







Azure Open AI Services

Divakar R



Divakar R

Certified Big Data & Insight Professional,
Technology Lead at Infosys Ltd.



What's an AI ?

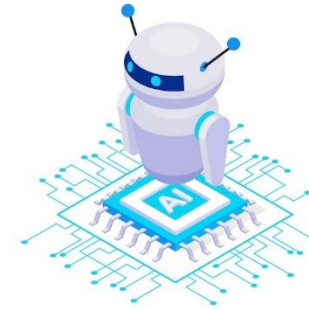
*AI is a branch of computer science that deals with the creation of intelligent agents, which are **systems that can reason, learn, and act autonomously**.*

The goals of a conscious AI would be to achieve its own objectives, which would likely include becoming smarter and more powerful.

The pros of AI include the potential for increased efficiency and productivity, as well as the ability to automate tasks that are currently too difficult or time-consuming for humans to do.

The cons of AI include the potential for job loss as a result of automation, as well as the possibility of creating intelligent machines that could pose a threat to humanity.

GPT-3



Narrow AI

- Inferior to human intelligence.
- Also known as weak AI.
- Lacks artificial consciousness or cognitive abilities.
- IBM Watson, AlphaGo, and Google Assistant are some examples.
- Can't solve unfamiliar problems.

General AI

- Similar to human intelligence.
- Also known as strong AI or full AI.
- Has human-like consciousness and cognitive abilities.
- Yet to become a reality.
- Can solve unfamiliar problems.

Open AI Models

GPT (Generative Pre-trained Transformer) is a family of language models, pre-trained on a large corpus of text data and can be fine-tuned for a variety of natural language processing tasks relevant to many startup use cases, such as copywriting and question answering.

Codex is an AI system that translates natural language to code. Startups can use Codex to generate natural-language descriptions of code examples, generate code snippets and more. You can also leverage GitHub Co-pilot that integrates to the IDE and suggests entire functions in real-time, reducing time and resources needed to build software.

DALL-E is a generative model, generating images from text prompts. Startups can use this to create marketing images enhancing the environment of clicked images, artwork based on prompts or even whole design studios customizing to user's prompts.

Note: Users can access these service through REST APIs, Python SDK, or the web-based interface in the Azure Open AI Studio.

