



OpenAI und ChatGPT: Neuer Support in der IT!



Thomas Jäkel



Lead Trainer Cloud Infrastructure

Microsoft Certified Trainer since 1999

github.com/www42/openAI





Agenda

Tag 1

Sprachmodelle

- Hello ChatGPT
- Wie funktionieren LLMs?
- Was kann ich damit machen?
- OpenAI API

Tag 2 Internet

Erweiterungen

- GPT 4
- OpenAI + Microsoft = Copilot
- Retrieval Augmented Generation (RAG)

Praktische Übungen mit OpenAI Account und Windows 11

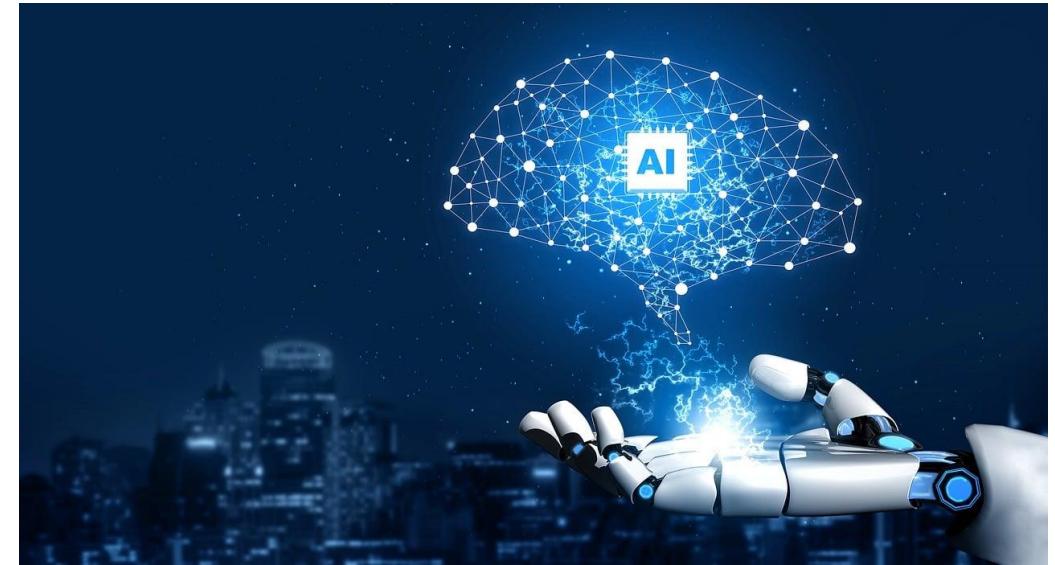


Künstliche Intelligenz?



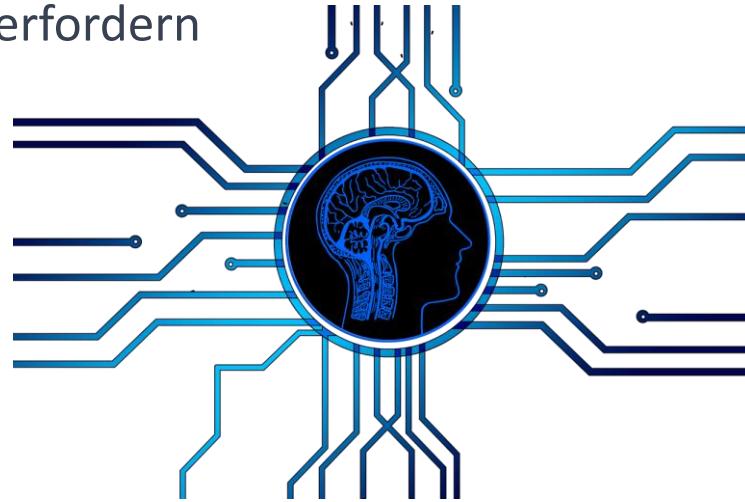
Was ist KI und warum ist sie wichtig?

- Künstliche Intelligenz (KI) hat in den letzten Jahren stark an Bedeutung gewonnen
- Sie ist zu einem Schlagwort in vielen Bereichen unseres Lebens geworden
- Diese Technologie hat das Potenzial, die Art und Weise, wie wir arbeiten, leben und kommunizieren, grundlegend zu verändern



Was ist Künstliche Intelligenz?

- Künstliche Intelligenz bezieht sich auf die Fähigkeit von Maschinen und Computern, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern
- Aktivitäten wie
 - Problemlösung
 - Mustererkennung
 - Sprachverarbeitung
 - Lernen und Entscheidungsfindung
- KI-Systeme basieren auf Algorithmen, die in der Lage sind, riesige Mengen an Daten zu analysieren und daraus Schlüsse zu ziehen
- Sie nutzt Techniken wie maschinelles Lernen, neuronale Netzwerke und tiefe Lernverfahren.





Arten von KI

- Schwache KI:
Diese Form der KI ist auf eine spezifische Aufgabe oder ein begrenztes Aufgabengebiet beschränkt. Ein gutes Beispiel hierfür sind Sprachassistenten wie Siri oder Chatbots. Sie können bestimmte Aufgaben effizient erledigen, sind jedoch nicht in der Lage, menschliche Intelligenz oder Bewusstsein zu reproduzieren.
- Starke KI:
Starke KI, auch als allgemeine KI bezeichnet, würde die Fähigkeit haben, wie ein Mensch zu denken und zu handeln. Diese Form der KI ist noch in der Entwicklungsphase und stellt eine Herausforderung dar, da sie ethische und philosophische Fragen aufwirft.



Warum ist KI wichtig?

