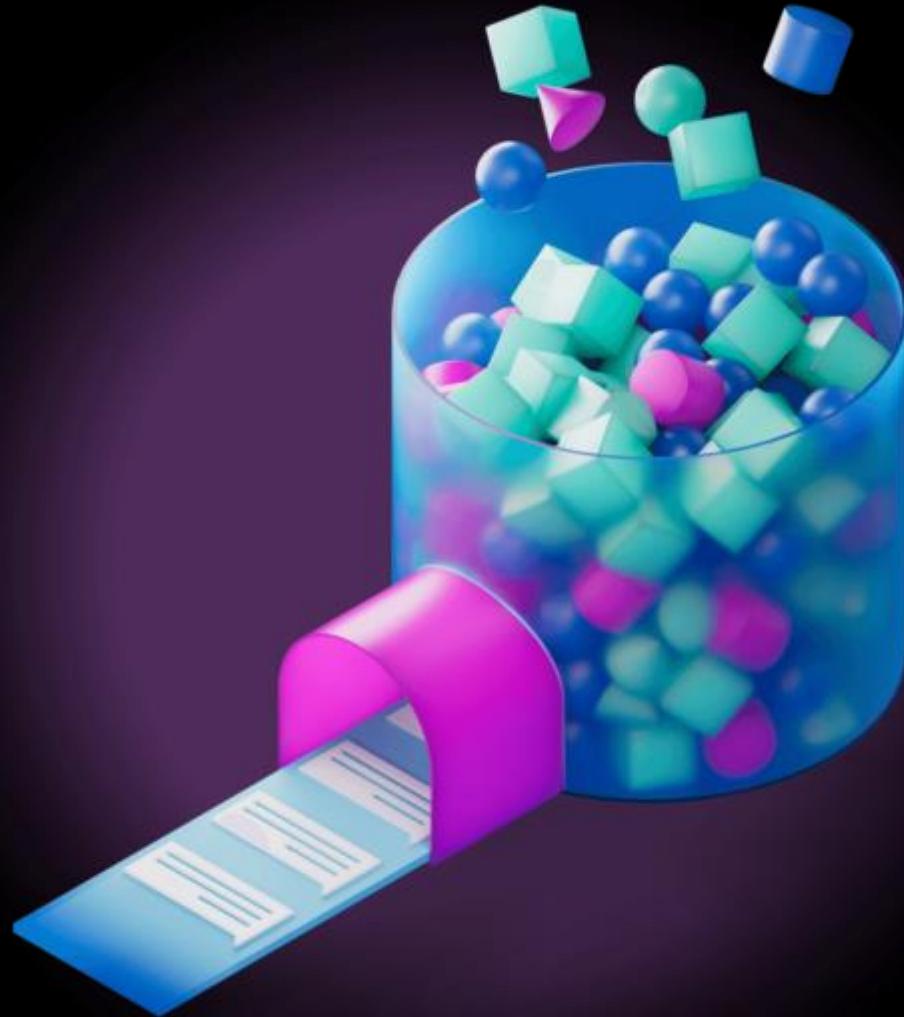
# 10 AI terms everyone should know

**Machine learning:** A field of computer science where people teach a computer system how to do something by training it to identify patterns and make predictions based on data

**Generative AI:** A type of AI that can make new things, not just provide information about existing things<sup>4</sup>. It can create things like pictures, music, text, videos and code.

**Large language models:** Computing systems that use machine learning techniques to process language and mimic the way humans communicate. They can do things like translate languages, answer questions, summarize text, and write stories.

**Responsible AI:** A set of principles and practices that guide people as they try to design AI systems that are safe and fair for everyone<sup>6</sup>. It involves understanding the data that was used to train the systems and finding ways to mitigate any biases or harms.



# 10 AI terms everyone should know

Generative Pre-trained Transformer (GPT)

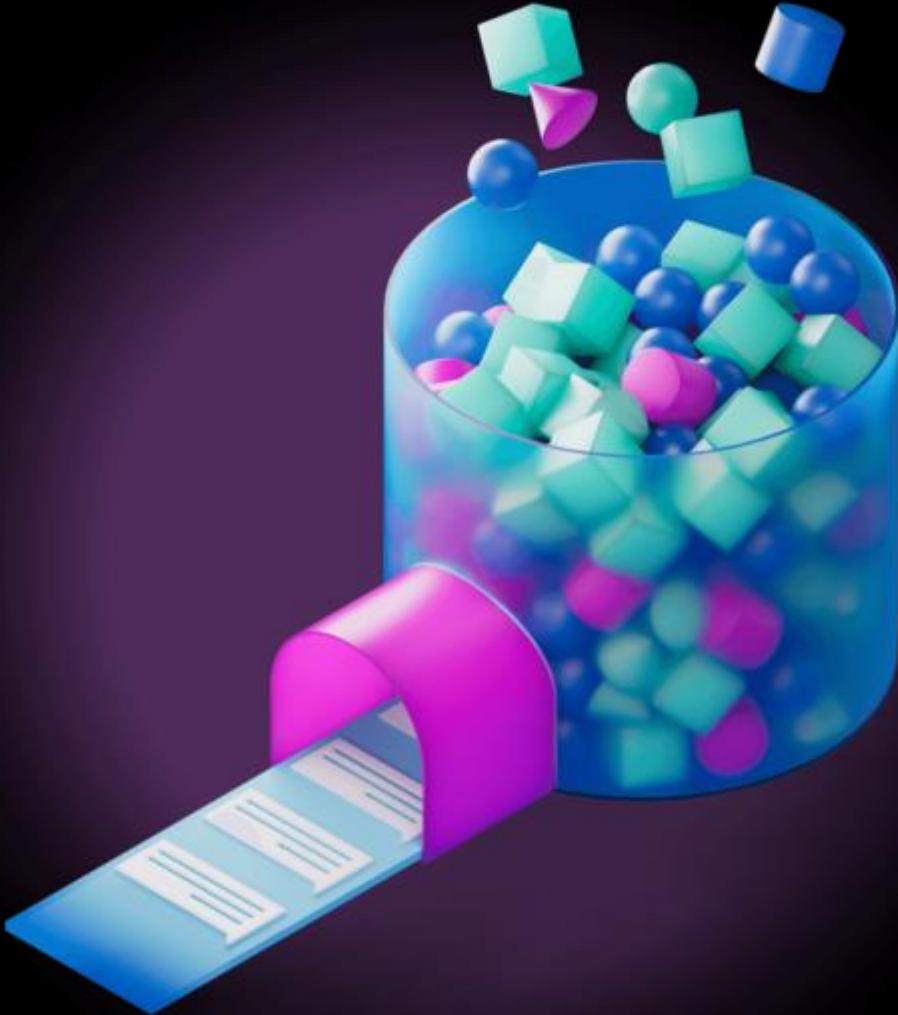
**Generative:** Generates new content

**Pre-Trained:** Trained on data before fine-tuning

**Transformer:** Understands relationships b/w words

"Few-shot" learning

Reinforcement Learning from Human Feedback (RLHF)



# 10 AI terms everyone should know

Extract the mailing address from this email:

Hi John Doe,

It was great to meet up at Build earlier this week. I thought the AI platform talk was great and I really enjoyed it.

I appreciate the offer for the book. If you are OK, you can mail it to me at home, or 123 Microsoft Way, Bellevue WA 92004.

Regards,

Chris Hoder

**Prompt**—Text input that provides some context to the engine on what is expecting.

**Completion**—Output that the model generates based on the prompt.

**Token** — partial or full words processed and produced by the GPT models

# 10 AI terms everyone should know

**Hallucinations:** When an AI model generates incorrect information but presents it as if it were a fact. This can happen due to insufficient, outdated, or low-quality training data, overfitting, use of idioms or slang expressions, or adversarial attacks.

**OpenAI:** A research organization that aims to create and promote artificial intelligence that can benefit humanity, without being constrained by profit or corporate agendas. It is backed by Microsoft and other prominent tech companies and individuals.