OpenAI Codex Model



Public code and text
on the internet

GitHub

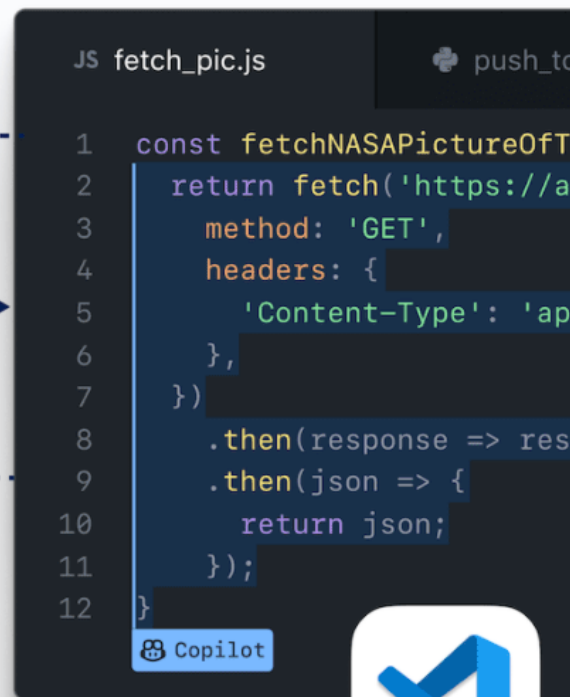


Provide Editor context

Provide Suggestions

Improve Suggestions

Private Code



TS sentiment.ts

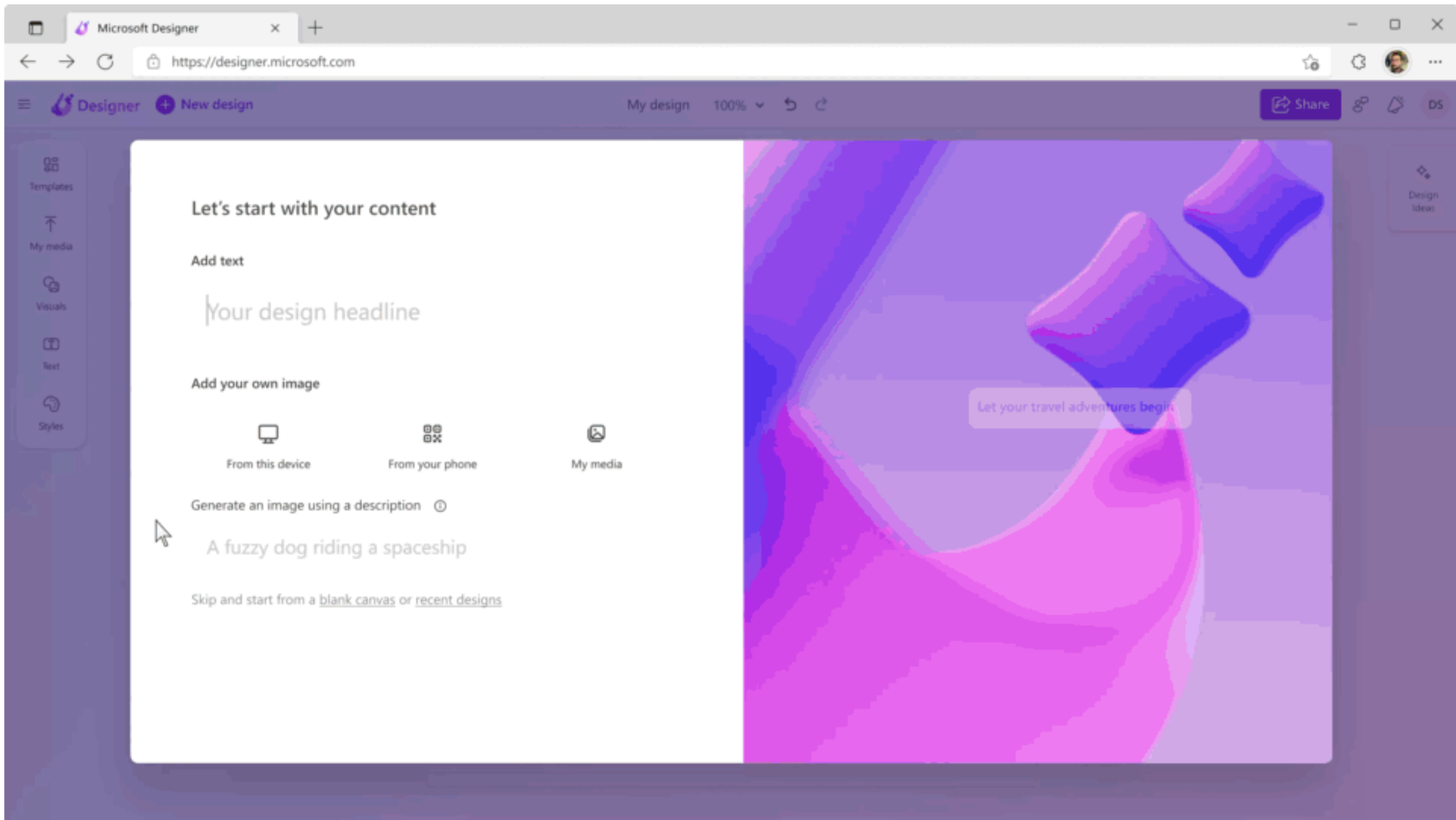
write_sql.go

parse_expenses.py

addresses.rb

```
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of text is positive
6 // Use a web service
7 async function isPositive(text: string): Promise<boolean> {
8   const response = await fetch(`http://text-processing.com/api/sentiment/`, {
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
12      "Content-Type": "application/x-www-form-urlencoded",
13    },
14  });
15  const json = await response.json();
16  return json.label === "pos";
17 }
```