- Automatisierung und Effizienz
 - KI kann Aufgaben automatisieren und Prozesse effizienter gestalten. Dies führt zu geringeren Kosten, Zeitersparnis und einer erhöhten Produktivität.
 - In der Industrie können Roboter und autonome Maschinen gefährliche oder repetitive Arbeiten übernehmen, wodurch menschliche Arbeitskräfte entlastet werden.
- Gesundheitswesen
 - Im Gesundheitswesen kann KI dazu beitragen, Diagnosen genauer und schneller zu stellen, personalisierte Behandlungspläne zu erstellen und medizinische Forschung voranzutreiben.
 - Maschinelles Lernen kann enorme Mengen an Patientendaten analysieren, um Muster und Trends zu identifizieren, die für die Verbesserung der Gesundheitsversorgung von entscheidender Bedeutung sind.



Warum ist KI wichtig?

- Bildung
 - KI kann personalisierte Bildungsmöglichkeiten bieten, indem sie den Lernfortschritt von Schülern verfolgt und individuelle Lehrpläne erstellt. Dadurch können Lehrer besser auf die Bedürfnisse ihrer Schüler eingehen und Bildungschancen gerechter gestalten.
- Finanzwesen
 - Im Finanzwesen hilft KI bei der Analyse von Marktbedingungen, um Investitionsentscheidungen zu optimieren.
 - Sie erkennt auch Betrugsversuche in Echtzeit und ermöglicht eine bessere Risikobewertung.
- Umweltschutz
 - KI kann zur Überwachung und Vorhersage von Umweltauswirkungen beitragen, um Naturkatastrophen zu verhindern oder den Umgang mit Ressourcen nachhaltiger zu gestalten.



Warum ist KI wichtig?

- Autonome Fahrzeuge
 - Die Entwicklung autonomer Fahrzeuge, die auf KI-Technologie basieren, hat das Potenzial, Unfälle zu reduzieren und den Verkehr effizienter zu gestalten.
- Sprachverarbeitung und Übersetzung
 - Durch Fortschritte in der Sprachverarbeitung können KI-Systeme Texte und Gesprochenes in verschiedene Sprachen übersetzen, was die Kommunikation über Sprachbarrieren hinweg erleichtert.





KI-Mythen

- Künstliche Intelligenz (KI) hat in den letzten Jahren erhebliche Fortschritte gemacht und spielt eine immer größere Rolle in unserer Gesellschaft. Doch abgesehen von wissenschaftlichen Fakten und technologischem Fortschritt sind KI-Mythen in der Populärkultur tief verwurzelt





Mythos 1: KIs sind bösartig und rebellisch

- Ein weitverbreiteter Mythos in der Populärkultur ist die Darstellung von KIs als bösartig und rebellisch, die sich gegen ihre menschlichen Schöpfer wenden.
- Dieses Motiv findet sich in Filmen wie "Terminator" und "Matrix".
- Obwohl es wahr ist, dass KI-Systeme missbraucht werden können, ist die Idee, dass KIs von Natur aus böse oder rebellisch sind, unrealistisch.
- In der Realität hängt das Verhalten von KIs von ihrem Design und ihrer Programmierung ab.



Mythos 2: KIs sind allwissend und unfehlbar

- Ein weiterer häufiger Mythos ist die Vorstellung, dass KIs allwissend und unfehlbar sind.
- Dieser Mythos wird in Filmen wie "A.I. Künstliche Intelligenz" dargestellt, in denen KIs als übermenschliche Wesen dargestellt werden.
- Tatsächlich sind KIs nicht unfehlbar und können auf falschen Informationen oder unzureichenden Daten trainiert werden.
- Die Leistung einer KI hängt von den Trainingsdaten und dem Algorithmus ab, der sie steuert.



Mythos 3: KIs sind etwas vollkommen Neues

- Künstliche Intelligenz (KI) als Forschungsgebiet existiert seit den 1950er Jahren
- Der Begriff "künstliche Intelligenz" wurde erstmals auf der berühmten Dartmouth-Konferenz von 1956 geprägt, bei der eine Gruppe von Forschern zusammenkam, um über die Möglichkeit zu diskutieren, Maschinen so zu programmieren, dass sie menschliche Intelligenz nachahmen.
- Es ist jedoch wichtig zu beachten, dass die Ideen und Konzepte, die der KI zugrunde liegen, in der Geschichte der Informatik und Mathematik noch weiter zurückreichen
- Alan Turing haben bereits in den 1930er Jahren theoretische Grundlagen für das Verständnis von Berechnung und Intelligenz geschaffen, die später in die Entwicklung der KI eingeflossen sind.
- In den letzten Jahrzehnten hat die KI erhebliche Fortschritte gemacht, insbesondere in Bereichen wie maschinelles Lernen, Deep Learning und neuronale Netze



Mythos 4: KIs sind autonome Entitäten

- Ein häufiger Mythos ist die Vorstellung, dass KIs eigenständige Entitäten sind, die unabhängig von menschlicher Kontrolle agieren.
- Dieser Mythos wird oft in dystopischen Geschichten dargestellt, in denen KIs die Welt beherrschen.
- Tatsächlich sind KIs Werkzeuge, die von Menschen entwickelt, trainiert und gesteuert werden.
- Sie agieren nur im Rahmen ihrer Programmierung und erfordern menschliche Aufsicht.



Mythos 5: KIs können Arbeitnehmer ersetzen

- In der Arbeitswelt gibt es oft die Vorstellung, dass KIs menschliche Arbeitnehmer vollständig ersetzen können.
- Während KI-Systeme in der Lage sind, bestimmte Aufgaben zu automatisieren, können sie nicht alle Aspekte menschlicher Arbeit und Kreativität ersetzen.
- Die Integration von KI in die Arbeitswelt erfordert oft eine Zusammenarbeit zwischen Mensch und Maschine.