**GPT:** A series of large language models developed by OpenAI that use deep neural networks to generate natural language texts. The latest version, GPT-4, has 175 billion parameters and can perform various natural language tasks, such as answering questions, writing essays, and creating code.

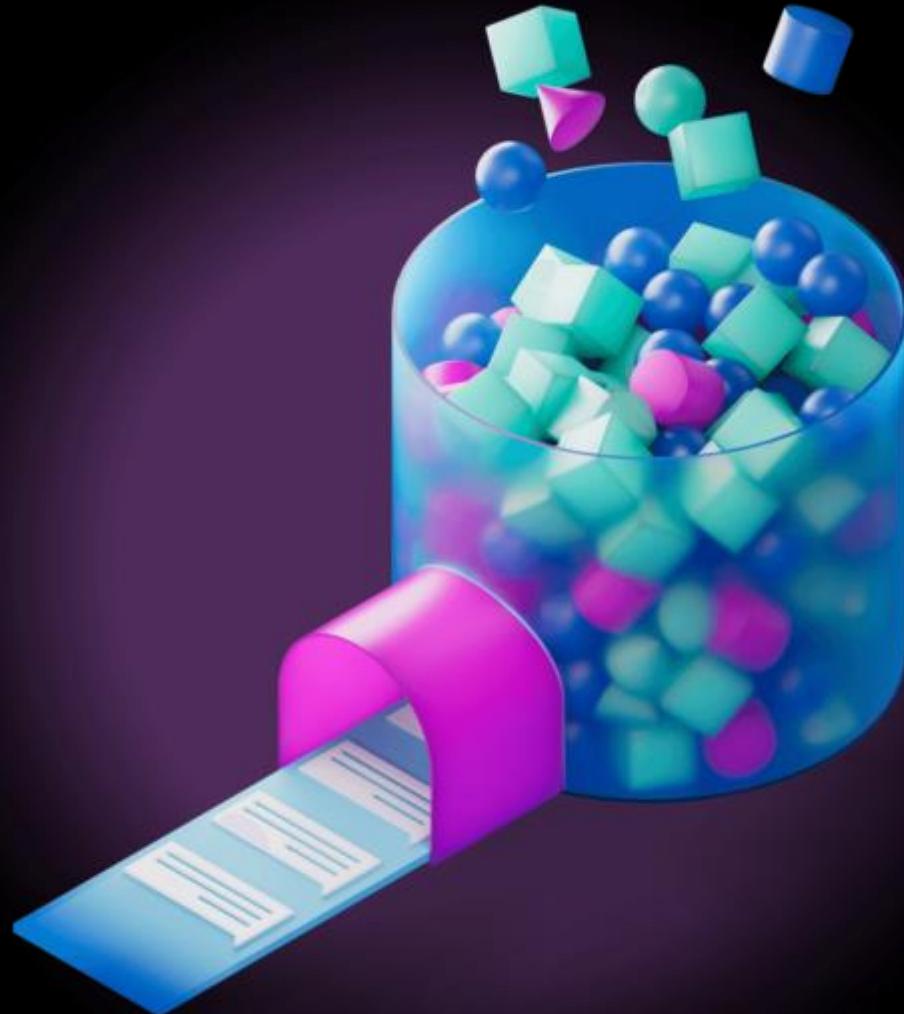
**ChatGPT:** A popular AI chatbot that runs on the GPT-3.5 model and can generate decent-quality content from a simple prompt given by the user. It has a simple user interface and some guardrails to prevent AI hallucinations and harmful outputs<sup>1</sup>.

**Gemini:** A rebranded version of Bard, a multimodal AI tool that can work with different types of data, such as text, images, and audio. It is integrated with the Google ecosystem and can help users with various tasks, such as creating presentations, editing videos, and composing music.

**Copilot:** A rebranded version of Bing Chat, an AI assistant that runs on the GPT-4 model and can help users with information, questions, and conversation. It is integrated with the Microsoft Edge browser and can also generate imaginative and innovative content, such as poems, stories, code, and songs<sup>1</sup>.

**Plugins:** Software components that can extend the functionality of an AI model or tool. They can help users customize their experience, access different features, or interact with other programs. They can also help developers fine-tune or improve their AI models.

**AI Model:** A mathematical representation of a real-world process or phenomenon that is created by an AI system using data and algorithms. An AI model can learn from data and make predictions or decisions based on it. An AI model can be trained, tested, evaluated, and deployed.



# What are large language models?

Large Language Models (i.e., "language calculators")

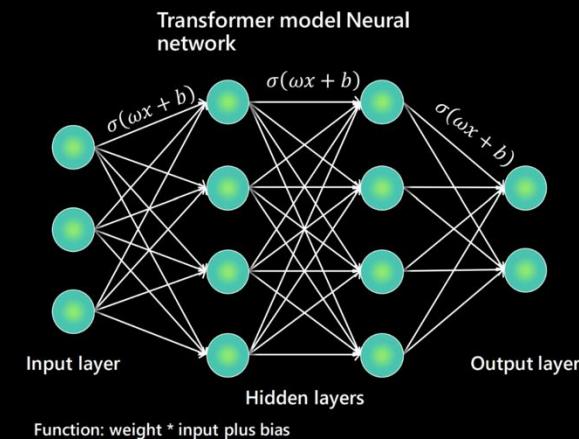
Large: More data than can be manually labeled

Language: Match context and words (e.g., word prediction, creative writing)

Model: Semi-supervised learning

A large language model (LLM) is a type of AI that can process and produce natural language text. It learns from a massive amount of data gathered from sources like books, articles, webpages, and images to discover patterns and rules of language.

How large are they?



BERT Large - 2018

**345M**

GPT2 - 2019

**1.5B**

GPT3 - 2020

**175B**

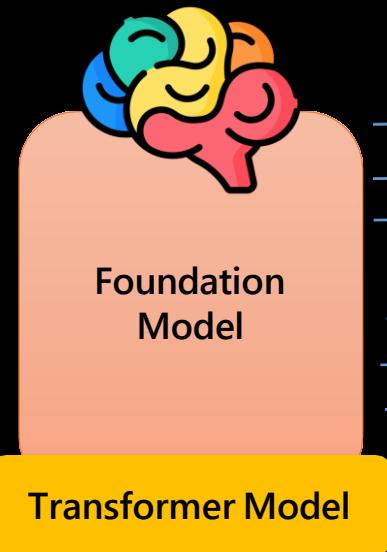
Turing Megatron NLG  
2021

**530B**

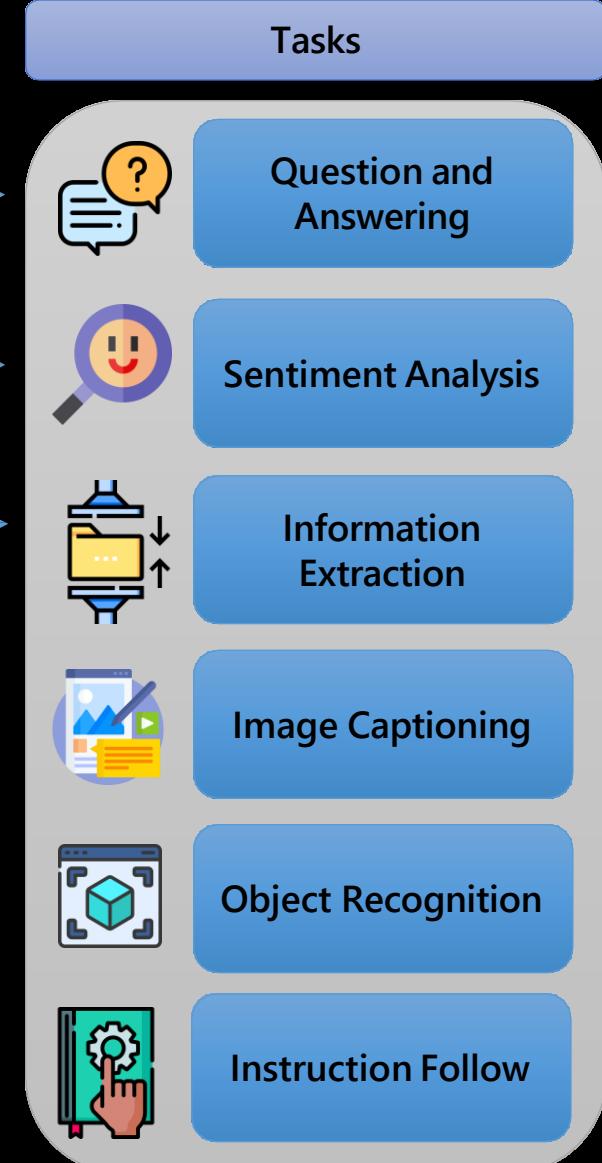
GPT4 – 2023

**1.4T** (estimated)

# Foundation Models



Adaptation







### Transformer Architecture

A new powerful model architecture is released that is the basis for LLMs

### GPT3

By scaling transformer models, special properties emerge such as few-shot learning. LLMs can be good at tasks without training.

