



global AI bootcamp

March 4th 2023, Torino - Italy

Azure OpenAI's API



Michele Aponte

CEO&CTO @ Ellycode / CEO @ Blexin / CTO @ Improove Skills
Founder @ Blazor Developer Italiani / MVP Developer Technologies & AI



apomic80



apomic80



apomic80



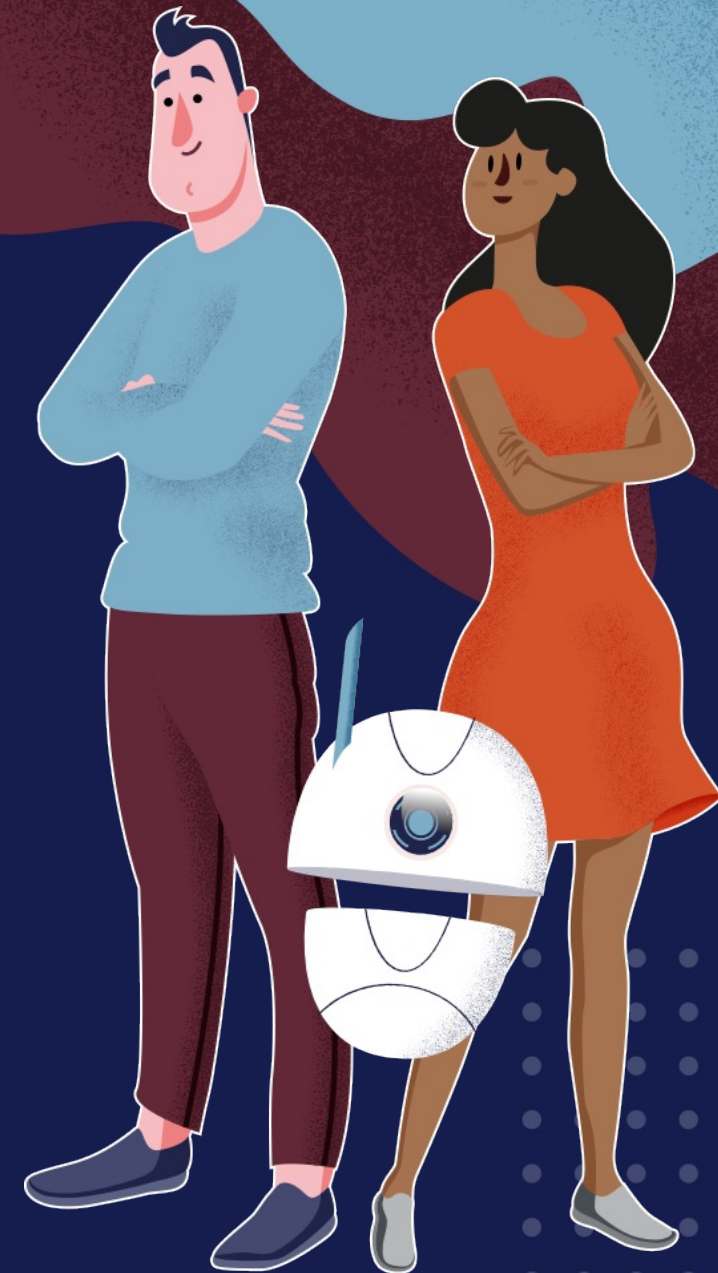
apomic80



apomic80

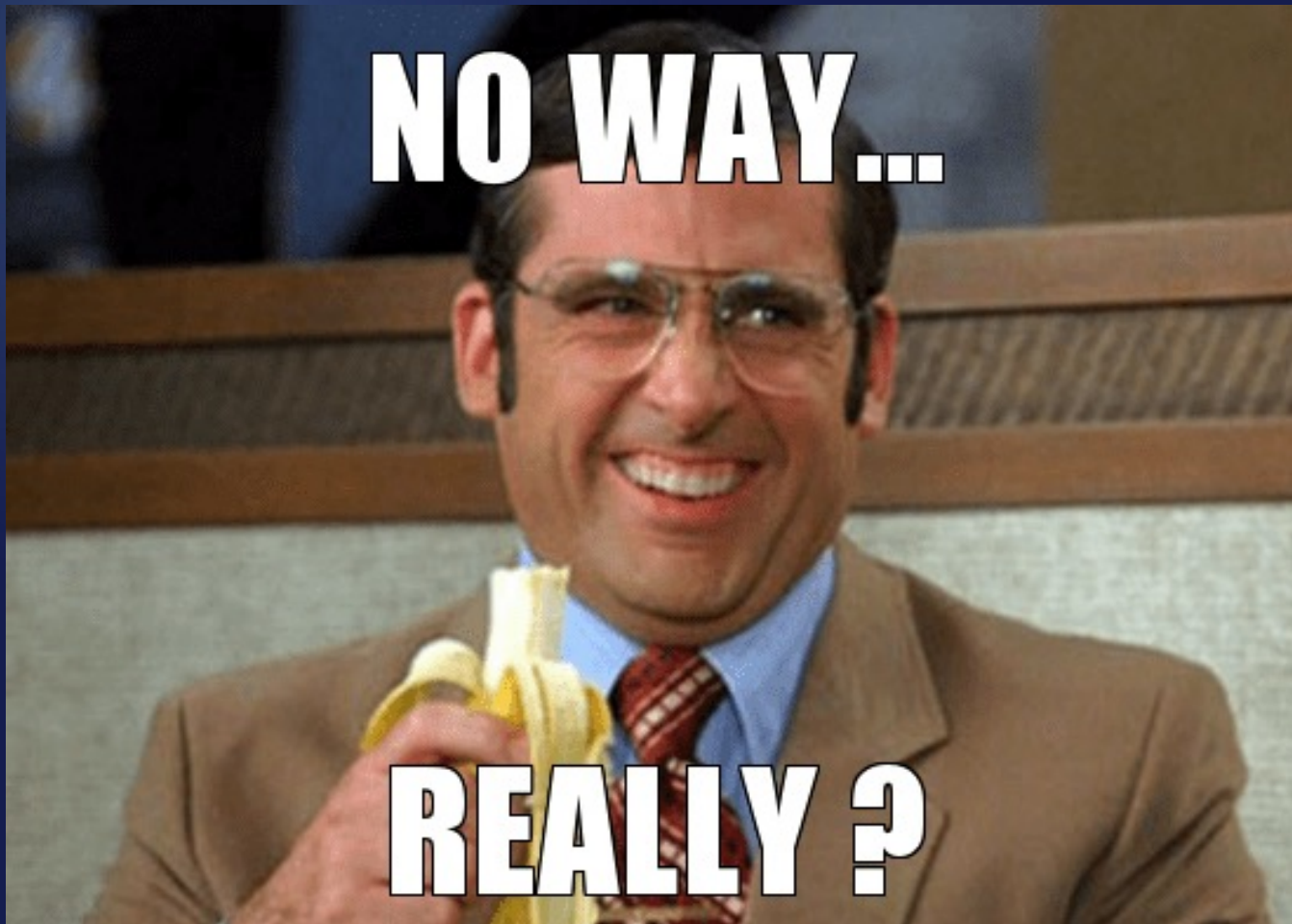


Deltatre
Innovation
Lab



#GlobalAIBootcamp 

Che cos'è OpenAI?



Avrete
sicuramente
sentito parlare
almeno di
ChatGPT....

Modelli di Generative AI

Artificial Intelligence

Machine Learning

Deep Learning

Generative AI

"Azure" OpenAI

Microsoft ha una **partnership** con OpenAI per fornire questo servizio tramite Azure a livello enterprise:

- Possiamo utilizzare l'infrastruttura di Azure, quindi **security, compliance** e la disponibilità da **diverse Region**.
- Usare i modelli OpenAI attraverso prodotti Microsoft come **Azure AI**.
- Sfruttare tutta la **potenza dei workloads** OpenAI



"Azure" OpenAI

Azure OpenAI è composta da **4 componenti** principali:

- Modelli Generative AI **già pronti**
- Possibilità di **personalizzare** i modelli attraverso l'operazione di **fine-tune** con i **propri dati**