Was ist OpenAI?



Sam Altman

1 billion = 10^9

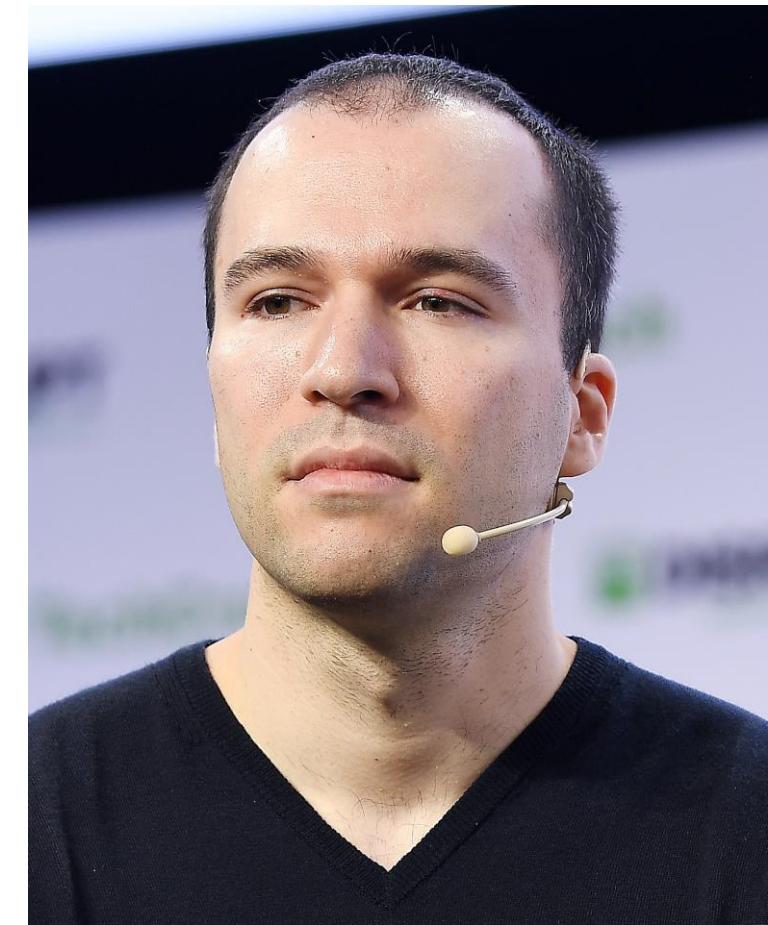
brainymotion

OpenAI Inc.

- 2015 -2018 Non-profit
- 2019 Capped profit
- 2020 GPT 3
- 2022 December ChatGPT
- 2023 Microsoft \$10 billion
- 2024 March GPT-4

Greg Brockman

ML





Meilensteine von OpenAI

- 2016: OpenAI veröffentlicht das erste Paper "Unsupervised Pretraining for Sequence to Sequence Learning".
- 2017: OpenAI veröffentlicht "Deep Reinforcement Learning from Human Feedback", ein bahnbrechendes Paper zur KI-gesteuerten Spieleentwicklung.
- 2018: OpenAI entwickelt die dritte Iteration des Sprachmodells GPT (Generative Pre-trained Transformer).
- 2019: OpenAI startet das Projekt "GPT-2", ein noch leistungsfähigeres Sprachmodell, das für Aufsehen sorgt.
- 2020: OpenAI stellt GPT-3 vor, das bis dato größte und fortschrittlichste Sprachmodell.



Fortschritte in der KI-Forschung

- Die KI-Technologien von OpenAI haben das Potenzial, verschiedene Bereiche der Gesellschaft zu revolutionieren.
- Sie könnten die Produktivität steigern, menschliche Arbeit unterstützen und neue Möglichkeiten für personalisierte Dienste schaffen.
- Gleichzeitig gibt es auch Herausforderungen und ethische Fragen, die im Zusammenhang mit dem Einsatz von KI-Technologien von OpenAI diskutiert werden müssen.