### ChatGPT

GPT-3.5 is trained with human feedback through reinforcement learning to achieve remarkable conversational abilities.

### GPT4

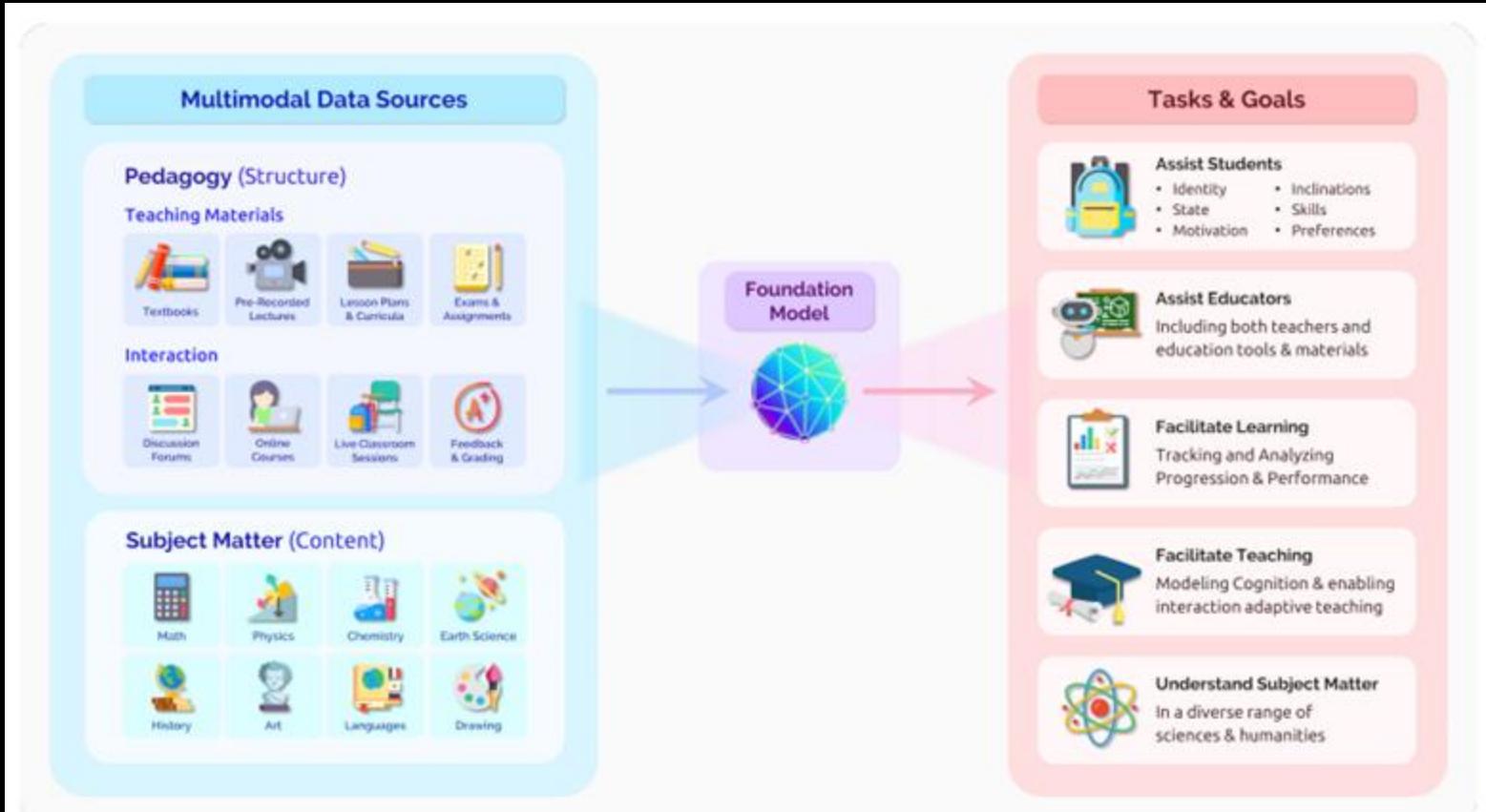
Shows strong performance on human tests and sparks of AGI.

