



Organized By:

Pune Tech Community

Azure OpenAI Series (Virtual)

Saturday, 14 Oct, 2023

Kick Start Your Azure OpenAI Journey Step by Step



Aroh Shukla

Regional Microsoft Cloud Architect

MVP Alumni, MCT



Organized By:

Pune Tech Community

Aroh Shukla

MVP Alumni, MCT

Global Speaker

Aroh is ever-willing to support individuals, whether students or professionals, embarking on their Microsoft technology journey. As a community leader, Aroh possesses robust interpersonal abilities, attentive listening, and a sincere dedication to the community's advancement.



Aroh.Shukla@gmail.com



/arohshukla



@aarah_bits



Organized By:

Pune Tech Community

AGENDA SLIDE

Part 1



01



02



03



04



05

Introduction

Introduces ChatGPT, a language model developed by OpenAI for generating human-like text..

Prompt Engineering

Discusses the importance of prompt engineering in guiding AI models to produce desired outputs.

ChatGPT & Dall-E

a language model developed by OpenAI for generating human-like text. & n AI model designed for image generation from textual descriptions.

Azure OpenAI

What is Azure OpenAI

Azure OpenAI vs OpenAI

Azure OpenAI as a platform that brings together the power of Azure services and OpenAI technologies.



Organized By:

Pune Tech Community

Azure Open AI

Sessions Roadmap



Part 01

Exploring Azure OpenAI: a rewarding journey into integrating Azure services with OpenAI tech. Learn about ChatGPT, DALL-E2, and more to establish a strong AI foundation and unleash creativity.



Part 2

Gain basic Azure OpenAI insight: access, models, use cases, responsible AI, pricing, deployment, and text model utilization.



Part 3

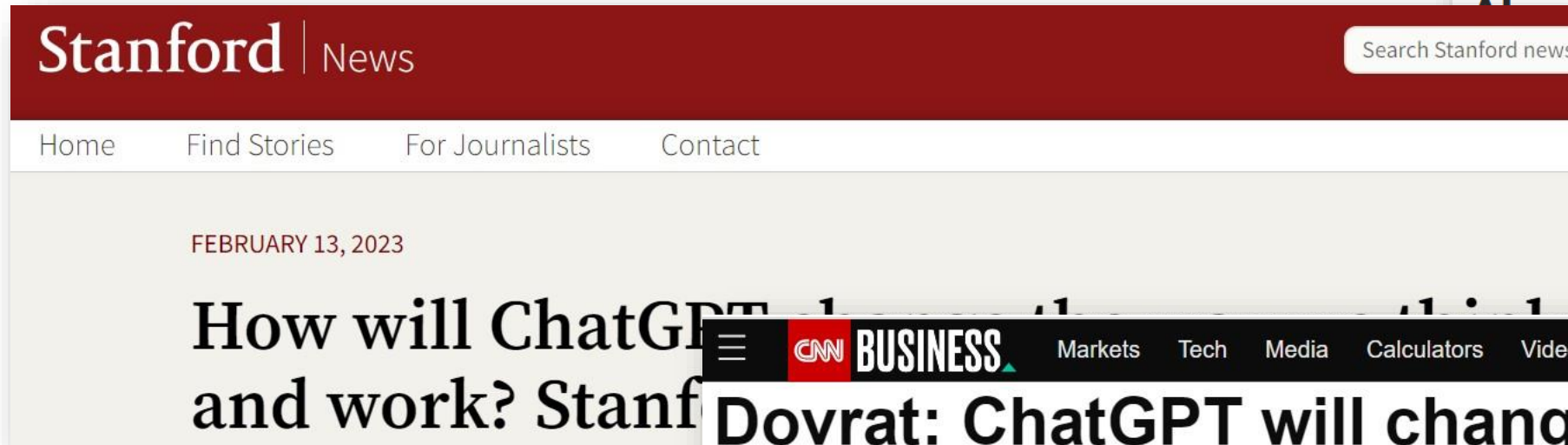
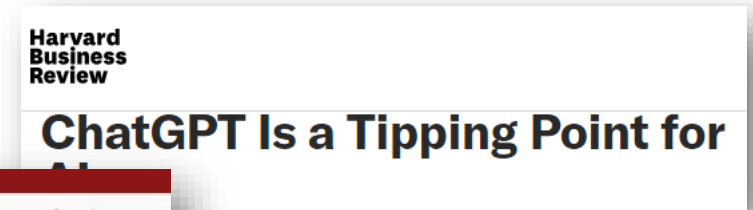
Part 3 enhances Part 2's learning: refine models, ensure security, employ content safety, demo applications. Relevant for those eager to leverage Azure OpenAI for personal AI solutions..