- Strumenti di **controllo** per implementare un uso **responsabile** dei modelli
- Enterprise-grade **security** con role-based access control (**RBAC**) e **private networks**



Cosa possiamo fare

Possiamo fare 3 categorie di cose:

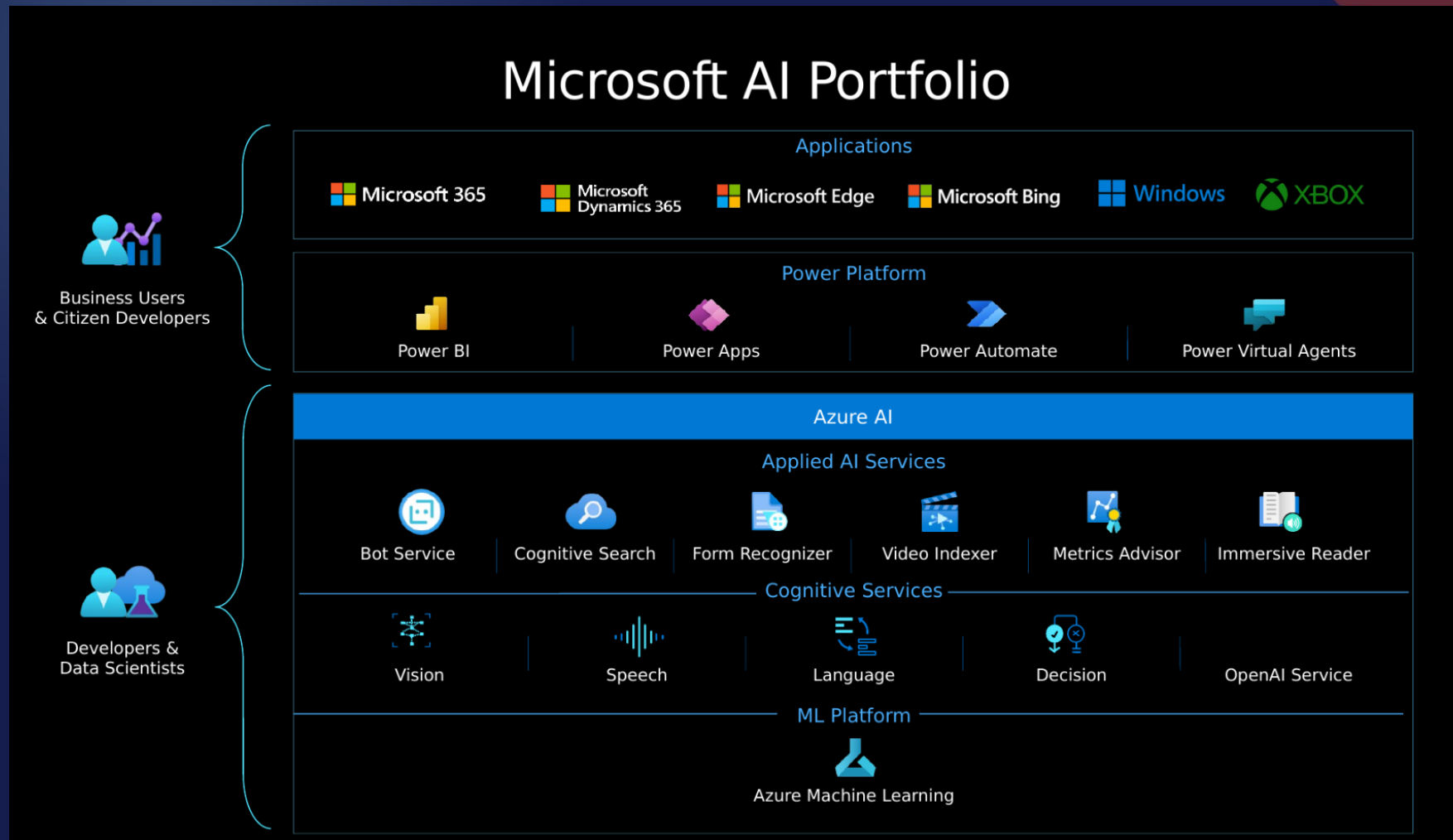
- Generare contenuti in Linguaggio Naturale
 - Text completion: generare e editare testo
 - Ricerca, classificazione e comparazione di testo
- Generare, editare e analizzare codice (in tantissimi linguaggi)
- Generare e editare immagini



Modelli OpenAI



Azure OpenAI's negli Azure AI services



Come possiamo utilizzare Azure OpenAI

Al momento è necessario **richiedere l'accesso** a Azure OpenAI compilando un form. Una volta avuto l'accesso:

- **Creare la risorsa** su Azure
- **Eseguire il deploy** di uno o più modelli (esistenti o personalizzati) utilizzando **Azure OpenAI Studio**



Come possiamo utilizzare Azure OpenAI

Cognitive Services | Azure OpenAI Studio - Preview

[Azure OpenAI Studio](#)

[Privacy & cookies](#)

Get started with Azure OpenAI

Perform a wide variety of natural language tasks with Azure OpenAI, including copywriting, summarization, parsing unstructured text, classification, and translation.