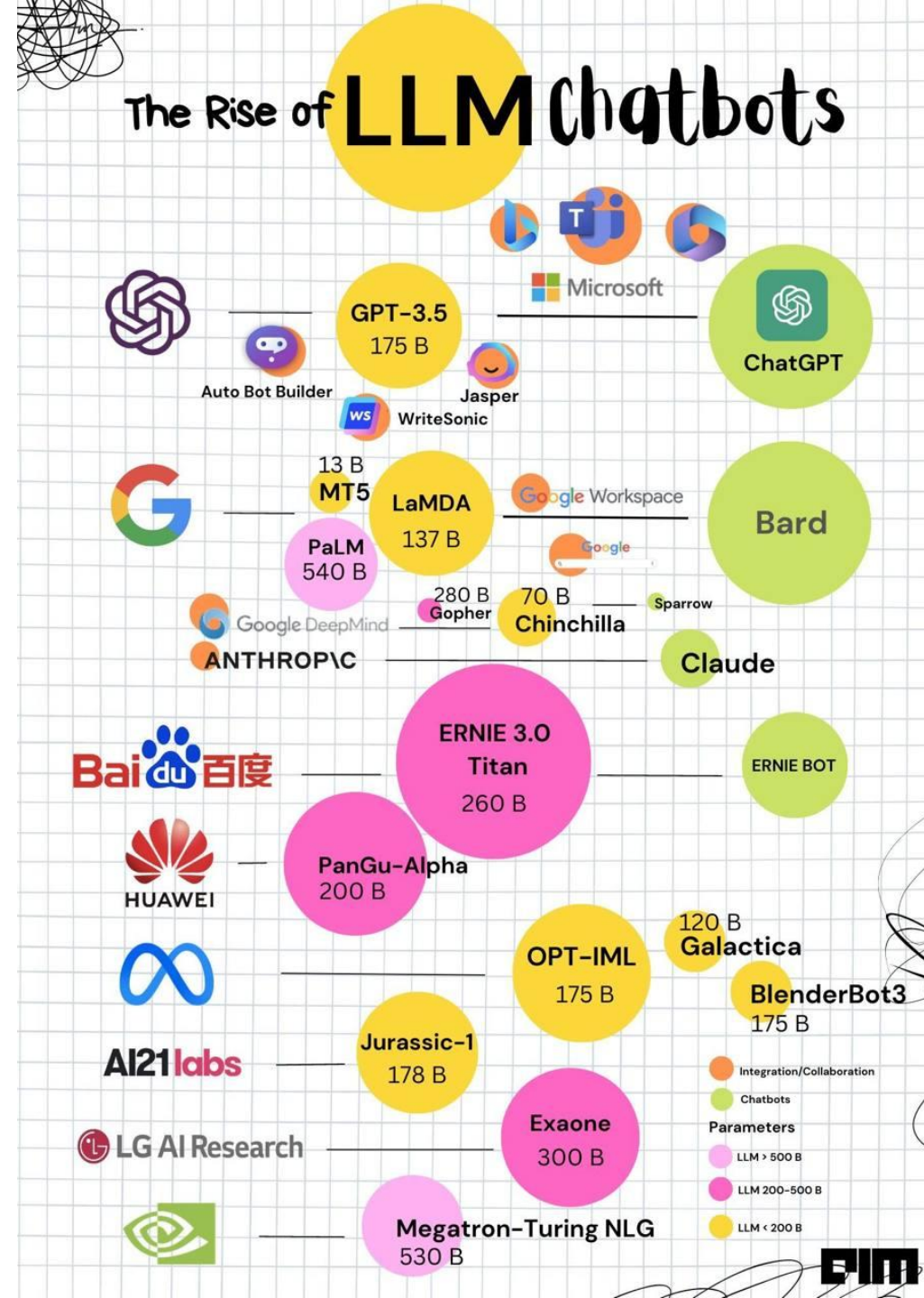
OpenAI

CHATGPT