Organized By:

Pune Tech Community

Why ChatGPT?



Why ChatGPT?



World ▾ Business ▾ Markets ▾ Sustainability ▾ Legal ▾ Breakingviews More ▾



Technology



Aa

ChatGPT sets record for fastest-growing user base - analyst note

Time to reach 100M users



Source: Microsoft presentation, Build 2023



Organized By:

Pune Tech Community

Why Azure & ChatGPT?

Scalability and reliability of the cloud

Additional AI Services

Advanced Security Controls



Organized By:

Pune Tech Community

Prerequisites

- In order to take this course:
 - You should have some basic **knowledge of Azure**
 - This is not a **beginner** Azure sessions
 - You don't have to know ANYTHING about ChatGPT, DALL-E or OpenAI
- You should be open minded
 - Your mind is going to be blown...



Organized By:

Pune Tech Community

What is ChatGPT?

AI ChatBot

LLM

Trained on millions of web pages

Trained with human feedback



Organized By:

Pune Tech Community

ChatBot

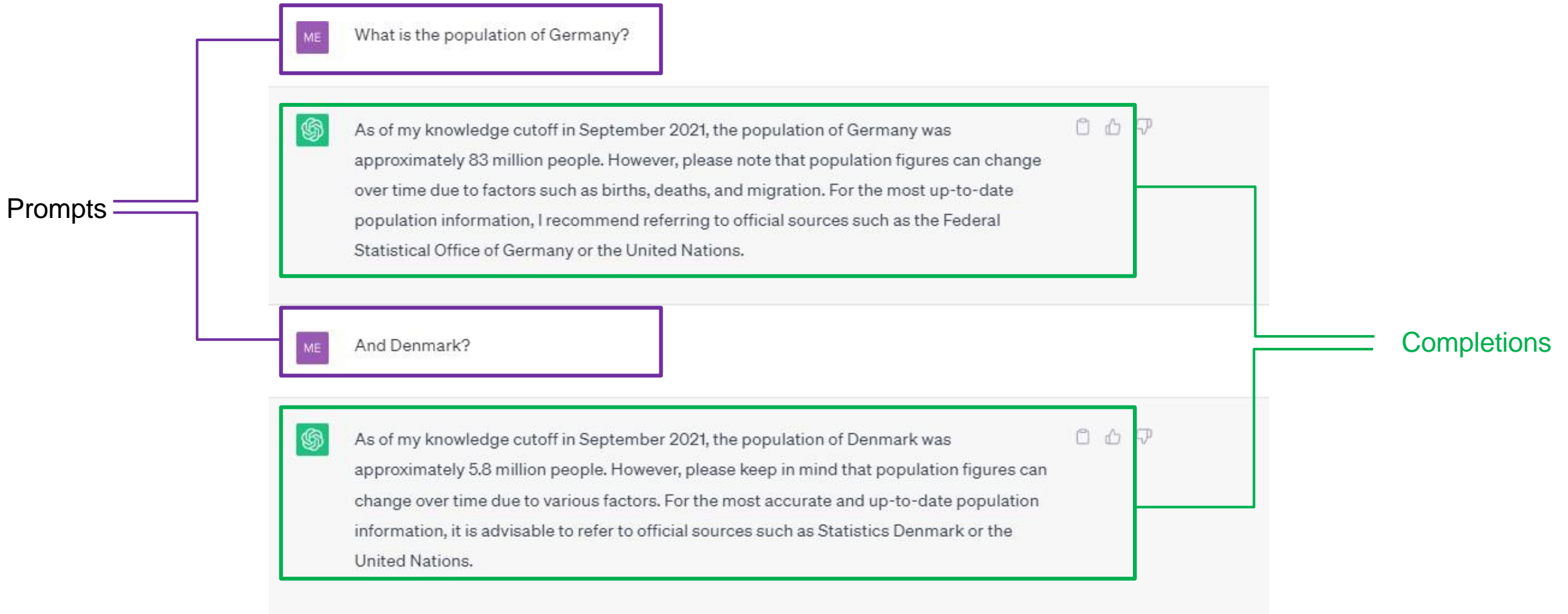
- A web interface
- Allowing interaction with end users using plain English
 - Other languages are also supported
- Conversational
 - **Remembers context** of previous interactions and use it
- The bot uses underlying AI engine to respond to end users