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](#)



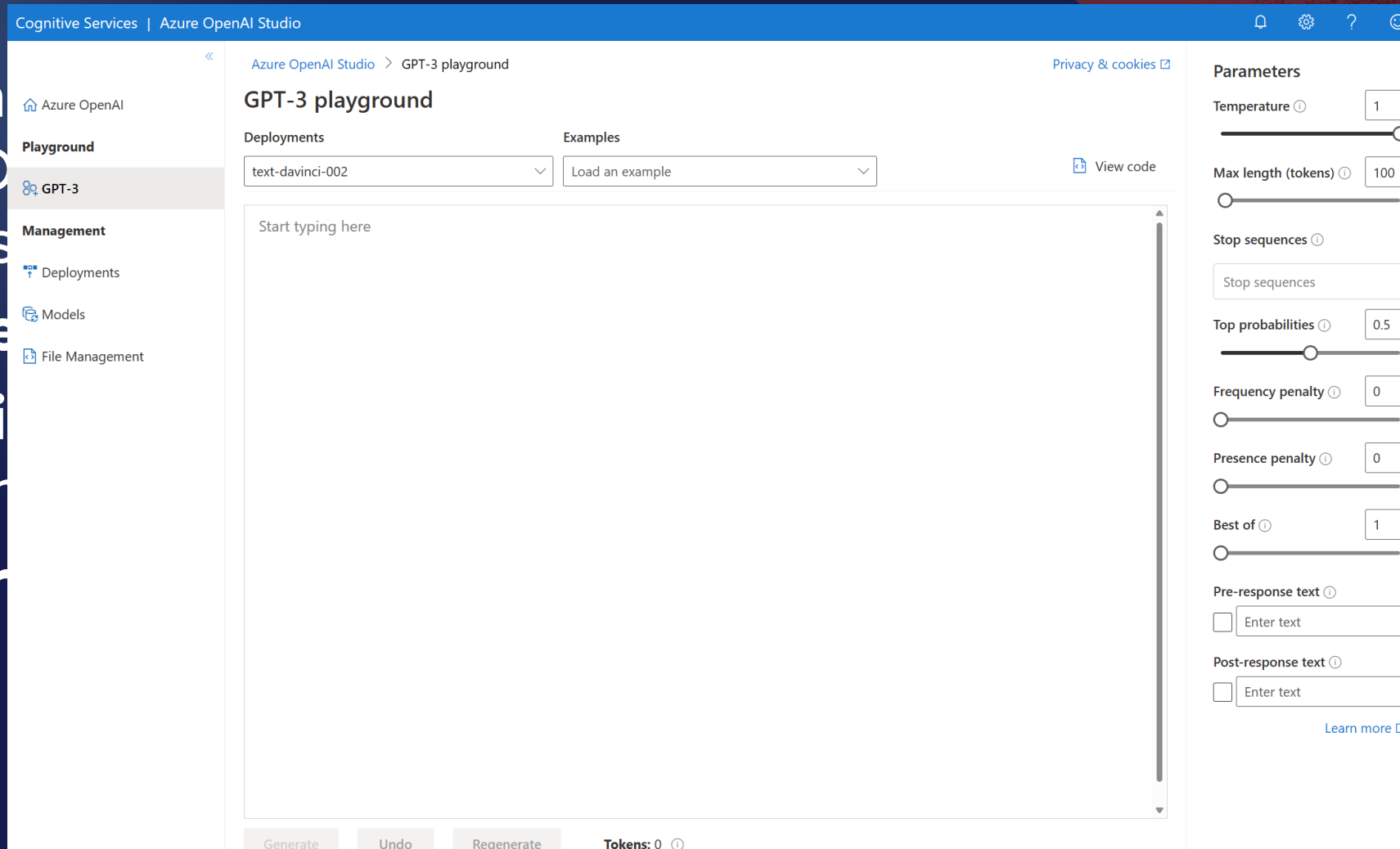
Generate New Product Names

Create product names from examples words.

[Learn more](#)