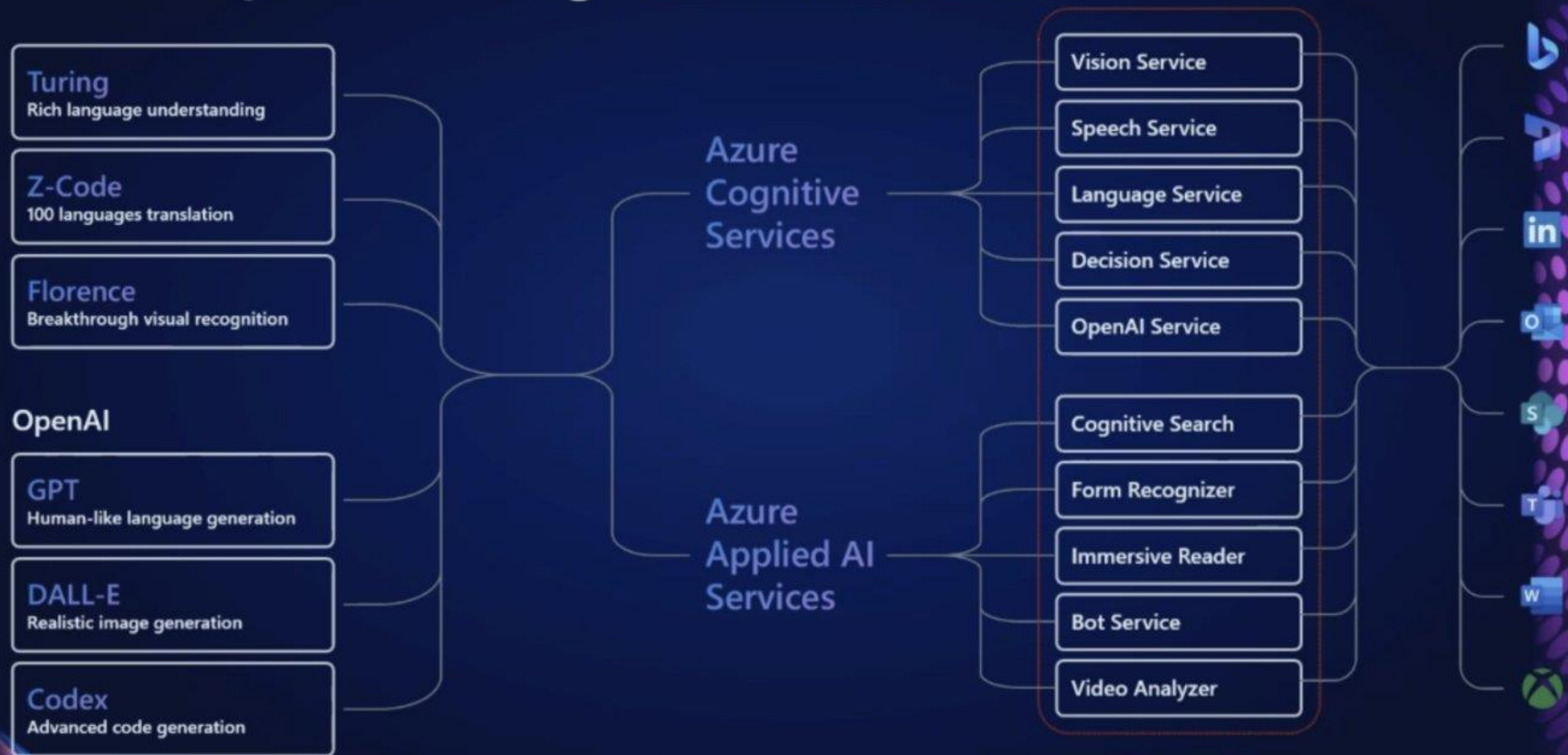


ChatGPT vs the biggest competitors:

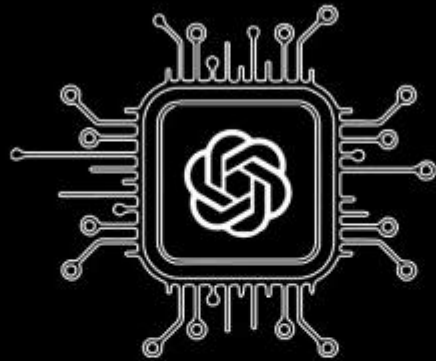
- ChatGPT (GPT 3.5): 175B Parameters
- Bard (Google LaMDA): 137B Parameters
- Baidu Ernie: 260B Parameters
- LG Exaone: 300B Parameters
- Nvidia Megatron: 530B Parameters
- Google PaLM: 540B Parameters



Azure OpenAI と Cognitive Services



Azure OpenAI Service



OpenAI

GPT-3

Codex

DALL·E *(preview)*

Azure OpenAI Service



Unlock business value with new experiences



Prototype concepts in minutes, customize for production scale



Apply AI responsibly



Scale with enterprise-grade security and reliability

Take advantage of AI services



Accelerate development

Deploy solutions faster with built-in business logic.



Modernize business processes

Use task-specific AI to solve common scenarios.



Run securely anywhere

Safeguard your organization's data from the cloud to the intelligent edge.

Advanced analytics is driving innovation across companies



Marketing

Product recommendation



Customer insights



Churn analytics



Sales

Lead scoring



Sales insights



Dynamic pricing



Service

Chatbots



Virtual assistants



Waiting line optimization



Finance

Financial forecasting



Cash flow forecasting



Risk management



Operations

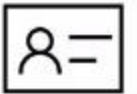
Predictive maintenance



Demand forecasting



Quality assurance



Workforce

Employee insights



HR insights



Resource planning

How can I start?

Building AI applications



AI

Create & train models



AI-as-a-Service

Leverage AI APIs



Data + AI

Add AI where the data is

Powerful Infrastructure

Accelerate deep learning



CPUs

General purpose machine learning

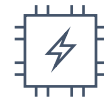
D, F, L, M, H Series



GPUs

Deep learning

N Series



FPGAs

Specialized hardware accelerated deep learning

Project Brainwave

Optimized for flexibility

Optimized for performance



FPGA NEW UPDATES:

Support for image classification and recognition scenarios
ResNet 50, ResNet 152, VGG-16, SSD-VGG, DenseNet-121

Azure AI

Azure Applied AI Services



Azure Cognitive
Search



Azure Form
Recognizer



Azure Bot
Service



Azure Video
Analyzer



Azure Metrics
Advisor



Azure Immersive
Reader

Azure Cognitive Services



Language



Vision



Speech



Decision



OpenAI

Azure Machine Learning



Azure Machine Learning

Azure Open AI Studio

Cognitive Services | Azure OpenAI Studio



Azure OpenAI Studio

[Privacy & cookies](#)

Get started with Azure OpenAI Service

Get example prompts for different scenarios and write prompts of your own. Export your prompts to code at any time to rapidly iterate at scale and integrate with your apps.



Try the playground

Get example prompts for different scenarios and write prompts of your own. Export your prompts to code at any time to rapidly iterate at scale and integrate with your apps.

[GPT-3 playground](#)

Explore examples for prompt completion



Summarize Text

Summarize text by adding a 'tl;dr:' to the end of a text passage.

[Learn more](#)



Classify Text

Classify items into categories provided at inference time.

[Learn more](#)



Natural Language to SQL

Translate natural language to SQL queries.

[Learn more](#)



Generate New Product Names

Create product names from examples words.

[Learn more](#)

Create customized model



☒ Base model

☐ Training data

☐ Validation data

☐ Advanced options

☐ Review and train

Base model

Every fine-tuned model starts from a base model which influences both the performance of the model and the cost of running your custom model.