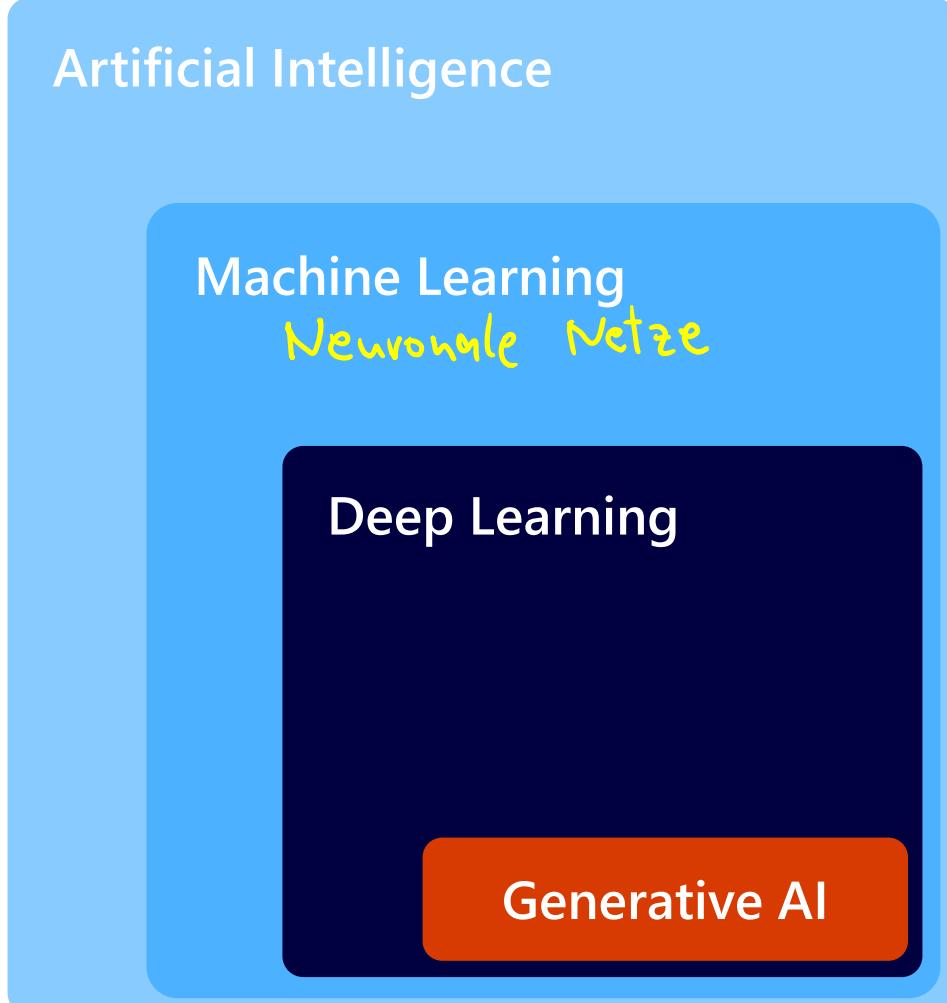


OpenAI's Zukunftsvision

- OpenAI strebt an, KI-Technologien zu entwickeln, die sicher, transparent und allgemein zugänglich sind.
- OpenAI setzt sich dafür ein, dass die Vorteile von KI gerecht auf die gesamte Menschheit verteilt werden.
- Die Zukunftsvision von OpenAI beinhaltet die Zusammenarbeit mit anderen Organisationen und die Förderung von Forschung und Entwicklung im Bereich der KI, um eine positive Zukunft zu gestalten.

Typische Modelle von OpenAI mit Neuronalen Netzen

What is generative AI?



1956

Artificial Intelligence

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

1997

Machine Learning

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

2017

Deep Learning

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

2021

Generative AI

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



Einführung in Neuronale Netze

- Neuronale Netze sind ein Kernkonzept in der Künstlichen Intelligenz.
- Sie bestehen aus künstlichen Neuronen, die miteinander verbunden sind und Informationen verarbeiten.
- Neuronale Netze werden verwendet, um komplexe Aufgaben wie Bilderkennung, Sprachverarbeitung und Entscheidungsfindung zu lösen.

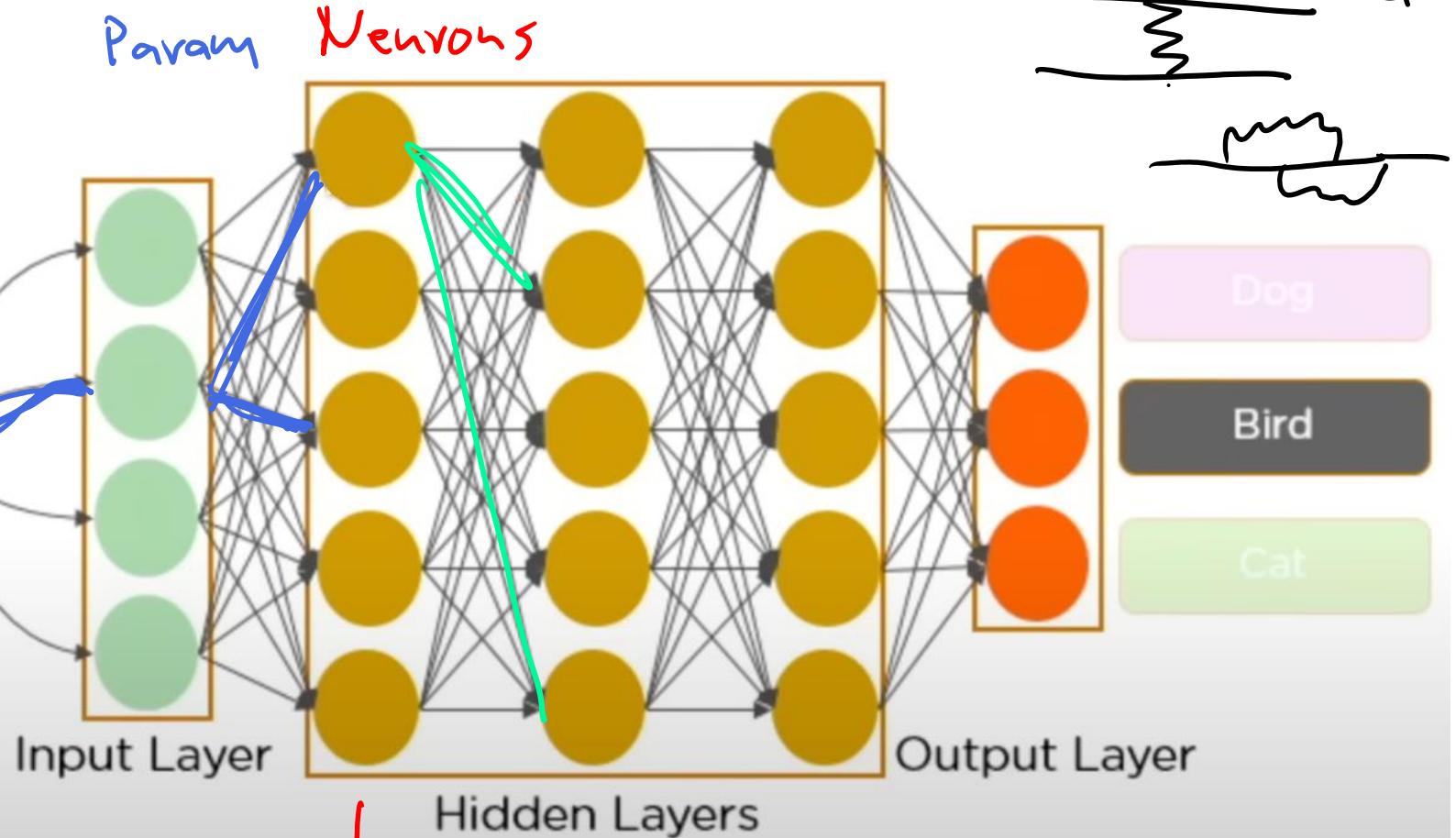
Gelman | Richard Feynman

Neuronale Netzwerke

LLM



Pixels of image fed as input





GPT (Generative Pre-trained Transformer)