Come possiamo utilizzare Azure OpenAI

Al momento
Azure OpenAI
l'accesso
- Creare
- Eseguire
personale
- Testare



Come possiamo utilizzare Azure OpenAI

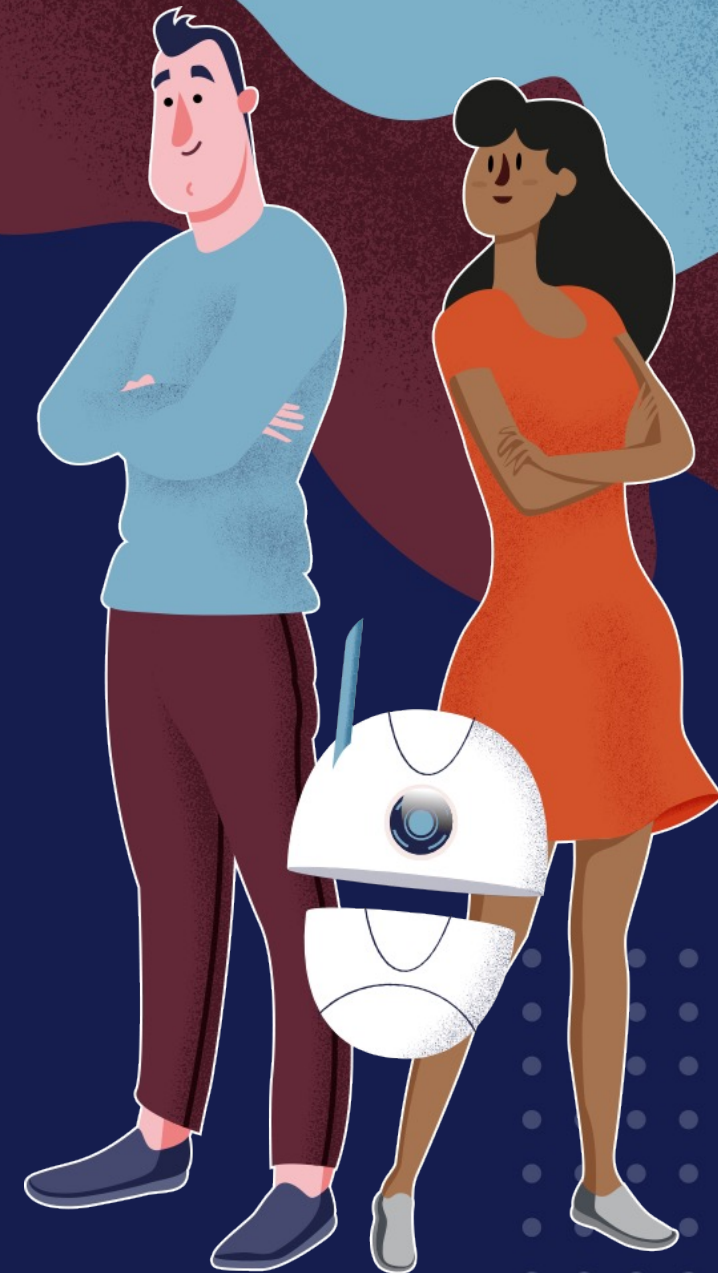
Al momento è necessario **richiedere l'accesso** a Azure OpenAI compilando un form. Una volta avuto l'accesso:

- **Creare la risorsa** su Azure
- **Eseguire il deploy** di uno o più modelli (esistenti o personalizzati) utilizzando **Azure OpenAI Studio**
- **Testare** il modello nel **Playground**
- **Utilizzare** il servizio mediante **libreria Python** o chiamata **REST** (no **NuGet** al momento ☹️)



DEMO 1

Creare un risorsa OpenAI



I modelli OpenAI di linguaggio naturale

I modelli Azure OpenAI capiscono il linguaggio naturale e forniscono risposte in linguaggio naturale:

- I modelli sono addestrati su **parole** o **blocchi di caratteri** chiamati **tokens**.
- Per esempio la parola **hamburger** viene spezzata nei tokens **ham**, **bur**, e **ger**, mentre una parola come **pera** è un singolo token.
- Questi tokens sono **mappati in vettori** usati per addestrare un modello di machine learning.
- Quando questi modelli ricevono una richiesta, **il testo passato come parametro viene spezzato in tokens**.

Modelli GPT per la generazione del linguaggio naturale

I modelli **Generative pre-trained transformer (GPT)** sono bravissimi sia a comprendere che a generare linguaggio naturale:

- I modelli GPT **sono numerati** in base alla versione, **GPT-2** o **GPT-3** ad esempio.
- **Azure OpenAI offre l'accesso a GPT-3** e presto supporterà GPT-3.5.

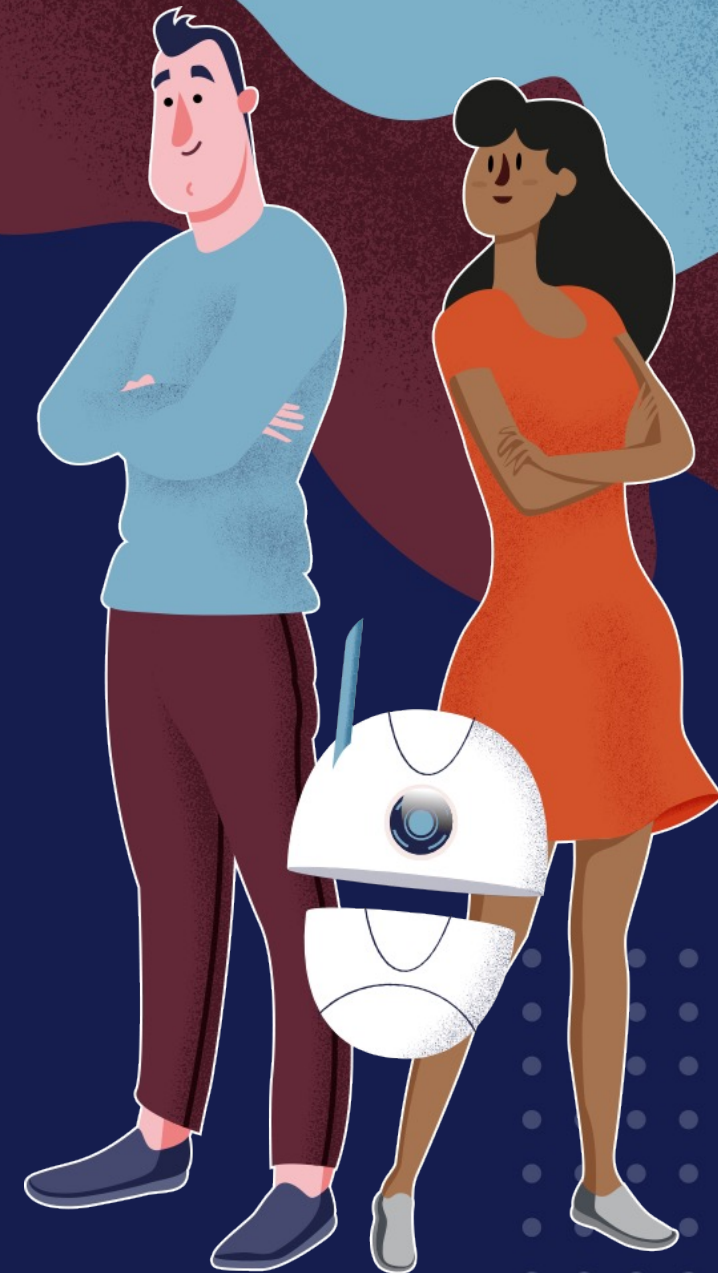
Cosa può fare un modello GPT-3?