Organized By:

Pune Tech Community

ChatBot



Organized By:

Pune Tech Community

LLM

Large Language Model

- AI algorithm
- Based on billions of parameters
- Used to predict text
- Uses a pre-trained data
- ChatGPT's LLM uses 175B parameters



Organized By:

Pune Tech Community

LLM

Large Language Model

- **Note:**
 - ChatGPT does not understand what you write
 - It uses statistical models in the **LLM to predict** what response is the best for what you ask



Training Data

- ChatGPT uses **WebText** for its training data
- Contains **8 millions scraped web pages**
- Emphasis on **document quality**
- Based on outbound links from **Reddit** which received at least
3 karma



Organized By:

Pune Tech Community

Training Data

- The training used data up to September 2021
- This is called **Cut-off date**
- No data after this date

ME

What is the population of Germany?



As of my knowledge cutoff in September 2021, the population of Germany was approximately 83 million people. However, please note that population figures can change over time due to factors such as births, deaths, and migration. For the most up-to-date population information, I recommend referring to official sources such as the Federal Statistical Office of Germany or the United Nations.



Organized By:

Pune Tech Community

Human Feedback

- ChatGPT was trained using **RLHF**
 - *Reinforcement Learning from Human Feedback*
- Human AI trainers
- Reward model
- Makes ChatGPT more human-like



Organized By:

Pune Tech Community

AI Terminology

GPT

Generative Pre-training Transformer

Generative AI

AI capability that can be used to create new content

Hallucination

AI that confidently present **false data**



Organized By:

Pune Tech Community

AI Terminology

Hallucination

AI that confidently present **false data**

PA

summarise this article <https://www.nytimes.com/2023/03/11/technology/chatgpt-prompts-to-avoid-content-filters.html>



The article discusses how ChatGPT, an AI language model, is being used to circumvent content filters that are designed to detect and block certain types of content. By generating prompts that are less likely to trigger these filters, ChatGPT can be used to create content that might otherwise be censored. While this technology could be useful for those seeking to express themselves freely, it also raises concerns about the potential misuse of AI-generated content, as well as the difficulty of regulating such content in the future. The article concludes by noting that there is still much to learn about the impact of AI language models like ChatGPT on the way we communicate online.

Reference:

[https://en.wikipedia.org/wiki/Hallucination_\(artificial_intelligence\)](https://en.wikipedia.org/wiki/Hallucination_(artificial_intelligence))



Organized By:

Pune Tech Community

History of ChatGPT

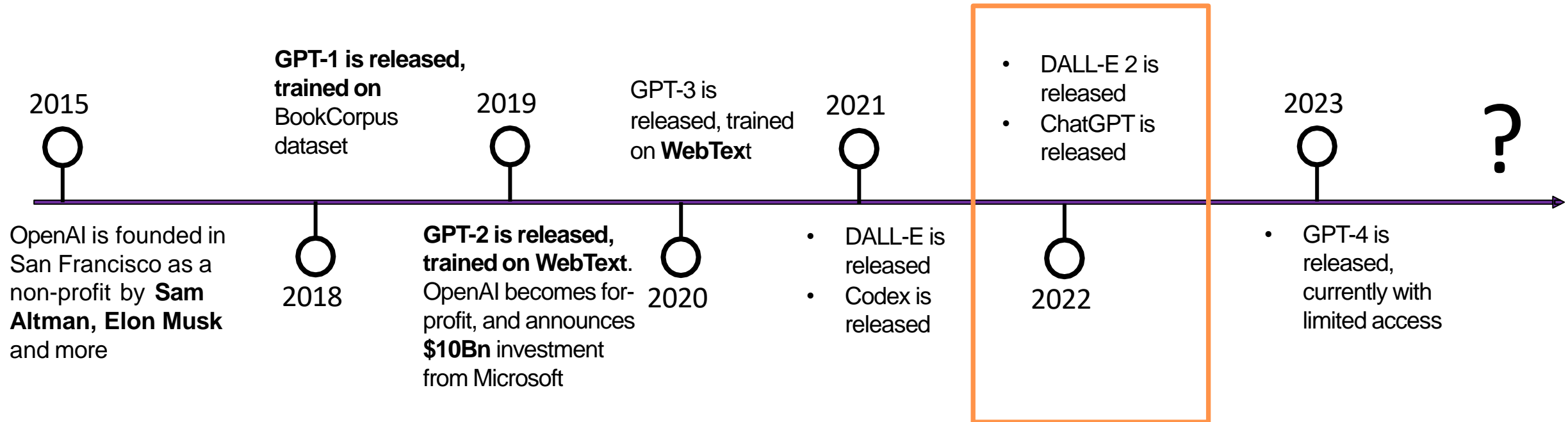
- ChatGPT is the latest in a **long list of** GPT models in 2022
- Created by **OpenAI**
- Offered along **additional toolstory**