- GPT ist eine Modellreihe von OpenAI, die auf dem Transformer-Modell basiert.
- Das Modell wird mit großen Textdatensätzen trainiert und kann darauf basierend Texte generieren.
- GPT-4, die neueste Version, ist besonders bekannt für ihre Fähigkeit, qualitativ hochwertige und kohärente Texte zu erzeugen.

Azure OpenAI model capabilities

- The service include multiple models, optimized for specific tasks
- Models generate responses based on natural language *prompts*

	Language Generation	Code Generation	Image Generation
Prompt:	Write a haiku about marmalade	Write a Python function to add two numbers	Paint a pink fox in a field in the style of Monet
Output:	<i>Orange sunrise, sweet Spread on toast with morning tea A marmalade treat</i>	def add_two_numbers(a, b): return a + b	 A vibrant, colorful illustration of a fox walking through a field of pink and yellow flowers. The fox is depicted in a dynamic pose, moving towards the right. The background is a soft-focus green and yellow, suggesting a natural outdoor setting.

Das Beste an KI ist ihre Fähigkeit ...

The best thing about AI is its ability to ...

The best thing about AI is its ability to ...

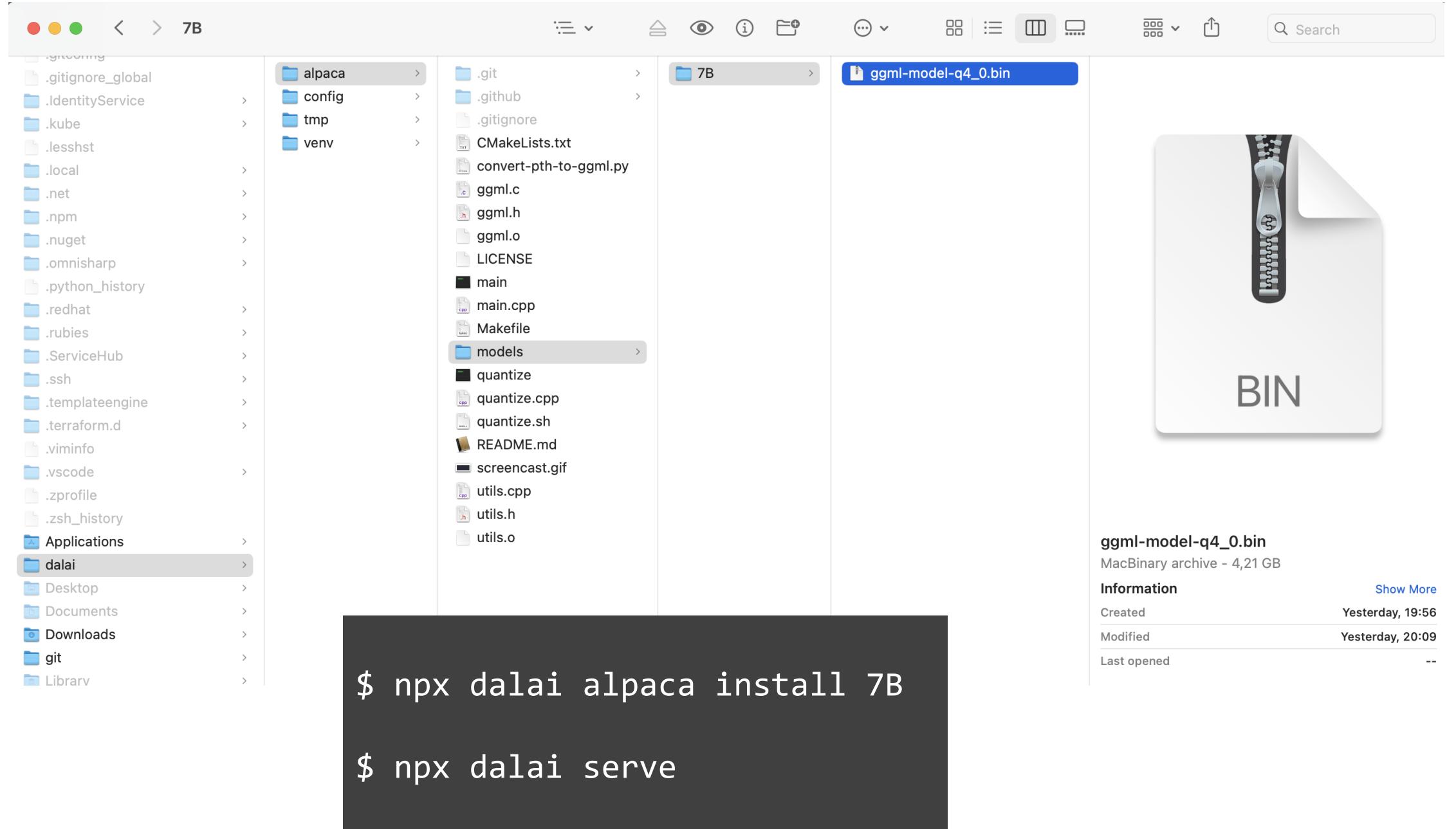


GPT (Generative Pre-trained Transformer)