Dato un testo in input, chiamato **prompt**, GPT prova a inferire, o indovinare, il contesto della domanda. Gli possiamo chiedere di:

- Eseguire la sintesi di un testo (**summarizing**)
- **Classificare** un contenuto
- **Generare** nomi o frasi a partire da una descrizione
- **Tradurre** testi
- **Rispondere** a domande
- **Suggerire** contenuti

DEMO 2

Usare il playground per generare contenuti

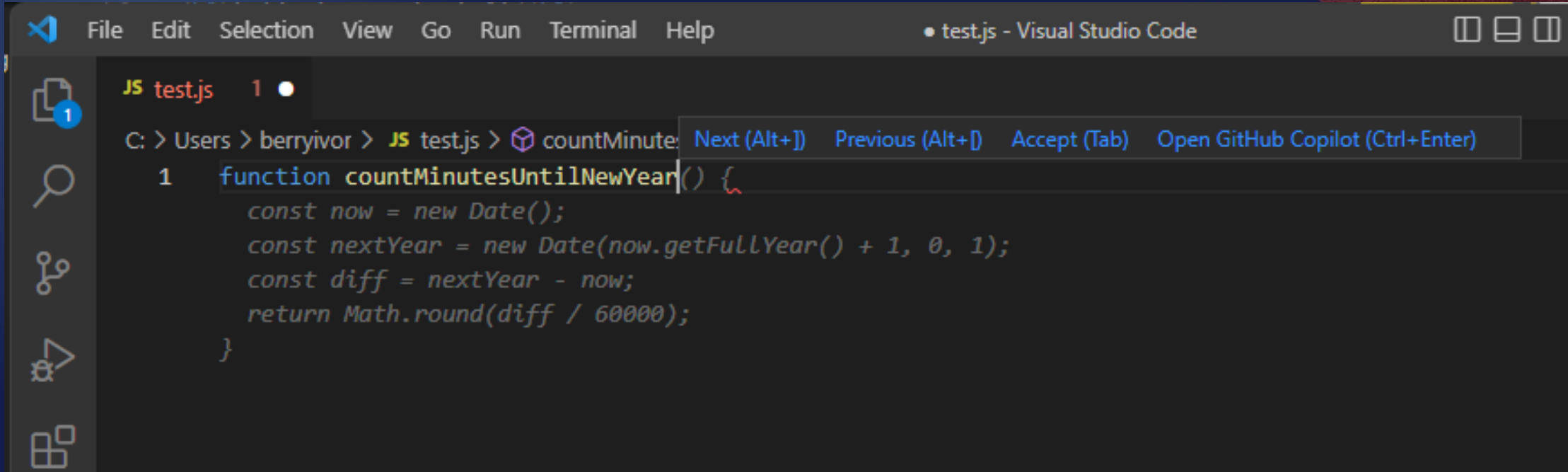


Generazione di codice con OpenAI

E' possibile chiedere, in linguaggio naturale, a un modello OpenAI anche di **generare del codice**:

- Questa famiglia di modelli è chiamata **Codex** e funziona con **tantissimi linguaggi** come C#, JavaScript, Perl, PHP, Python
- Un esempio molto valido è **GitHub Copilot**, nato dalla collaborazione tra **GitHub** e **OpenAI**, un assistente nella scrittura del codice fornito come **plugin** in **Visual Studio Code**

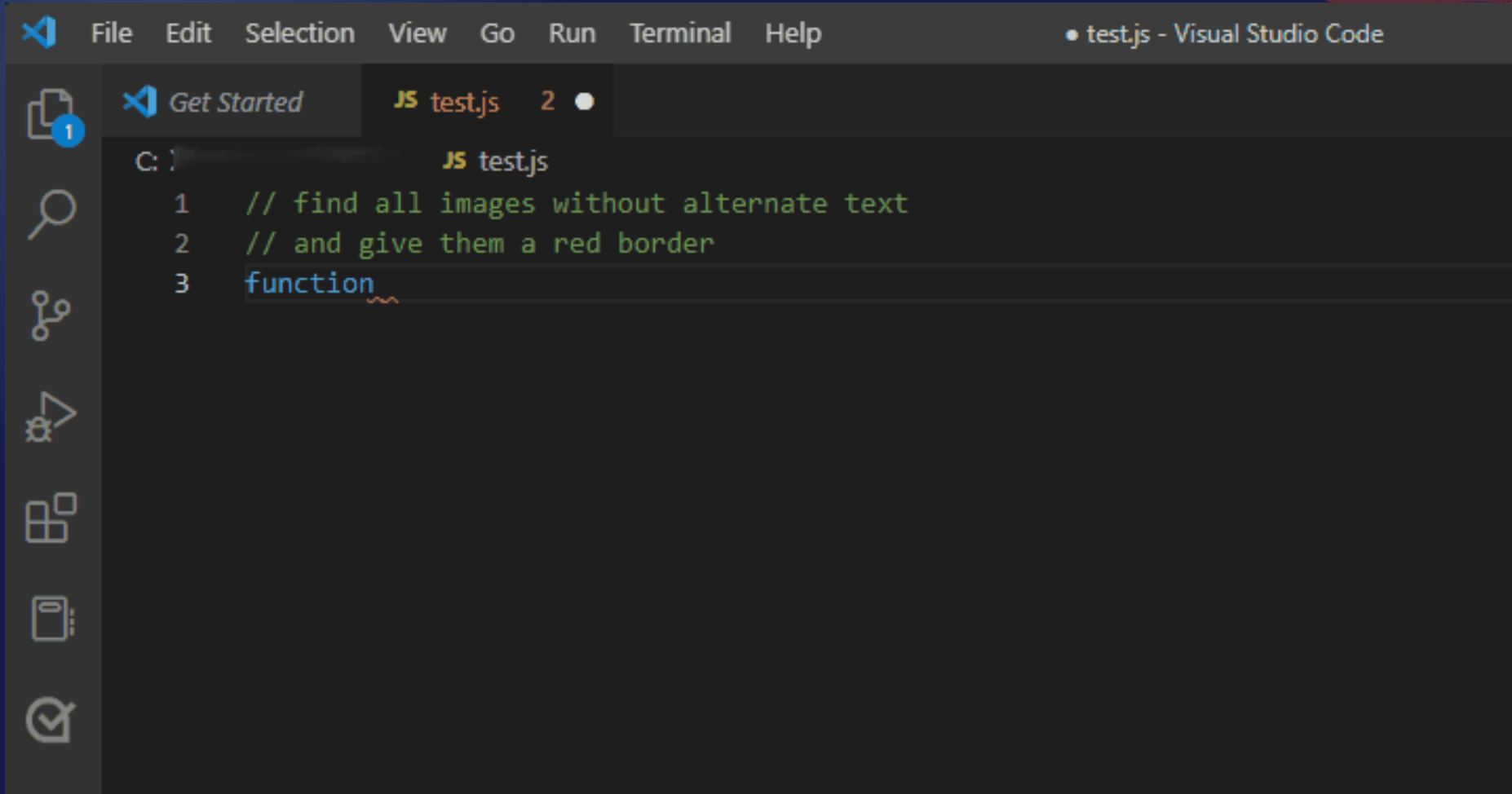
GitHub Copilot



```
JS test.js 1 ●
C: > Users > berryivor > JS test.js > countMinutesUntilNewYear()
1 function countMinutesUntilNewYear() {
    const now = new Date();
    const nextYear = new Date(now.getFullYear() + 1, 0, 1);
    const diff = nextYear - now;
    return Math.round(diff / 60000);
}
```