2017

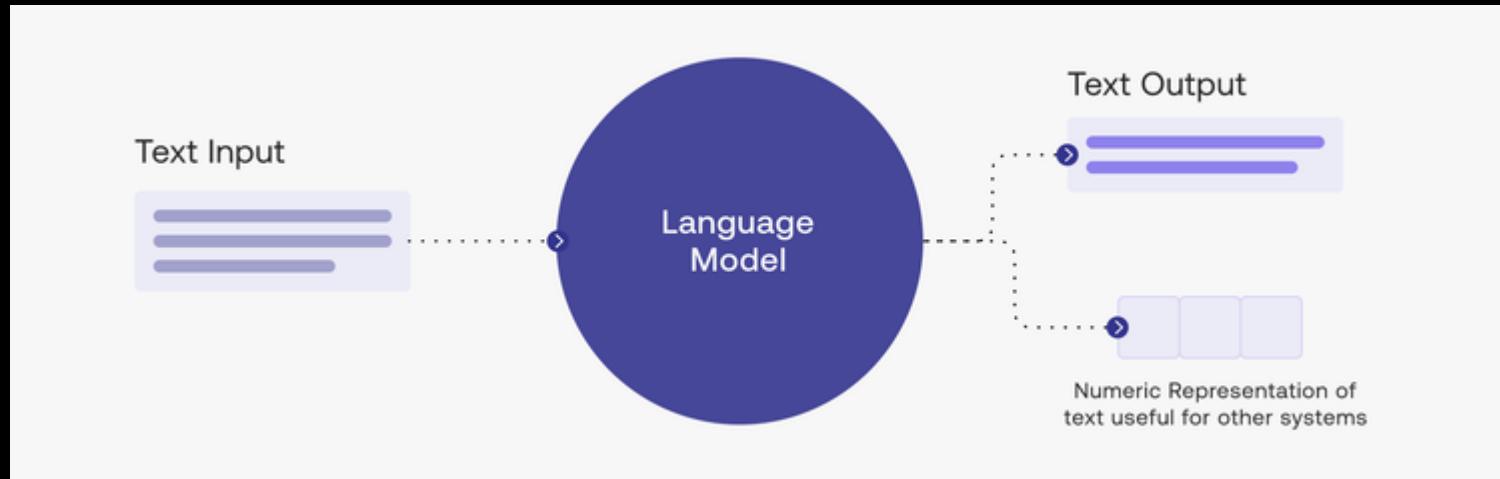
2020

2022

2023

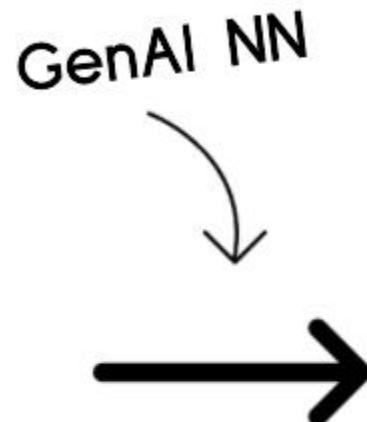


# How to use LLM



**"A painting of a  
flying dog"**

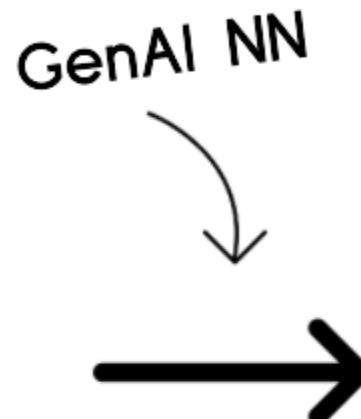
Prompt



Image

**"Write a letter to  
my teacher about  
me missing class  
today"**

↑  
Prompt



**Dear Teacher,**

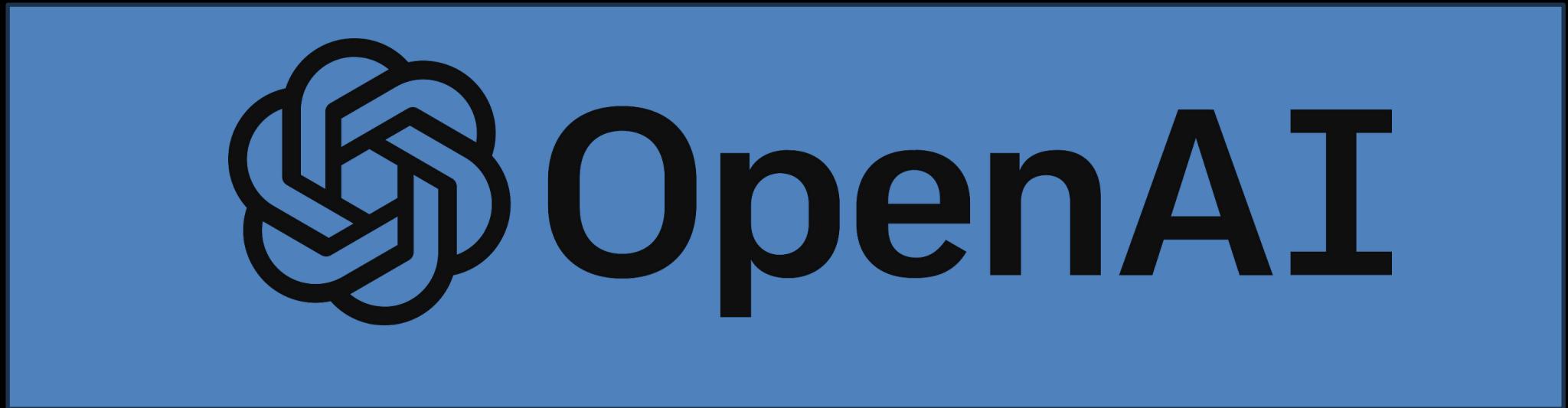
**I am writing to inform you  
that I was unable to attend  
your class today due to a  
personal emergenc..**

↑  
Text

# Introduction to OpenAI

## What is OpenAI?

OpenAI is a private research laboratory that aims to develop and direct artificial intelligence (AI) in ways that benefit humanity as a whole. The company was founded by Elon Musk, Sam Altman and others in 2015 and is headquartered in San Francisco.



# OpenAI API Models

(as of July 2023)

- **GPT-4:** OpenAI's latest large multimodal model. Significant improvement in "complex reasoning tasks" compared to GPT-3.5 models. Optimized for chat. (**Input:** text/image. **Output:** text.)
- **GPT-3.5-turbo:** Model running the free version of ChatGPT. Optimized for chat. (**Input:** text. **Output:** text.)
- **Whisper:** General purpose speech recognition model. Can do transcription, translation, and detect language. Model is open source. (**Input:** audio. **Output:** text.)
- **Embeddings:** Generates (useful) vector representation of text which can be used for downstream tasks e.g. text similarity. (**Input:** text. **Output:** numerical vector.)
- **GPT-3 models:** Original ChatGPT models. Less powerful than GPT-3.5 models, however, GPT-3 models are cheaper. (**Input:** text. **Output:** text.)
- **Moderation:** A fine-tuned model to detect NSFW content. (**Input:** text. **Output:** Category (e.g. hate, harassment, ...))
- **DALL-E:** Creates realistic images based on user prompts. (**Input:** text or image. **Output:** image.)

# Creating video from text

Sora is an AI model that can create realistic and imaginative scenes from text instructions.

<https://cdn.openai.com/sora/videos/tokyo-walk.mp4>

# Introduction to Azure OpenAI?

Azure OpenAI Service is Microsoft's cloud solution for deploying, customizing, and hosting large language models. It brings together the best of OpenAI's cutting edge models and APIs with the security and scalability of the Azure cloud platform. Microsoft's partnership with OpenAI enables Azure OpenAI users to access the latest language model innovations.

Azure OpenAI supports many models that can serve different needs. These models include:

- GPT-4 models are the latest generation of generative pretrained (GPT) models that can generate natural language and code completions based on natural language prompts.
- GPT 3.5 models can generate natural language and code completions based on natural language prompts. In particular, GPT-3.5-turbo models are optimized for chat-based interactions and work well in most generative AI scenarios.
- Embeddings models convert text into numeric vectors, and are useful in language analytics scenarios such as comparing text sources for similarities.
- DALL-E models are used to generate images based on natural language prompts. Currently, DALL-E models are in preview. DALL-E models aren't listed in the Azure OpenAI Studio interface and don't need to be explicitly deployed.



## GPT-3.5 and GPT-4

Text

## Prompt

Write a tagline for an ice cream shop.

## Response

We serve up smiles with every scoop!

## ChatGPT

Conversation

## Prompt

I'm having trouble getting my Xbox to turn on.

## Response

There are a few things you can try to troubleshoot this issue ... ...

## Prompt

Thanks! That worked. What games do you recommend for my 14-year-old?

## Response

Here are a few games that you might consider: ...

## Codex

Code

## Prompt

```
Table customers, columns =  
[CustomerId, FirstName,  
LastName, Company, Address,  
City, State, Country,  
PostalCode]
```

Create a SQL query for all customers in Texas named Jane  
query =

## Response

```
SELECT *  
FROM customers  
WHERE State = 'TX' AND  
FirstName = 'Jane'
```

## DALL-E 2

Images

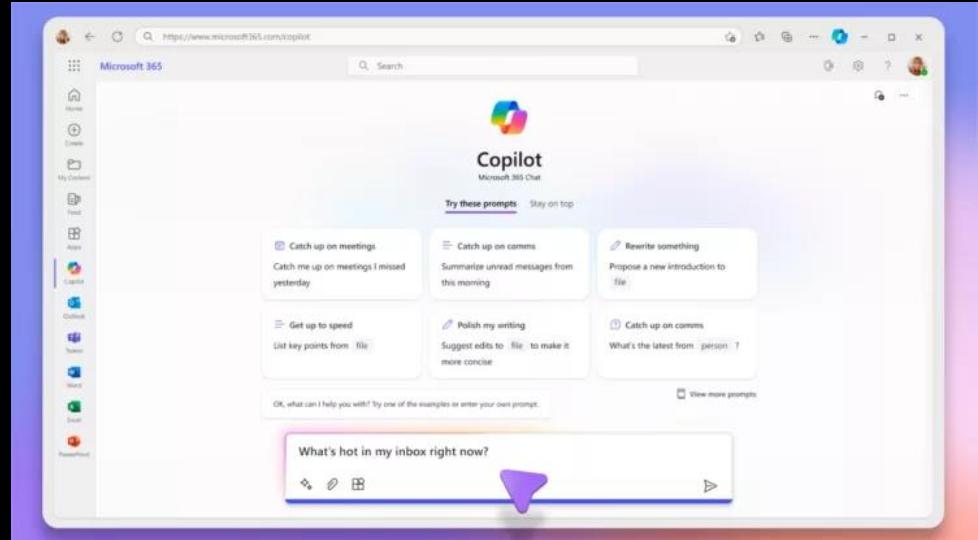
## Prompt

A ball of fire with vibrant colors to show the speed of innovation at our media and entertainment company