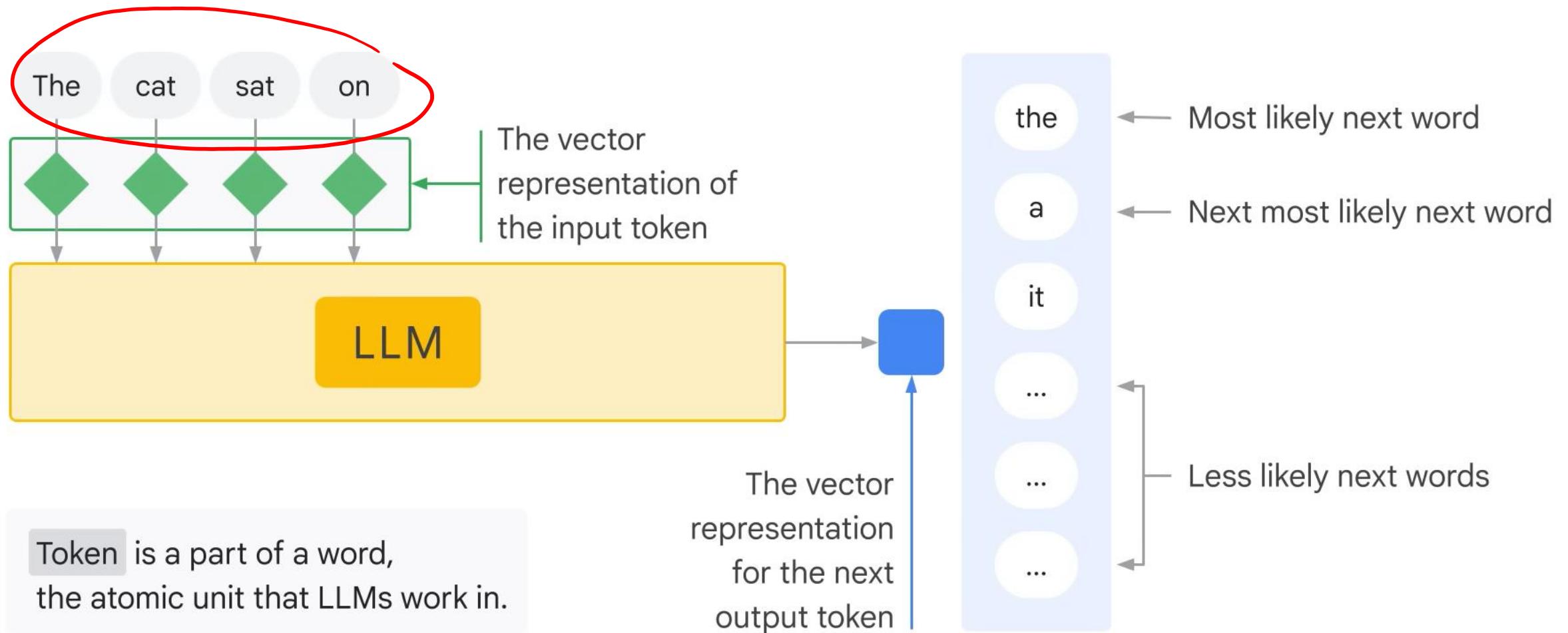
• GPT 2	2019	1.5	$\times 10^9$	Parameters
• GPT 3	2020	17	$\times 10^9$	
• GPT 4	2023	<u>120 ?</u>	$\times 10^9$	



Vergleich: LLaMA 2 Alpaca



Generic language model - A next word predictor...



Chat.openai.com

← ⏪ ⌂ https://platform.openai.com/tokenizer

[Overview](#)[Documentation](#)[API reference](#)[Examples](#)

Tokenizer

The GPT family of models process text using **tokens**, which are common sequences of characters found in text. The models understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.

You can use the tool below to understand how a piece of text would be tokenized by the API, and the total count of tokens in that piece of text.

GPT-3 Codex

```
Zum Kampf der Wagen und Gesänge,  
Der auf Corinthus Landesenge  
Der Griechen Stämme froh vereint,  
Zog Ibycus, der Götterfreund.
```

5

Clear

Show example

Tokens

54

Characters

128

```
Zum Kampf der Wagen und Gesänge,  
Der auf Corinthus Landesenge  
Der Griechen Stämme froh vereint,  
Zog Ibycus, der Götterfreund.
```

5

TEXT

TOKEN IDS

Chat GPT API

en

A helpful rule of thumb is that one token generally corresponds to ~4 characters of text for common English text. This translates to roughly $\frac{3}{4}$ of a word (so 100 tokens \approx 75 words).

If you need a programmatic interface for tokenizing text, check out our [tiktoken](#) package for Python. For JavaScript, the [gpt-3-encoder](#) package for node.js works for most GPT-3 models.

The screenshot shows the OpenPlayground AI interface. At the top, the URL <https://nat.dev> is highlighted with a red circle. The main content area displays a generated text about artificial intelligence, with the word "automate" highlighted in yellow. Below the text, a bar chart shows the distribution of words: "automate = 48.47%", "create = 4.03%", "make = 3.79%", "process = 4.66%", and "quickly = 6.98%". The total logprob is -0.72 on 1 token, with 67.93% probability covered in top 5 logits.