Next (Alt+J) Previous (Alt+I) Accept (Tab) Open GitHub Copilot (Ctrl+Enter)

GitHub Copilot



The screenshot shows the Visual Studio Code interface with the menu bar (File, Edit, Selection, View, Go, Run, Terminal, Help) and a tab for 'test.js - Visual Studio Code'. The left sidebar contains icons for Explorer (with a '1' badge), Search, Source Control, Run and Debug, Extensions, Accounts, and Testing. The main editor area displays the file 'test.js' with the following code:

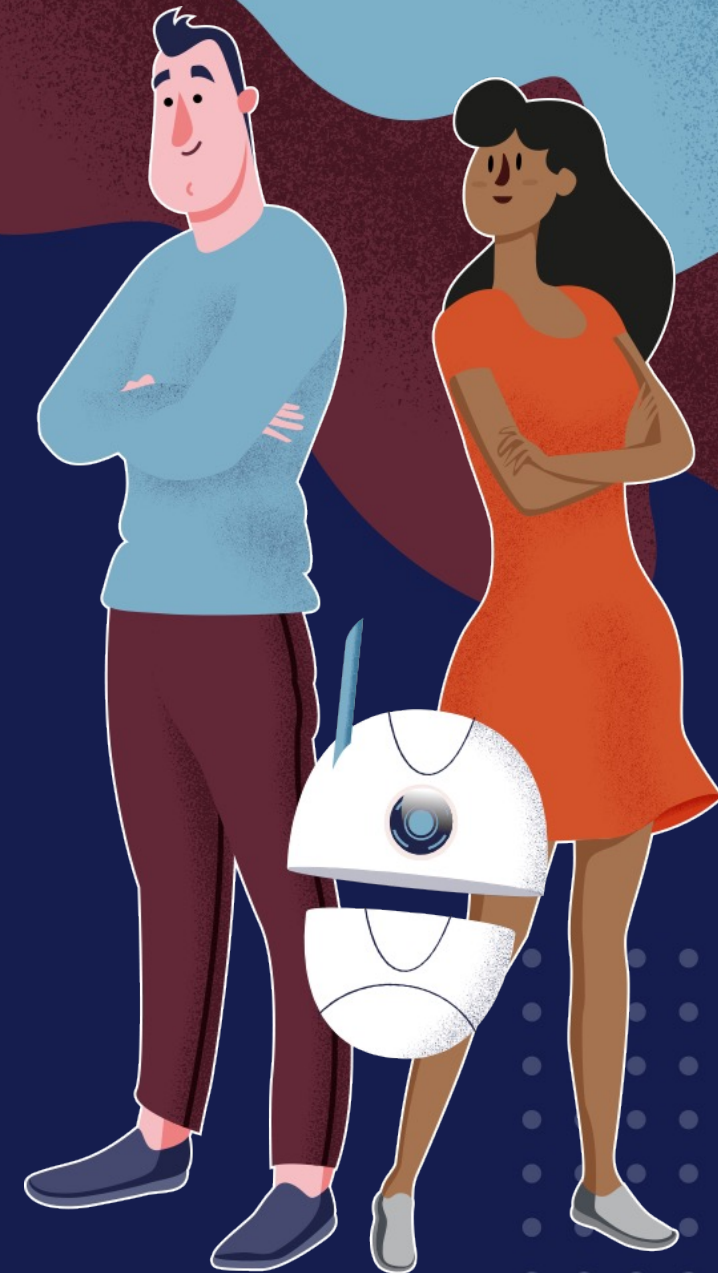
```
C:\> JS test.js  
1 // find all images without alternate text  
2 // and give them a red border  
3 function
```

GitHub Copilot suggestions are visible as a light blue bar at the bottom of the editor, showing the following code:

```
function findImagesWithoutAlternateText() {  
  const images = document.getElementsByTagName('img');  
  for (let i = 0; i < images.length; i++) {  
    if (!images[i].alt) {  
      images[i].border = '2px solid red';  
    }  
  }  
}
```


DEMO 3

Usare il playground per generare codice



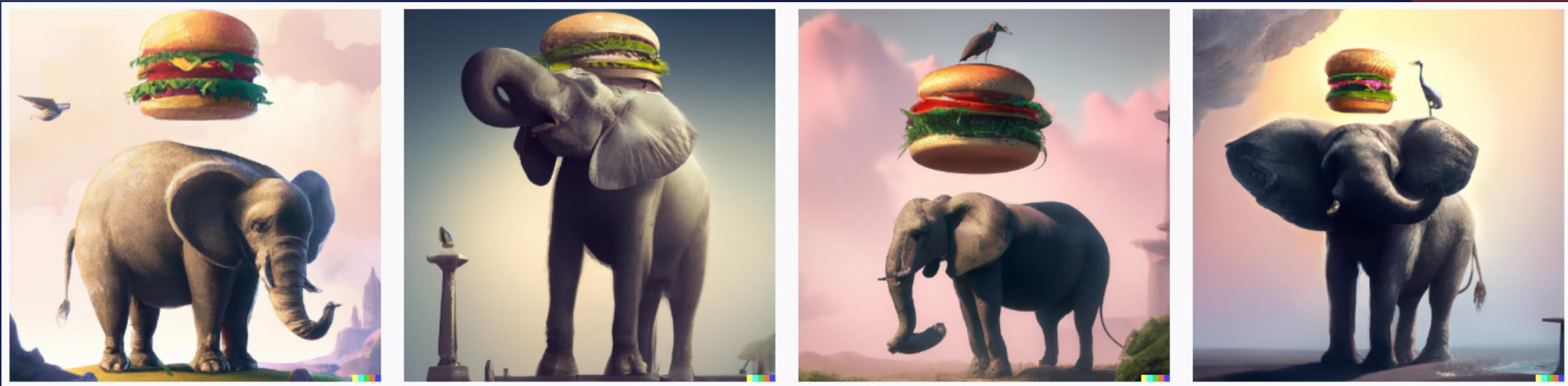
Generare immagini con i modelli OpenAI

E' possibile generare o editare immagini a partire da una richiesta in linguaggio naturale:

- Questa tipologia di modello si chiama **DALL-E**, e anche in questo caso il nome è seguito dal numero di versione, ad esempio DALL-E 2
- **L'accesso** a questi modelli al momento **é solo su invito** (sono ancora in attesa 😞), ma possiamo farci una idea del suo funzionamento

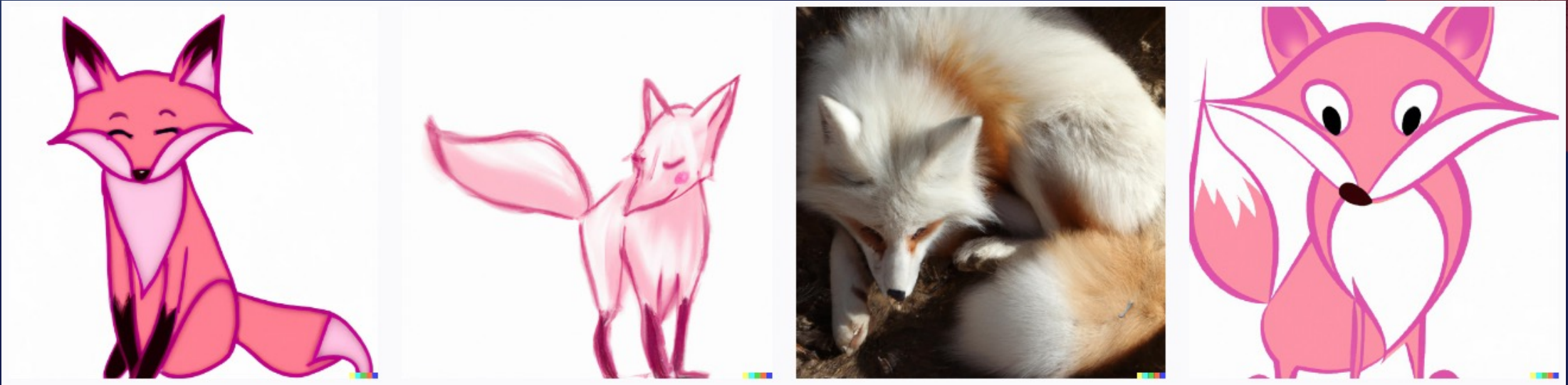
Generare una immagine da un testo

“an elephant standing with a burger on top, style digital art”



Generare una immagine da un testo

"a pink fox"



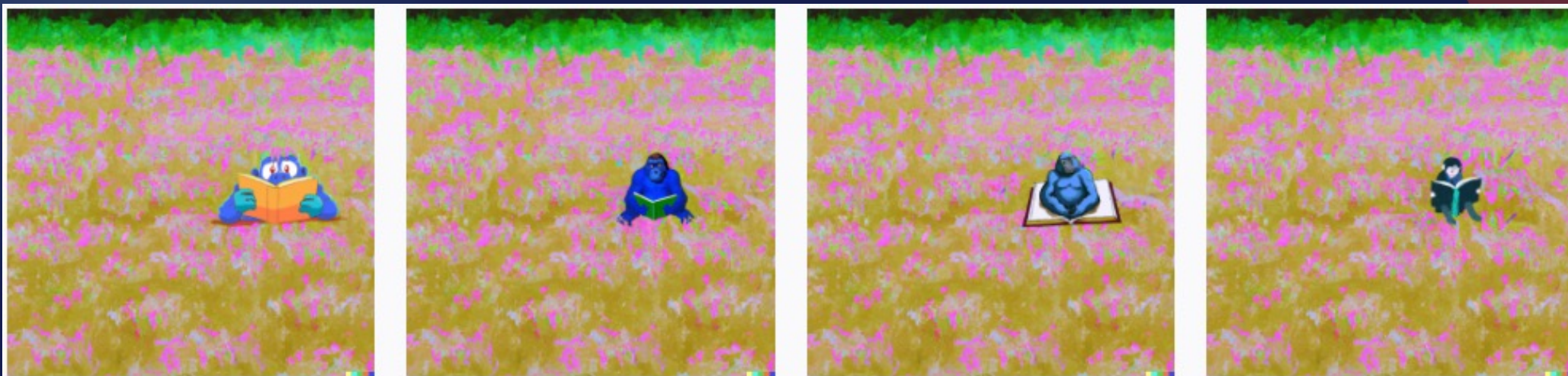
Generare una immagine da un testo

“a pink fox running through a field, in the style of Monet”



Modificare una immagine

“blue gorilla reading a book in a field”



Accesso alle API di Azure OpenAI

Come tutti i servizi Azure, possiamo accedere a Azure OpenAI utilizzando delle API REST. Ci servono:

- Un **client HTTP** nel linguaggio scelto
- **L'endpoint del servizio** (<https://<nome>.openai.azure.com/>)
- **Nome** che abbiamo dato al **deploy** che abbiamo fatto e che vogliamo utilizzare
- Una delle due **key** della risorsa