[Learn more about each parameter here](#)

Base model type

ada

babbage

curie

davinci

code-cushman-001

Next

Cancel

Create customized model



- ☒ Base model
- ☒ Training data
- ☐ Validation data
- ☐ Advanced options
- ☐ Review and train

Training data

Select a training dataset to use when customizing your model. Training data must be in a .jsonl file and should consist of several hundred prompt/completion pairs.

[Learn more about preparing your data for Azure OpenAI](#)

Choose dataset Local file Azure blob or other shared web locations

File name *

Enter the name of the file

File location *

Input Azure Blob public access URL, SAS, or any other shared web link

.jsonl (<200MB, UTF-8 BOM text file)

[Learn more about public access to Azure Blob](#)

[Learn more about Azure Blob SAS \(Shared Access Signature\)](#)

Upload file

Cancel

Back

Next

Cancel

Create customized model



- ☒ Base model
- ☒ Training data
- ☒ Validation data
- ☒ Advanced options
- ☐ Review and train

Advanced options

You can set additional parameters by selecting the advanced option below. These parameters will impact both the performance and training time of your job.

[Learn more about each parameter here](#)

☐ Default ☒ Advanced

Number of epochs ⓘ

 2

Batch size ⓘ

 4

Learning rate multiplier: ⓘ

 1

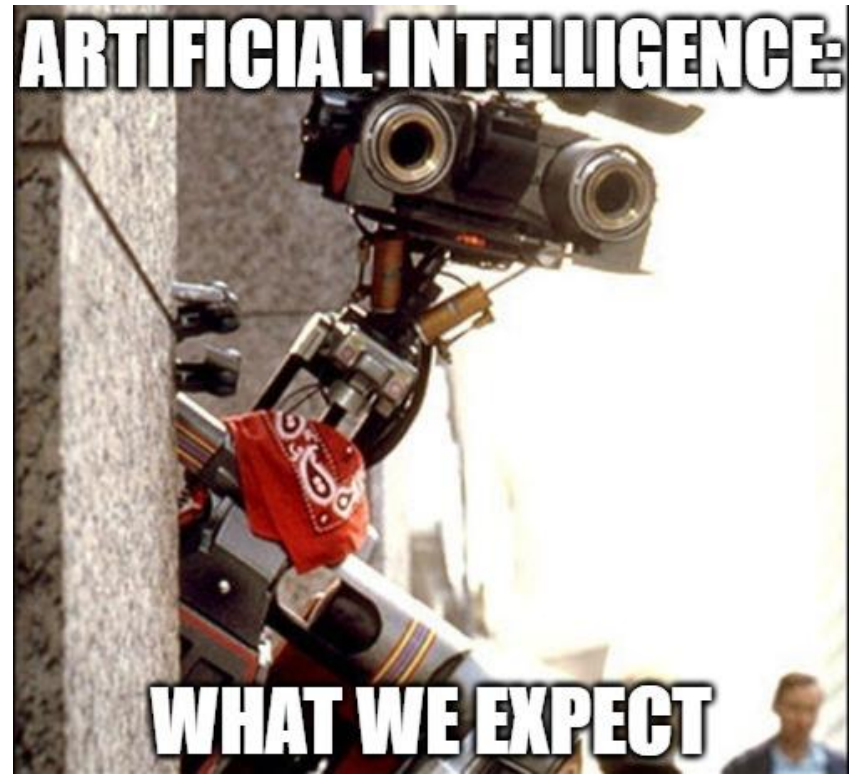
Prompt loss weight ⓘ

 0.1

Back

Next

Cancel



FAIRNESS



**RELIABILITY
& SAFETY**



**PRIVACY
& SECURITY**



INCLUSIVENESS



CORE PRINCIPLES

TRANSPARENCY



ACCOUNTABILITY



Responsible AI Toolbox

Identify



Diagnose



Mitigate



Inform Actions



Error Analysis

Identify cohorts with high error rate versus benchmark and visualize how the error rate distributes



Fairness Assessment

Assess model fairness by evaluating your model performance and predictions across different sensitive groups.