The right side of the interface contains a "Model" configuration panel with the following settings:

- Model: text-davinci-003
- Maximum Length: 256
- Temperature: 0.5 (highlighted with a red circle)
- Top P: 1
- Frequency Penalty: 0
- Presence Penalty: 0
- Stop Sequences: Enter sequence and press Tab
- Show Probabilities: checked
- Highlight Models: checked

A blue arrow points from the "Temperature" slider in the Model panel to the highlighted "Temperature" value in the text output. A red arrow points from the "Temperature" slider in the Model panel to the "Temperature" value in the text output.

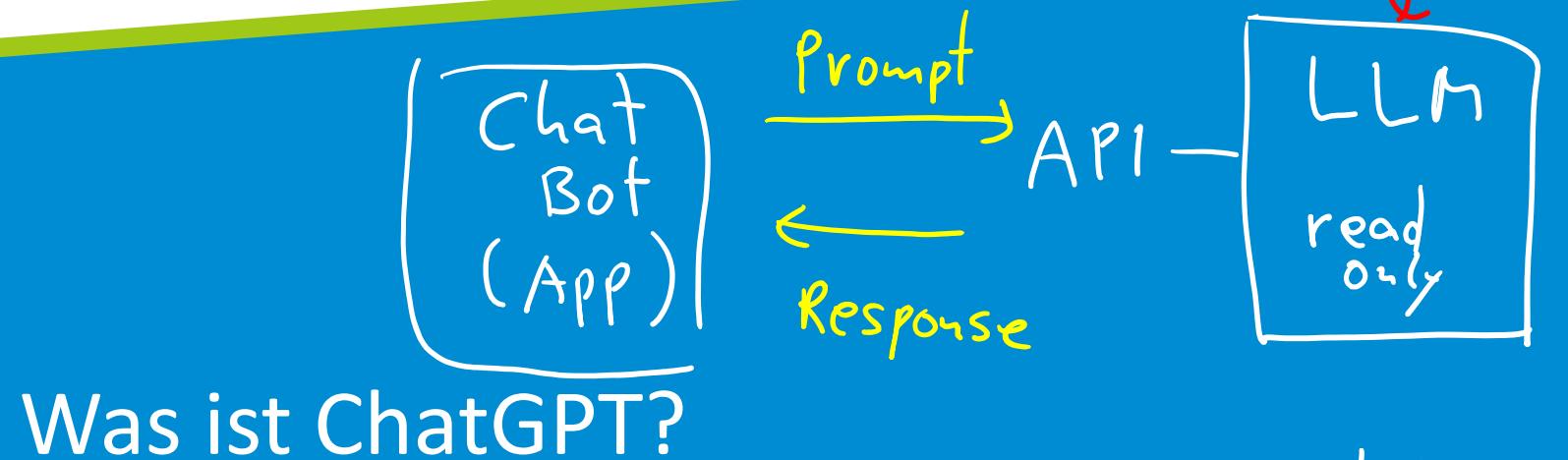
At the bottom left, there are buttons for "Submit", "Undo", and "Redo".



Zusammenfassung

- OpenAI hat eine Vielzahl von Modellen entwickelt, die auf Neuronalen Netzen basieren.
- Diese Modelle können Texte generieren, Bilder erstellen, die Verbindung zwischen Text und Bildern herstellen und sogar in Spielen eingesetzt werden.
- Die Forschung und Entwicklung von OpenAI auf dem Gebiet der Neuronalen Netze hat das Potenzial, viele Bereiche zu beeinflussen und die KI-Technologie weiter voranzutreiben.

Training
+ supervised
Learning
34 Tage 900 x
A100



Was ist ChatGPT?

4 GByte
(LLAMA 2)



Die Geschichte von ChatGPT

- ChatGPT ist ein Sprachmodell, das auf der GPT-3.5-Architektur von OpenAI basiert.
- Das Modell wurde entwickelt, um natürliche und interaktive Gespräche mit Benutzern zu führen.
- ChatGPT basiert auf umfangreichem Texttraining und kann verschiedene Aufgaben und Fragen bearbeiten.



Die Entwicklung von ChatGPT

- Das erste GPT-Modell wurde 2018 von OpenAI veröffentlicht und beeindruckte die Fachwelt durch seine Fähigkeit, qualitativ hochwertigen Text zu generieren.
- Durch kontinuierliches Training und Verbesserungen wurde das Modell immer leistungsfähiger und vielseitiger.
- ChatGPT wurde entwickelt, um eine verbesserte Chatbot-Erfahrung zu bieten und in der Lage zu sein, komplexe Anfragen zu verstehen und zu beantworten.



Einsatzmöglichkeiten von ChatGPT

- ChatGPT findet Anwendung in verschiedenen Bereichen wie Kundensupport, Informationssuche und Unterhaltung.
- Es kann als persönlicher Assistent dienen, der Fragen beantwortet, Ratschläge gibt und Informationen bereitstellt.
- ChatGPT kann auch als kreative Schreibhilfe, zum Erstellen von Texten und zum Generieren von Ideen verwendet werden.



Weiterentwicklungen und Herausforderungen

- OpenAI arbeitet kontinuierlich an Verbesserungen von ChatGPT und der KI-Technologie im Allgemeinen.
- Herausforderungen bestehen darin, das Modell genauer, verständlicher und weniger anfällig für Fehlinformationen zu machen.
- OpenAI bemüht sich auch um den verantwortungsvollen Einsatz von ChatGPT und berücksichtigt ethische Fragen und mögliche Missbrauchsrisiken.