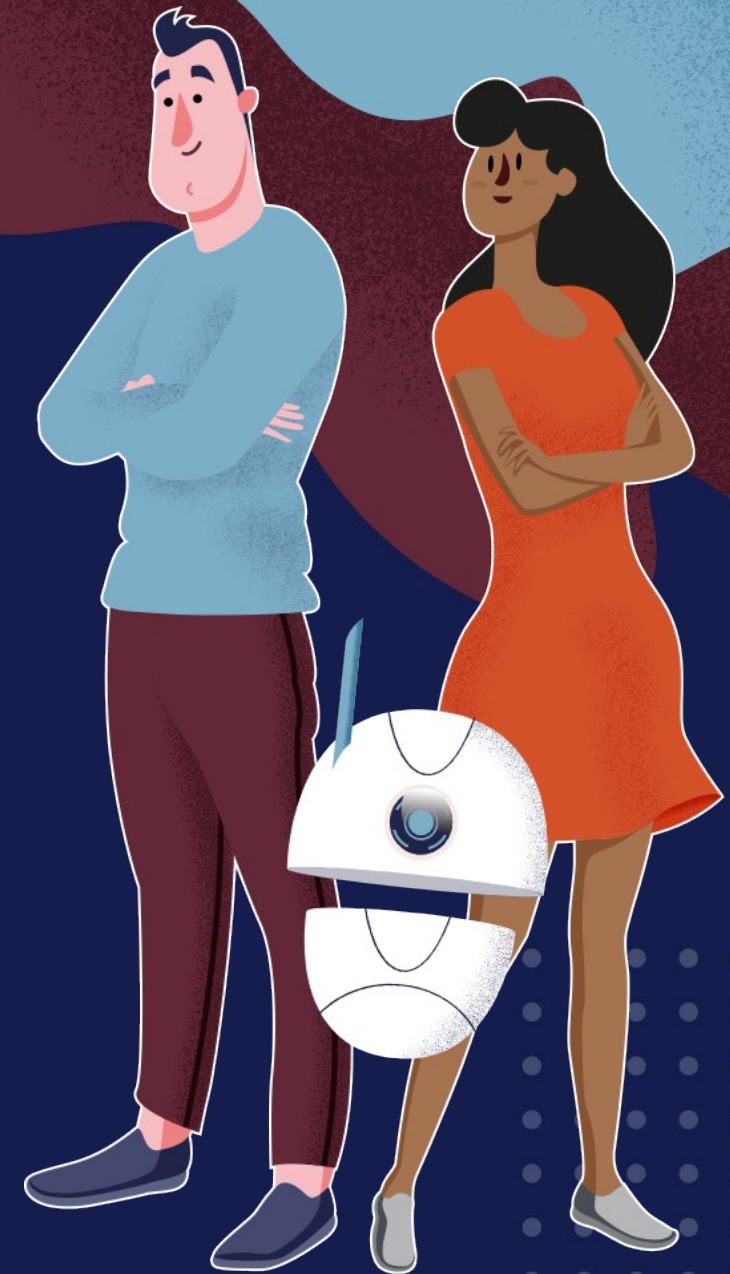
Accesso alle API di Azure OpenAI

A questo punto ci basta:

- Eseguire una **POST HTTP** all'endpoint così costruito:
 - endpoint + "openai/deployments/" + nome_deploy + "completions?api-version=2022-12-01"
- Aggiungere un **header HTTP "api-key"** con la key della risorsa
- Passare nel **body della Request** almeno la proprietà **"prompt"**
- Eseguire il **parsing della risposta JSON** per recuperare le informazioni fornite dal servizio

DEMO 4

Costruiamo il nostro playground
utilizzando C# e Blazor



Personalizzare i modelli OpenAI

I modelli Azure OpenAI, al pari di molti Cognitive Services, **possono essere personalizzati** eseguendo il cosiddetto **fine-tuning**:

- Questo ci permette di **utilizzare i nostri dati** per addestrare i modelli
- E' molto importante **verificare i risultati prodotti** da questo tipo di modelli
- I modelli di Azure OpenAI sono addestrati con dati che arrivano fino a **Ottobre 2019**, quindi non contengono informazioni successive a quella data.
- <https://learn.microsoft.com/en-us/azure/cognitive-services/openai/how-to/fine-tuning>

Quanto mi costi?

Pricing details:

Instance	Series	Models	Inferencing per 1,000 tokens	Hosting per hour
Standard	Base Series	Ada	€0.000369	N/A
		Babbage	€0.000462	N/A
		Curie	€0.001845	N/A
		Davinci	€0.018450	N/A
	Base Series Fine-tuned	Ada	€0.000369	€0.0462
		Babbage	€0.000462	€0.0738
		Curie	€0.001845	€0.2214
		Davinci	€0.018450	€31.3639
	Codex Series	Code-Cushman	€0.022140	N/A
		Code-Davinci	€0.092247	N/A
	Codex Series Fine-tuned	Code-Cushman	€0.022140	€0.4982/hour
	Embeddings	Ada	€0.003690	N/A
		Babbage	€0.004613	N/A
		Curie	€0.018450	N/A
		Davinci	€0.184494	N/A

Accesso limitato a Azure OpenAI

Una delle ragioni per le quali l'accesso a Azure OpenAI è attualmente limitato è garantire un uso etico e controllato di questi servizi:

- Bisogna sottomettere una richiesta sia per sperimentare con il servizio sia per usarla in produzione, spiegando il caso d'uso che si vuole implementare.
- E' richiesta una ulteriore registrazione per accedere ai tool di content filters e abuse monitoring.

Takeaways

- Possiamo **dare valore alle nostre applicazioni** con la potenza dei modelli OpenAI
- Se **utilizziamo** già **Azure** possiamo sfruttare tutto quello che già conosciamo per dare robustezza agli **scenari enterprise**
- Possiamo facilmente integrare i servizi Azure OpenAI con un semplice **client HTTP** (in attesa di un pacchetto NuGet ufficiale)
- Occhio ai **costi** e alla **“degenerazione”** che può avere l’uso di questi servizi, l’approccio alla **AI** deve essere sempre **responsabile!**

Link utili

- Percorso Learn: <https://learn.microsoft.com/en-us/training/modules/explore-azure-openai/>
- Documentazione ufficiale: <https://learn.microsoft.com/en-us/azure/cognitive-services/openai/>
- Reference API: <https://learn.microsoft.com/en-us/azure/cognitive-services/openai/reference#management-apis>

Questions & Answers



Michele Aponte

CEO&CTO @ Ellycode / CEO @ Blexin / CTO @ Improove Skills

Founder @ Blazor Developer Italiani / MVP Developer Technologies & AI



apomic80



apomic80



apomic80



apomic80



apomic80