## Response

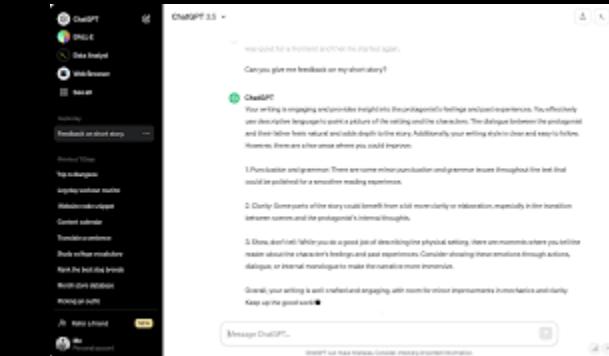


## Microsoft Product based on Azure OpenAI



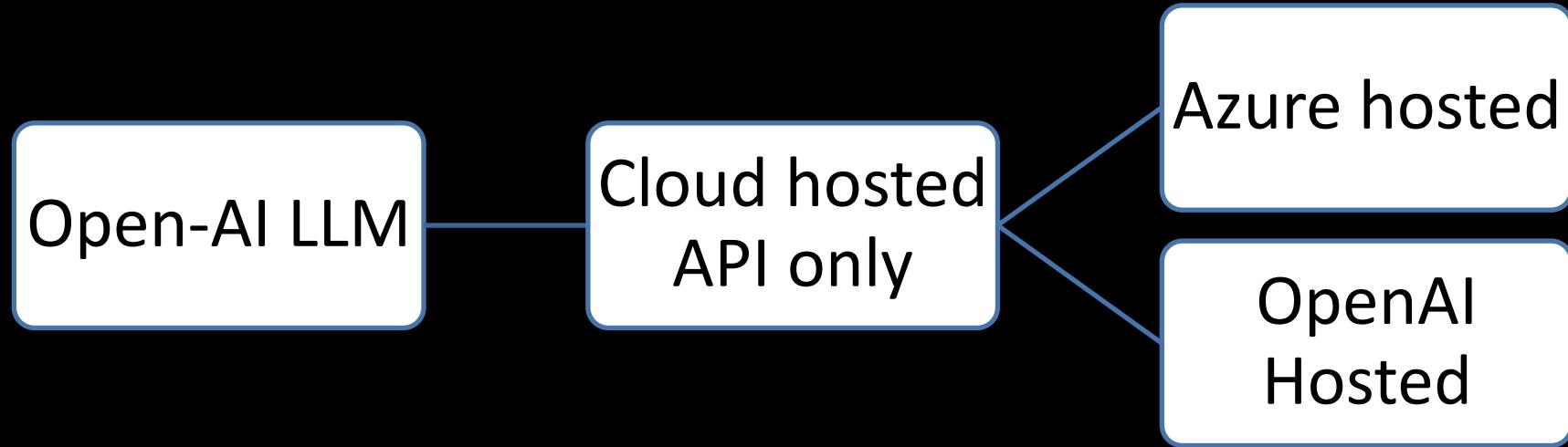
Copilots

## Open Product based on OpenAI Model



GPTs and ChatGPT

## Access to OpenAI



Model	Azure OpenAI Service	OpenAI API	Description	Factors	Azure OpenAI	OpenAI
	YES	NO	A large multimodal model that can understand and generate natural language or code, and perform complex tasks with greater accuracy than any previous model.		A service provided by Microsoft Azure that co-develops the API's with OpenAI.	A research laboratory that develops and publishes new AI Models.
GPT-4	YES	YES	A set of models that improve on GPT-3 and can understand and generate natural language or code.	GPT-4 support	Supports the latest GPT-4 models, which are currently in preview.	Supports the GPT-4 models but you require ChatGPT plus subscription to access the models.
GPT-3.5	YES	YES	A model that can generate and execute code for a wide range of programming languages and tasks.	Security	Provides enterprise-grade security features such as encryption, authentication and role-based access control.	Provides basic security features such as API key and content filtering.
Codex	YES	YES	A model that can generate and edit images given a natural language prompt.	VNET and private endpoint support	Supports virtual network and private endpoint integration as part of Azure Cognitive Services.	Does not support virtual network and private endpoint integration
DALL-E	YES	YES	A model that can convert audio into text.	Access	Access is limited and requires an application process that evaluates the use cases and responsible AI principles	Access is open and requires a subscription plan that varies by model and usage
Whisper	YES	NO	A set of models that can convert text into a numerical form.	Learning Resources	Provides documentation, tutorials, examples, and best practices for using the service	Provides documentation, playgrounds, examples, and blog posts for using the service
Embeddings	YES	NO	A fine-tuned model that can detect whether text may be sensitive or unsafe.	Customer Support	Provides support options through Azure Cognitive Services support guide	Provides support options through OpenAI platform support guide
Moderation	YES	NO	A set of models that can understand and generate natural language.	Fine-tuned Model creation	Allows fine-tuning of some models to customise them for specific tasks	Allows fine-tuning of some models to customise them for specific task
GPT-3	YES	YES		Pricing for ChatGPT	ChatGPT is not available in Azure OpenAI service	ChatGPT is available in OpenAI service with a subscription plan of \$20 per month for ChatGPT Plus program
				Prompt and completion for ChatGPT	ChatGPT is not available in Azure OpenAI service	ChatGPT is available in OpenAI service with a web-based interface that allows users to interact with the model by providing prompts and receiving completions

# Why Azure OpenAI

- Security – AAD RBAC
- Security – secure access via Microsoft private network
  
- Latency - DC location
  
- Privacy/compliance – opt out from data being stored
- Privacy/compliance – Azure compliance certifications
  
- Monitoring – integrates with Azure Monitor
- Harmful content filtering

# Why Azure OpenAI

## OpenAI vs Azure OpenAI

*What's the difference?*

	Azure OpenAI	OpenAI
<i>Data Privacy</i>		
Doesn't use your data for model training	✓	✓
Can opt-out of API data retention	✓	✓
<i>Data Residency</i>		
Can choose between different regions	✓	✗
<i>Infrastructure Security</i>		
Virtual Networking	✓	✗
Private Link	✓	✗
Access Control	✓	✗

# Why Azure OpenAI

- Security – AAD RBAC
- Security – secure access via Microsoft private network
- Latency - DC location
- Privacy/compliance – opt out from data being stored
- Privacy/compliance – Azure compliance certifications
- Monitoring – integrates with Azure Monitor
- Harmful content filtering

# Why Azure OpenAI

- Security – AAD RBAC
- Security – secure access via Microsoft private network
  
- Latency - DC location
  
- Privacy/compliance – opt out from data being stored
- Privacy/compliance – Azure compliance certifications
  
- Monitoring – integrates with Azure Monitor
- Harmful content filtering

# Subfields of Artificial Intelligence

- Machine Learning (ML)

This subfield of AI aims to develop algorithms that enable machines to learn from data without explicit programming.

- Natural Language Processing (NLP)

Natural language processing aims to teach machines to understand and generate human language. Speech recognition, language translation, and text analysis are some methods used in NLP.

- Neural Networks

A neural network is an algorithm inspired by the arrangement and function of the human brain.