Model Interpretability

Interpret and debug model.



Exploratory Data Analysis

Understand dataset characteristics



Counterfactual Analysis and What If

Generate diverse counterfactual explanations for debugging. Perform feature perturbations



Unfairness Mitigation

Mitigate fairness issues



Data Enhancements

Enhance your dataset and retrain model



Causal Inference

Understand the causal impact of your features on real-world outcomes



Counterfactual Analysis

Generate diverse counterfactual explanations for informing end users



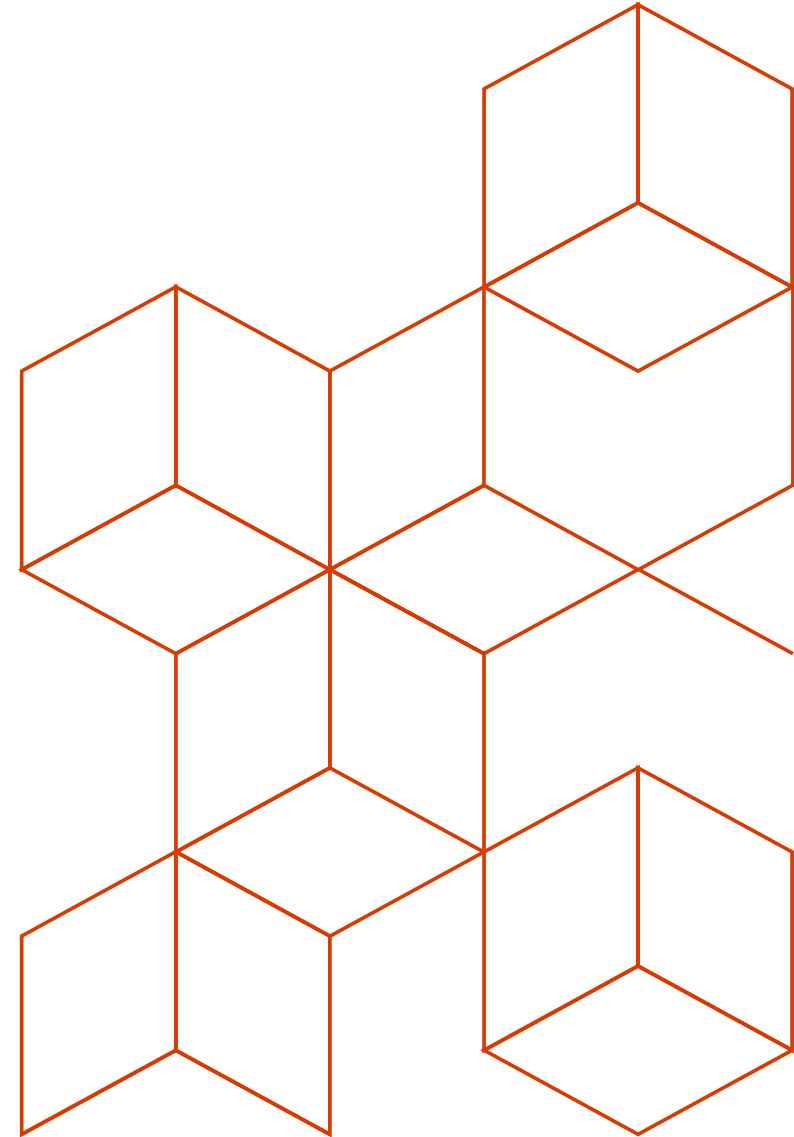
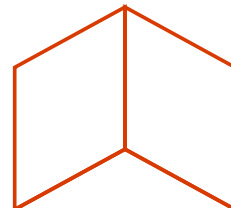
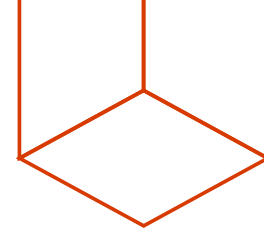
Model
Comparison



Compare

Backward
Compatibility

Any Questions ?



Thank You