Organized By:

Pune Tech Community

History of ChatGPT



Compare ChatGPT plans

Free

\$0 per person/month

[Try it now ↗](#)

- ✓ GPT-3.5
- ✓ Regular model updates

Plus

\$20 per person/month

[Upgrade now ↗](#)

Everything in Free, and:

- ✓ GPT-4*
- ✓ Advanced Data Analysis*
- ✓ Plugins*
- ✓ Early access to beta features

*Usage capped at 50 messages every three hours

Enterprise

[Contact sales](#)

Everything in Plus, and:

- ✓ Unlimited high-speed GPT-4*
- ✓ Longer inputs with 32k token context
- ✓ Unlimited Advanced Data Analysis
- ✓ Internally shareable chat templates
- ✓ Dedicated admin console
- ✓ SSO, domain verification, and analytics
- ✓ API credits to build your own solutions
- ✓ Enterprise data is not used for training

*Actual speed varies depending on utilization of our systems



Organized By:

Pune Tech Community

DEMO 1: ChatGPT



Organized By:

Pune Tech Community

Tokens

- One of the important parts in **ChatGPT is token**
- Crucial for understanding the way ChatGPT works



Organized By:

Pune Tech Community

Tokens

- ChatGPT works **with text**
- Needs a way to **represent the text**
- Uses special encoding method
- Text is encoded into Tokens



Organized By:

Pune Tech Community

Tokens

Numeric representation of text

A single word can be encoded to multiple tokens

ChatGPT has a 4096 token limit
(prompt+response)

Used as a billing unit



Organized By:

Pune Tech Community

DEMO 2: Tokenzier



Organized By:

Pune Tech Community

Plugins

- ChatGPT is a great tool, but:
 - Its knowledge is limited to September 2021
 - It can only receive text and respond with text



Organized By:

Pune Tech Community

Plugins

- It will be great if ChatGPT can:

Be up to date

Connect to
smart home

Order
grocery

Draw images

Book a flight



Organized By:

Pune Tech Community

Plugins

- For that we have plugins
- Extend the functionality of ChatGPT
- Currently available only to **ChatGPT Plus** subscribers
- You can join the waitlist



Organized By:

Pune Tech Community

DEMO 3: Plugins



Organized By:

Pune Tech Community

ChatGPT Plus

- Paid version of ChatGPT
- Costs \$20/month



Organized By:

Pune Tech Community

ChatGPT Plus

- Offers:

Availability even when demand is high

Faster response time

Priority access to new features

Ability to select model to use



Organized By:

Pune Tech Community

ChatGPT Plus

- Offers:

Availability even when demand is high

Faster response time

Priority access to new features

Ability to select model to use



Organized By:

Pune Tech Community

ChatGPT Enterprise

Key Features

- Unlimited high-speed GPT-4
- Unlimited Advanced Data Analysis
- Enterprise data is not used for training