- Expert Systems

- Robotics

- Computer Vision

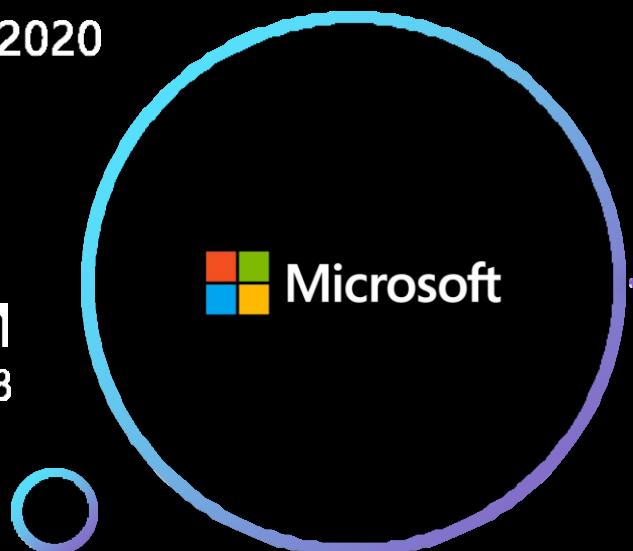
# Large Language Foundation models

Microsoft AI at Scale

## Project Turing

17B  
Turing-NLG | 2020

340M  
BERT-Large | 2018



Our mission is to expand the boundaries of natural language understanding.

machine reading comprehension, question answering, transfer learning, reinforcement learning, computer vision, and even building interpretable models

Turing-NLG

Megatron NLG Turing

T-ULRv5

DeepSpeed

ZeRO

**175B**

GPT-3 | 2020

**17B**

Turing-NLG | 2020

**340M**

BERT-Large | 2018

OpenAI



Microsoft



**530B**

Megatron-Turing  
NLG | 2021



**175B**

GPT-3 | 2020

**17B**

Turing-NLG | 2020

**340M**

BERT Large | 2018

# Our partnership with OpenAI



*Ensure that artificial  
general intelligence (AGI)  
benefits humanity.*



*Empower every person and  
organization on the planet  
to achieve more*

---

**GPT-3x & 4x**

Generate and Understand Text

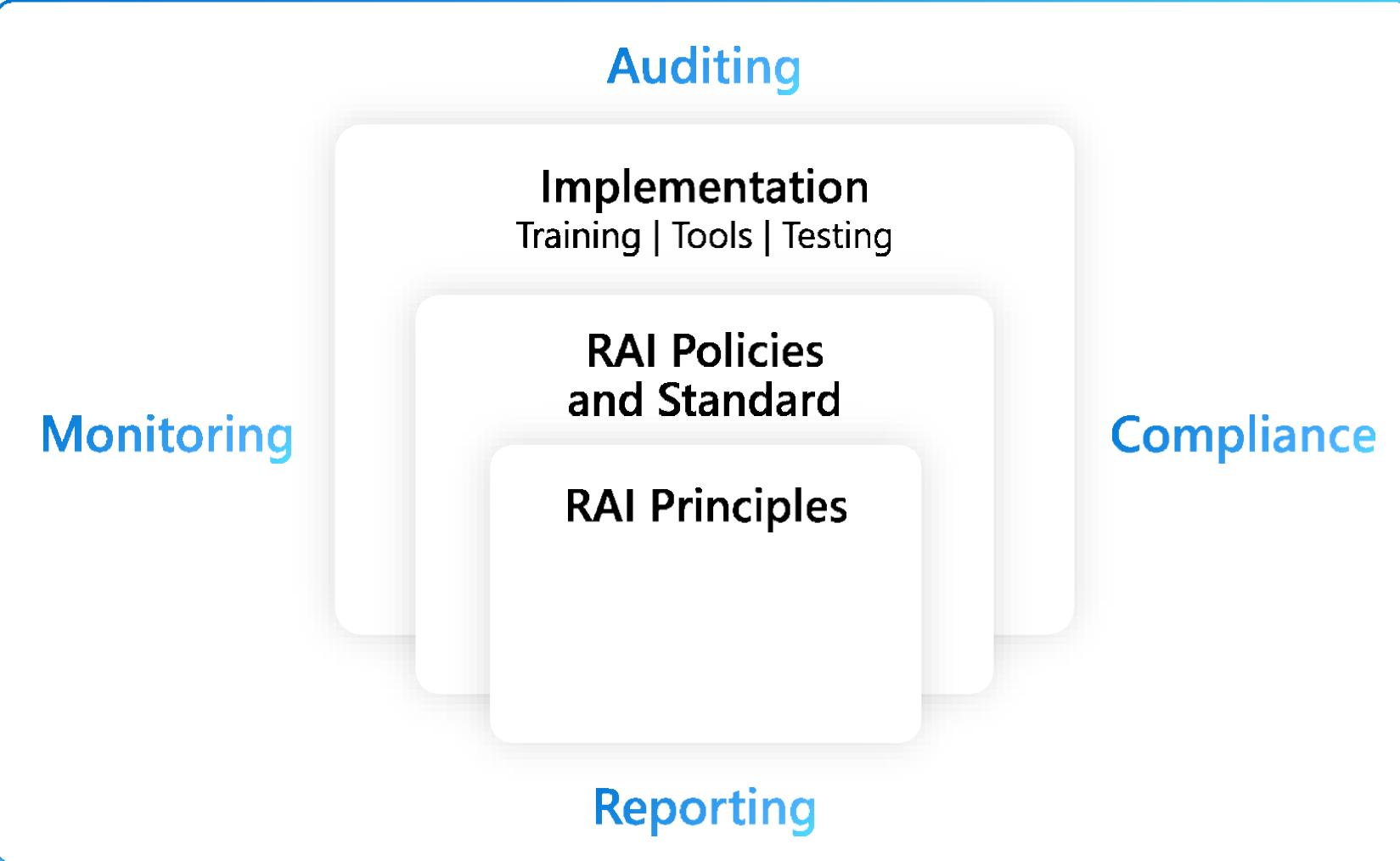
**DALL·E**

Generate images from text prompts

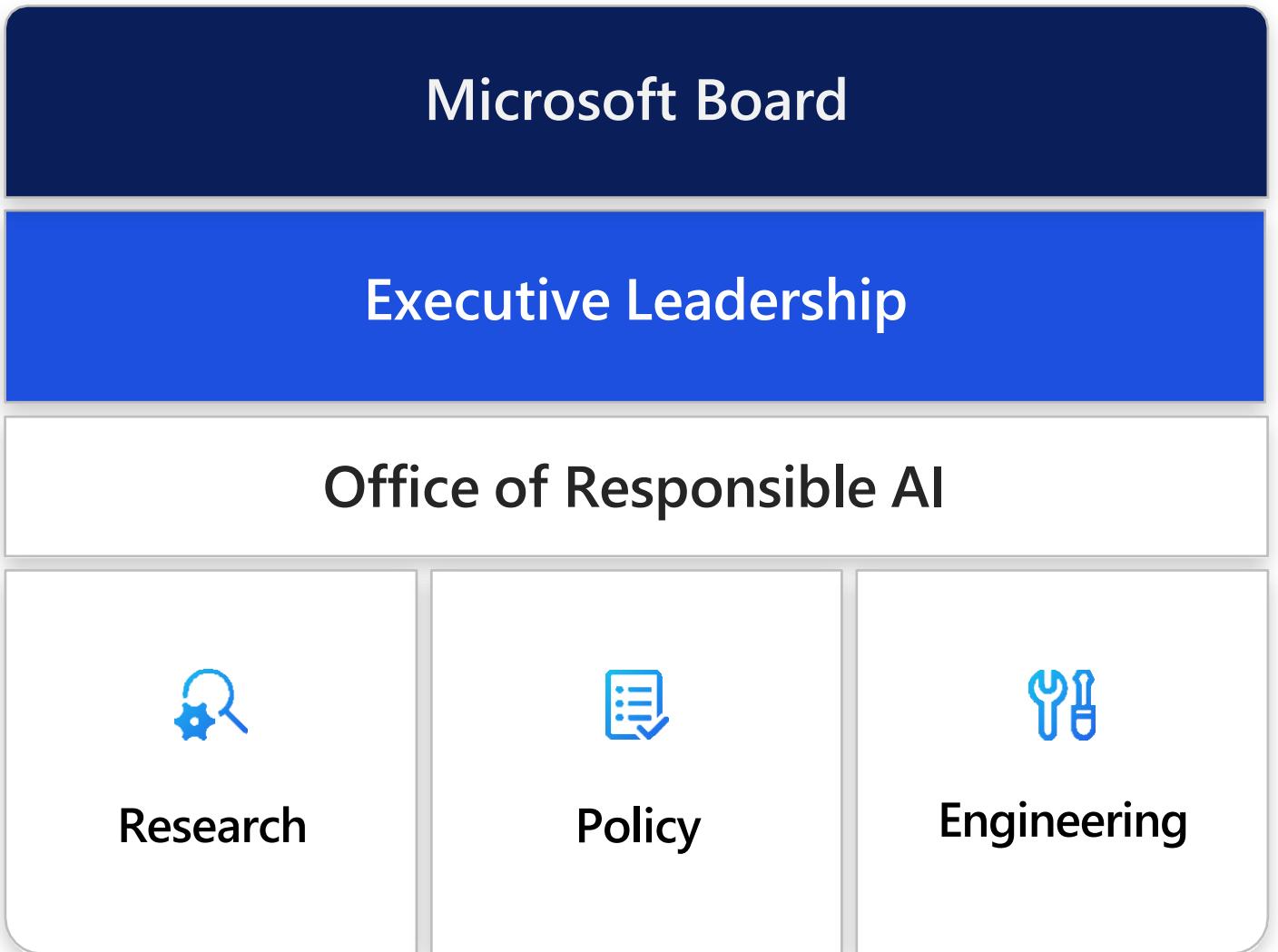
**ChatGPT**

Generate and Understand Code

# Responsible AI Governance Framework



# Our ecosystem



# The Anatomy of the Responsible AI Standard

Principles

> Which **enduring values** guide our responsible AI work?