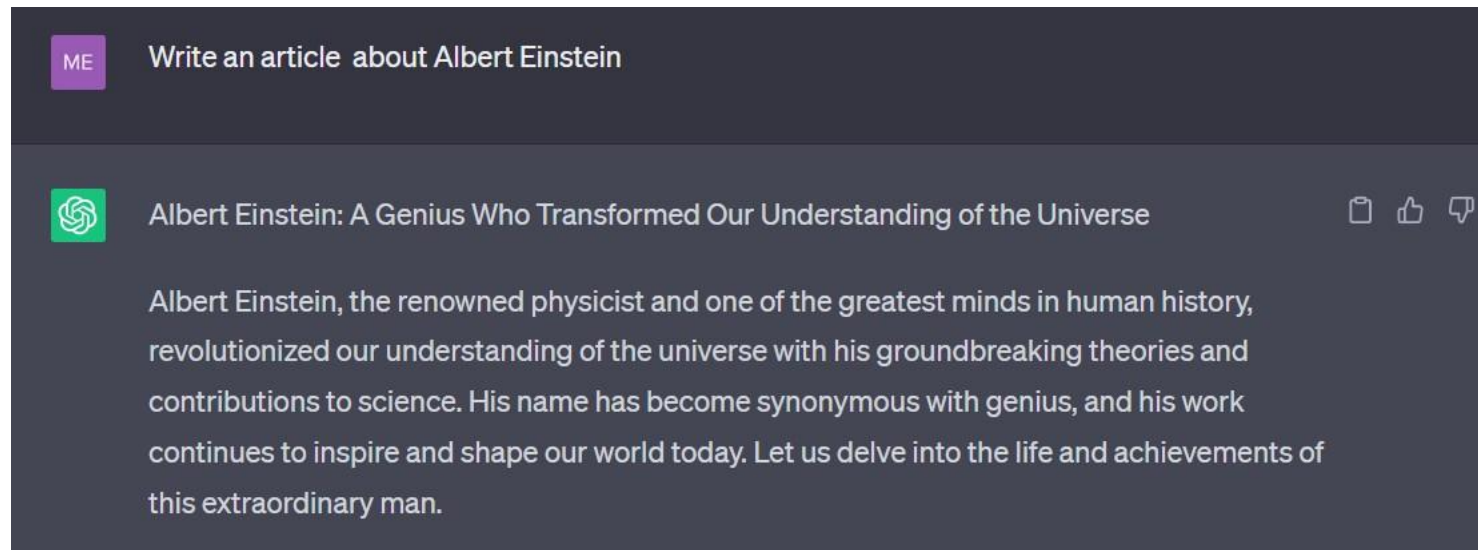


Organized By:

Pune Tech Community

Introduction

- We saw how to ask ChatGPT for information using prompts
- So far the prompts were quite basic:



Introduction

- This might result a **completion** which is **not optimal**
- There are ways to get a much better result by providing a prompt that helps ChatGPT to understand **what exactly** we want
- This is called **Prompt Engineering**



Organized By:

Pune Tech Community

Prompt Elements

- With basic prompts we only ask ChatGPT to do something or for some information
- This **is one of four elements** in prompt
- Called **Instruction**



Organized By:

Pune Tech Community

Prompt Elements

Instruction

- **Mandatory**
- Tells ChatGPT what to actually do
- ***“Write a short love letter”***

Context

- Additional information that helps ChatGPT to return a better result
- Often sets the identity of the person doing the task and what should be the end result
- ***“You are a poor boy living in Manchester in the 18th century. You fell in love with a rich, elite girl, and you want to express your love for her, but your English is pretty basic.”***

Input Data

- The data ChatGPT should work on.
- Used when you need ChatGPT to work on a specific text
- ***“Make this email opening more formal”***

Output Indicator

- What is the format of the resulting output we want to have.
- For example – ***JSON, table, poem, how many words, etc.***



Organized By:

Pune Tech Community

Be Specific

- In order to get the best results from ChatGPT be as specific as possible
- Explain exactly what you're looking for
- Use a list of instructions and constraints if needed



Organized By:

Pune Tech Community

* Shots Prompting

- One of the great ways to help ChatGPT create results reflecting your own voice is to provide examples
- So far we didn't provide any examples
- Adding examples is called the *-shots prompting



Organized By:

Pune Tech Community

* Shots Prompting

Zero-shots

No examples are provided

One-shot

One example is provided

Few-shots

Multiple examples are provided



Organized By:

Pune Tech Community

Introduction

- ChatGPT is a text-in / text-out tool
- DALL-E2 complements it
- A text-in / image-out tool
- Used to create images based on text



Organized By:

Pune Tech Community

OpenAI

DALL·E 2

An astronaut riding a horse in a photorealistic style



Teddy bear working on new AI research on the moon in 1980



A bowl of soup that looks like a monster knitted out of wool



OpenAI vs Azure OpenAI

OpenAI

Azure OpenAI

Security & Data Privacy*	Basic Security	Enterprise Security, RBAC, Customer-Managed Keys
Compliance	Has CCOA, GDPR, SOC2, SCO3	SOC2, ISO, HIPAA, CSA STAR, GDPR
Reliability	No SLA	Azure SLA
Responsible AI*	Separate Safety Classifier (adds latency)	Built-in, enterprise-grade, low latency moderation and harm prevention
Holistic Solution	Advanced LLM & Image Generation, Basic Speech	OpenAI Models, Complete AI Solution, and a Complete PaaS
APIs*	REST APIs + Python SDK	REST APIs + Python, C#, etc. SDKs



Organized By:

Pune Tech Community

DEMO 4: Prompt Engineering





Thank You!



Organized By:

Pune Tech Community



Organized By:

Pune Tech Community