Goals

> What are the **outcomes** that we need to secure?

Requirements

> What are the **steps we must take** to secure the Goals?

Tools and Practices

> Which **aids** can help us meet the Requirements?

# The Standard's Goals at a Glance

## Accountability

- A1: Impact Assessment
- A2: Oversight of significant adverse impacts
- A3: Fit for purpose
- A4: Data governance and management
- A5: Human oversight and control

## Transparency

- T1: System intelligibility for decision making
- T2: Communication to stakeholders
- T3: Disclosure of AI interaction

## Fairness

- F1: Quality of service
- F2: Allocation of resources and opportunities
- F3: Minimization of stereotyping, demeaning, and erasing outputs

## Reliability & Safety

- RS1: Reliability and safety guidance
- RS2: Failures and remediations
- RS3: Ongoing monitoring, feedback, and evaluation

## Privacy & Security

- PS1: Privacy Standard compliance
- PS2: Security Policy compliance

## Inclusiveness

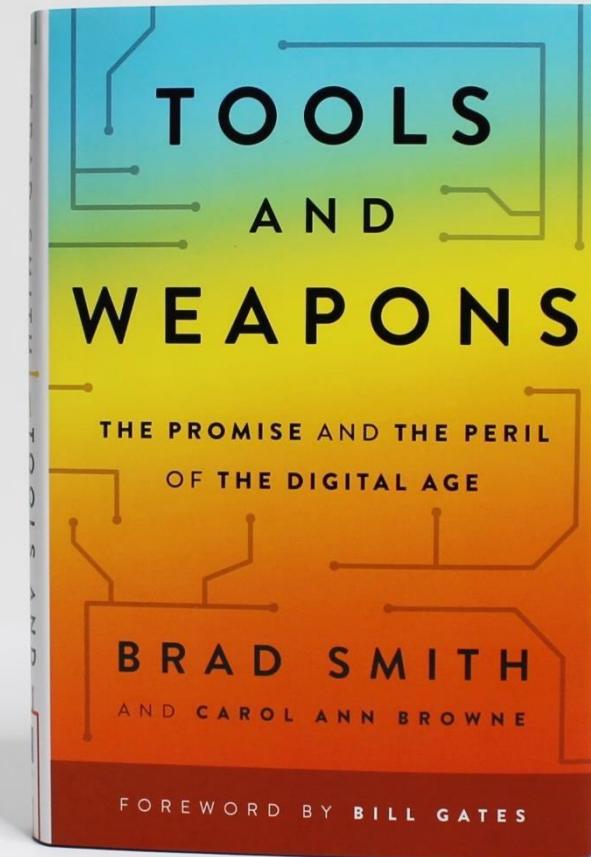
- I1: Accessibility Standards compliance

# Why responsible AI?

“When your technology changes the world, you bear a responsibility to help address the world that you have helped create.”

**Brad Smith**

President and Chief Legal Officer, Microsoft

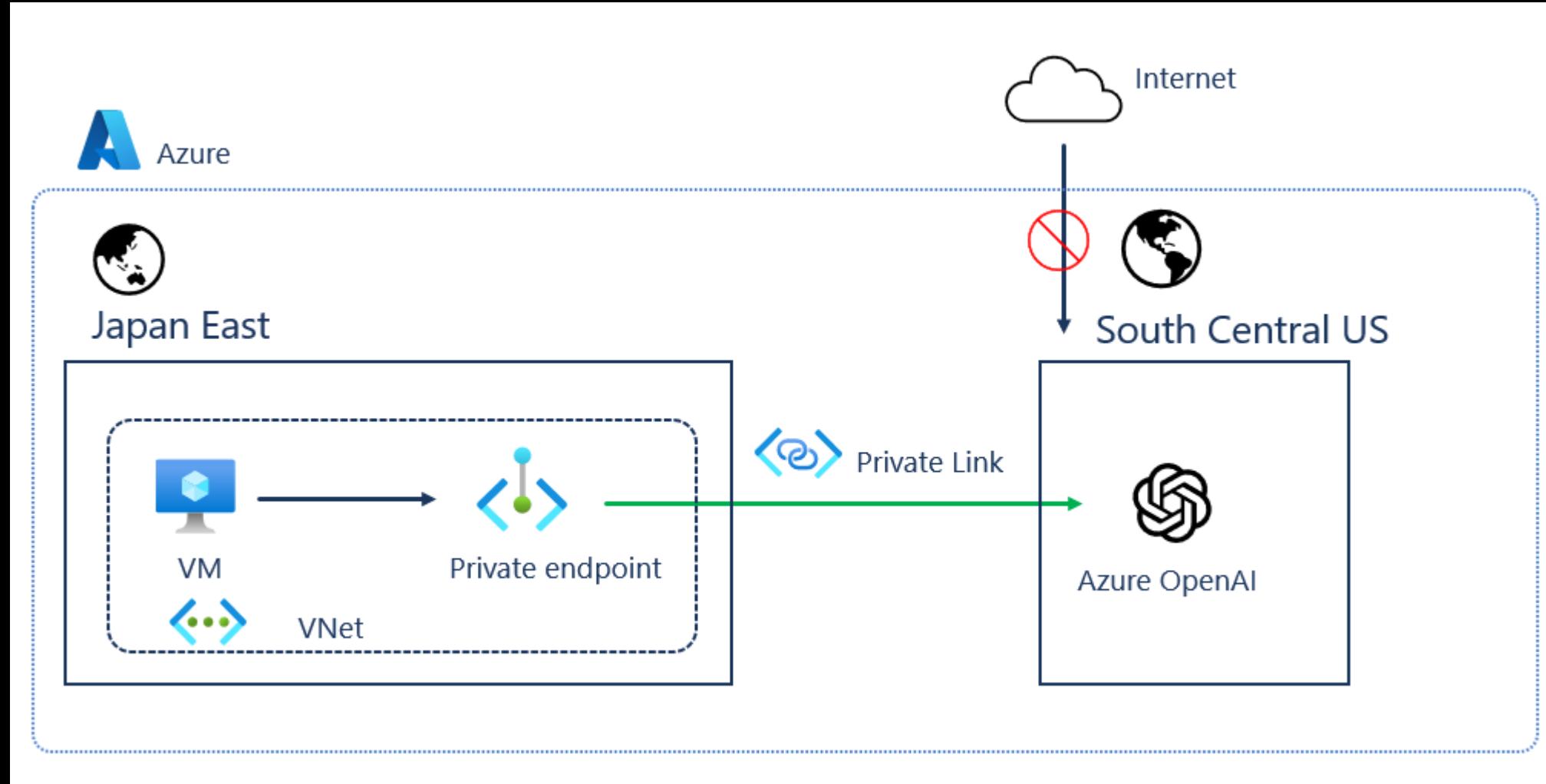


[Responsible AI principles from Microsoft](#)

# Why Azure OpenAI

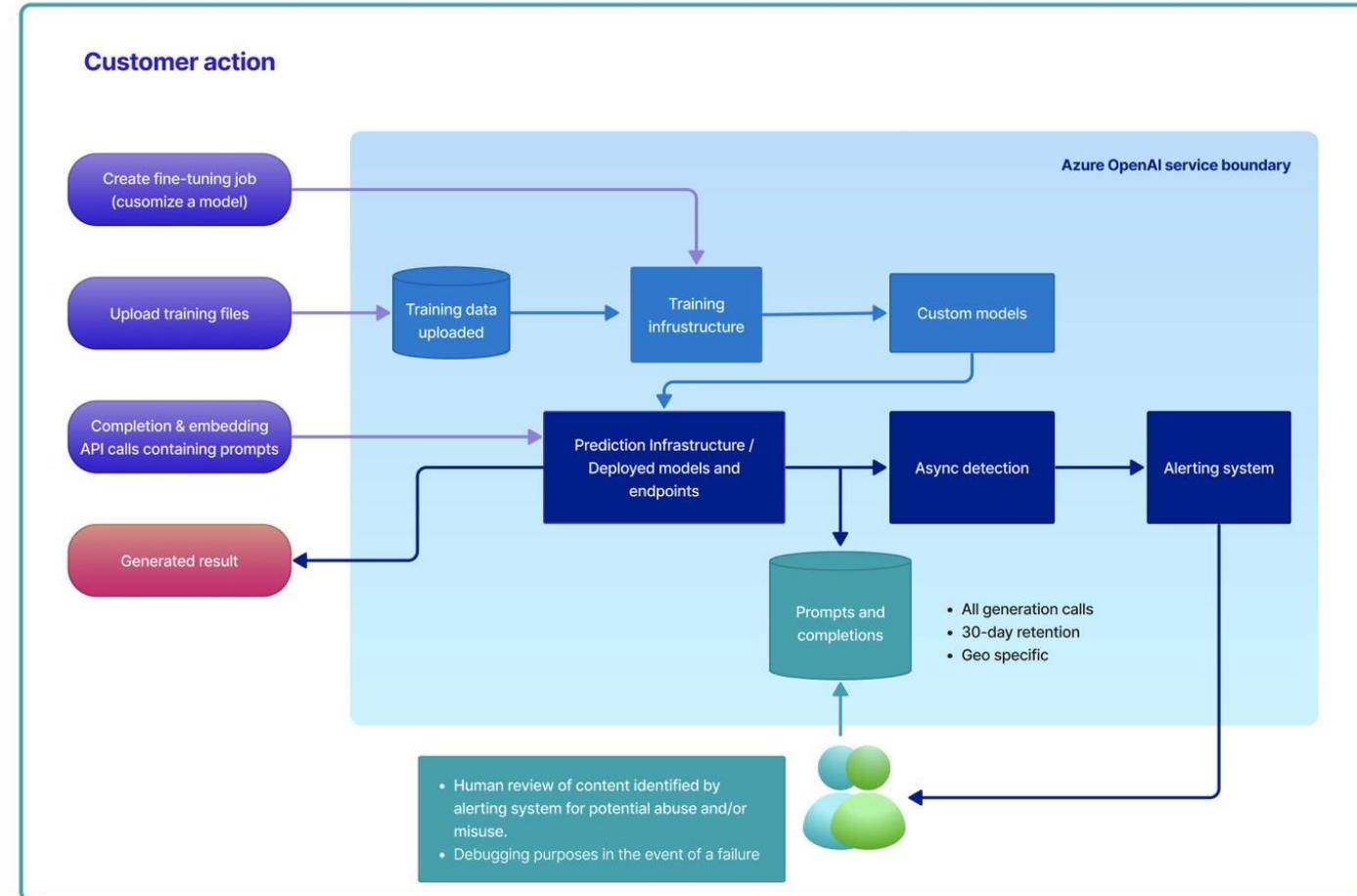
- Security – AAD RBAC
- Security – secure access via Microsoft private network
  
- Latency - DC location
  
- Privacy/compliance – opt out from data being stored
- Privacy/compliance – Azure compliance certifications
  
- Monitoring – integrates with Azure Monitor
- Harmful content filtering

# Why Azure OpenAI



# Data privacy

# How Azure OpenAI processes data



# Azure OpenAI Data Processing

## CAN A CUSTOMER OPT OUT OF HAVING PROMPTS-AND-COMPLETIONS LOGGED FOR CONTENT FILTERING & ABUSE MONITORING?

- Customers can apply to have abuse monitoring and/or content filtering configured off by completing this [Limited Access form](#).
- Once approved, a customer can determine if they want Microsoft to configure off content filtering, abuse monitoring, or both.
  - If abuse monitoring is configured off, no prompt and completion data is stored, and no human review is possible. This configuration happens at the subscription level, and the [product documentation](#) explains how a customer can [verify that logging is off](#).
- Because the [content filtering](#) system does not store prompt and completion data, content filtering can remain active and no prompt and completion data will be stored.

# Uses and Capabilities

# Unlock new use cases



## Content generation

Call Center Analytics: Automatically generate responses to customer inquiries



## Code generation

Aircraft company using Azure OpenAI Service to convert natural language to SQL for aircraft telemetry data

Consulting service using Azure OpenAI Service to convert natural language to query proprietary data models



## Semantic search

Financial services firm using Azure OpenAI Service to improve search capabilities and the conversational quality of a customer's Bot experience.

Insurance companies extract information from volumes of unstructured data to automate claim handling processes



## Summarization

International insurance company using Azure OpenAI Service to provide summaries of call center customer support conversation logs

Global bank using Azure OpenAI Service to summarize financial reporting and analyst articles

Government agency using Azure OpenAI Service to extract and summarize key information from their extensive library of rural development reports

Financial services using Azure OpenAI Service to summarize financial reporting for peer risk analysis and customer conversation summarization

# Examples of multiple model use cases



## End-to-end call center analytics

- Classification
- Sentiment
- Entity extraction summarization
- Email generation

## Customer 360

- Hyper-personalisation using timely Summarization of customer queries & trends
- Search
- Content generation

## Business process automation

- Search through structured & unstructured documentation
- Generate Code to query data models
- Content Generation



# Azure OpenAI Service

## Use cases



### Codex

- Natural Language to Code
- Natural Language to SQL
- Code to Natural Language
- Code documentation
- Refactoring

### DALL·E 2

- Creative ideation
- Podcast and music playlist images
- Content syndication
- Marketing campaign personalization
- Hyper-personalization



# Azure OpenAI Service

## Use cases

### Language

Reason over structured and unstructured data:  
**Classification, Sentiment, Entity Extraction, Search**

- Product feedback sentiment
- Customer and employee feedback classification
- Claims and risk analyses
- Support emails and call transcripts
- Social media trends

### Writing assistance

- Marketing copy/email taglines
- Long format text
- Paragraphs from bullets



# Azure OpenAI Service

## Use cases

### Language

#### Summarization

- Call center call transcripts
- Subject Matter Expert Documents
  - Competitive analysis
  - Peer Analysis
  - Technical reports
- Product and service feedback
- Social media trends

#### Conversational AI

- Smart assists for call centers
- Tech support chat bots
- Virtual assistants



# Azure OpenAI Service

## Use cases

### Language

#### Use Cases that use multiple model capabilities

- Contact Centers
  - Classification—route mails to appropriate team
  - Sentiment—prioritize angry customers
  - Entity extraction and search—analyze liability and risk
  - Mail and call transcript summarization
  - Customer response email generation
- Rapid response marketing campaigns: classification, sentiment, summarization, content generation, image generation

# A Additional considerations

This is only a REST API. It needs to be connected to applications and/or business processes to be useful.

These models use input from instructions and examples in the prompt to identify the task.

The model completes the task by predicting the most probable next text.

In the next section(s), we will cover three main approaches for in-context learning.

There are some additional considerations when choosing a use case. [Learn more](#)

# Q&A

## Azure GPT

| Ask the team anything



Usama Wahab Khan  
Twitter : @usamawahabkhan  
LinkedIn : Usamawahabkhan