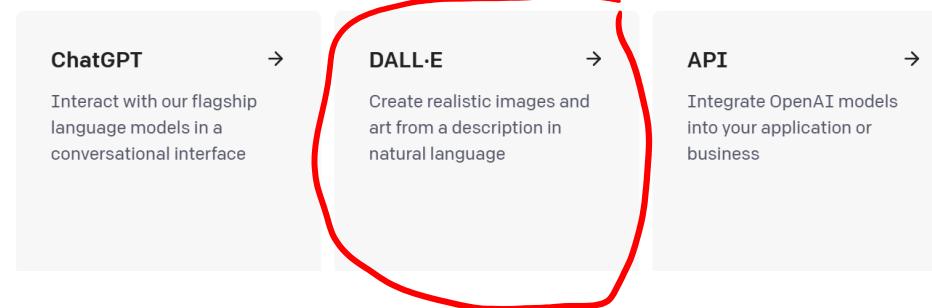


Die Zukunft von ChatGPT

- OpenAI plant, ChatGPT weiterhin zu verbessern und neue Versionen mit noch fortschrittlicheren Fähigkeiten zu veröffentlichen.
- Die Zukunft von ChatGPT könnte mehr Interaktivität, besseres Verständnis von Kontext und spezifischeren Einsatz in verschiedenen Branchen umfassen.
- OpenAI strebt an, die KI-Technologie zugänglicher zu machen und die Vorteile von ChatGPT breit zu verteilen, während gleichzeitig mögliche Herausforderungen angegangen werden.



Drei verschiedene Produkte



Your plan

ChatGPT Plus USD \$20/mo

Your current plan

- Access to GPT-4, our most capable model
- Faster response speed
- Exclusive access to features like Plugins and Advanced Data Analysis

[Manage my subscription](#)

[I need help with a billing issue](#)

This screenshot shows the "Your plan" section of the OpenAI website. It displays the "ChatGPT Plus" plan at USD \$20 per month. The "Your current plan" section lists three key benefits: access to GPT-4, faster response speed, and exclusive features like Plugins and Advanced Data Analysis. Below this, there are links to manage the subscription and handle billing issues.

Buy credits

115 credits

\$15 USD

– 1 +

All credits are shared within the trainymotion organization.

[Cancel](#) [Continue](#)

This screenshot shows a modal window titled "Buy credits". It displays a balance of 115 credits and a price of \$15 USD. There are buttons to decrease or increase the quantity. A note states that all credits are shared within the organization. At the bottom are "Cancel" and "Continue" buttons.

Billing overview

Free trial

Credit remaining ⓘ \$0.00

[Start payment plan](#)

Set up payment plan

Note: This card will be saved to your account and can be removed at any time.

Card information

Name on card

Billing address

Country

Address line 1

Address line 2

City

Postal code

State, county, province, or region

[Cancel](#) [Continue](#)

This screenshot shows the "Billing overview" section and a partially open "Set up payment plan" modal. It indicates a free trial with \$0.00 credit remaining. The payment plan modal asks for credit card details, including card number, name on card, billing address, city, state, and postal code. It also includes a note about saving the card to the account.

Die Oberfläche

- URL: <https://chat.openai.com/>
- Kostenlosen Account erstellen
- Jederzeit möglich, auf kostenpflichtig zu wechseln



Create your account

Note that phone verification may be required for signup. Your number will only be used to verify your identity for security purposes.

Email address

Continue

Already have an account? [Log in](#)

OR

Continue with Google

Continue with Microsoft Account

Continue with Apple



ChatGPT →

Interact with our flagship language models in a conversational interface

DALL·E →

Create realistic images and art from a description in natural language

API →

Integrate OpenAI models into your application or business

Die Oberfläche

The screenshot shows the ChatGPT interface. On the left, there's a sidebar with a dark background containing a list of recent conversations and various settings. The main area shows a conversation between a user (RG) and the AI. RG asks "was ist openai?", and the AI responds with a detailed explanation of what OpenAI is, its history, and its mission. The AI also mentions its developments in AI technology and its goal to democratize AI. At the bottom, there's a footer with a link to the ChatGPT terms of service.

+ New chat

OpenAI for Artificial Intelligence

Neuronale Netze erklärt

OpenAI and Bing Partnership

OpenAI - KI-Forschungsunternehmen

Clear conversations

Upgrade to Plus NEW

Dark mode

Updates & FAQ

Log out

RG was ist openai?

OpenAI ist ein forschungsorientiertes Unternehmen im Bereich der künstlichen Intelligenz (KI), das im Jahr 2015 gegründet wurde. Es hat seinen Hauptsitz in San Francisco, Kalifornien. Die Organisation wurde von einer Gruppe von bekannten Persönlichkeiten im Bereich der Technologie und KI gegründet, darunter Elon Musk, Sam Altman, Greg Brockman und andere.

OpenAI zielt darauf ab, die Forschung im Bereich der KI zu fördern und neue Technologien zu entwickeln, die das Potenzial haben, die Welt auf transformative Weise zu verändern. Das Unternehmen konzentriert sich auf eine breite Palette von Anwendungen, darunter die Entwicklung von KI-Systemen, die in der Lage sind, natürliche Sprache zu verstehen und zu erzeugen, sowie auf KI-Systeme, die in der Lage sind, komplexe Probleme in Bereichen wie Robotik, Automatisierung, Gesundheitswesen und Bildung zu lösen.

OpenAI hat eine Reihe von bedeutenden Entwicklungen auf dem Gebiet der KI erreicht, darunter die Entwicklung von Sprachmodellen wie GPT-3, die eine große Anzahl von natürlichen Sprachanwendungen ermöglichen. Es hat auch an der Entwicklung von Deep Learning-Algorithmen und anderen Technologien gearbeitet, die bei der Lösung von Problemen in verschiedenen Branchen und Sektoren eingesetzt werden können.

Das Ziel von OpenAI ist es, die KI-Forschung und -Entwicklung zu demokratisieren und der Welt Zugang zu leistungsstarken, menschenähnlichen KI-Systemen zu ermöglichen, die das Potenzial haben, das Leben der Menschen auf der ganzen Welt zu verbessern.

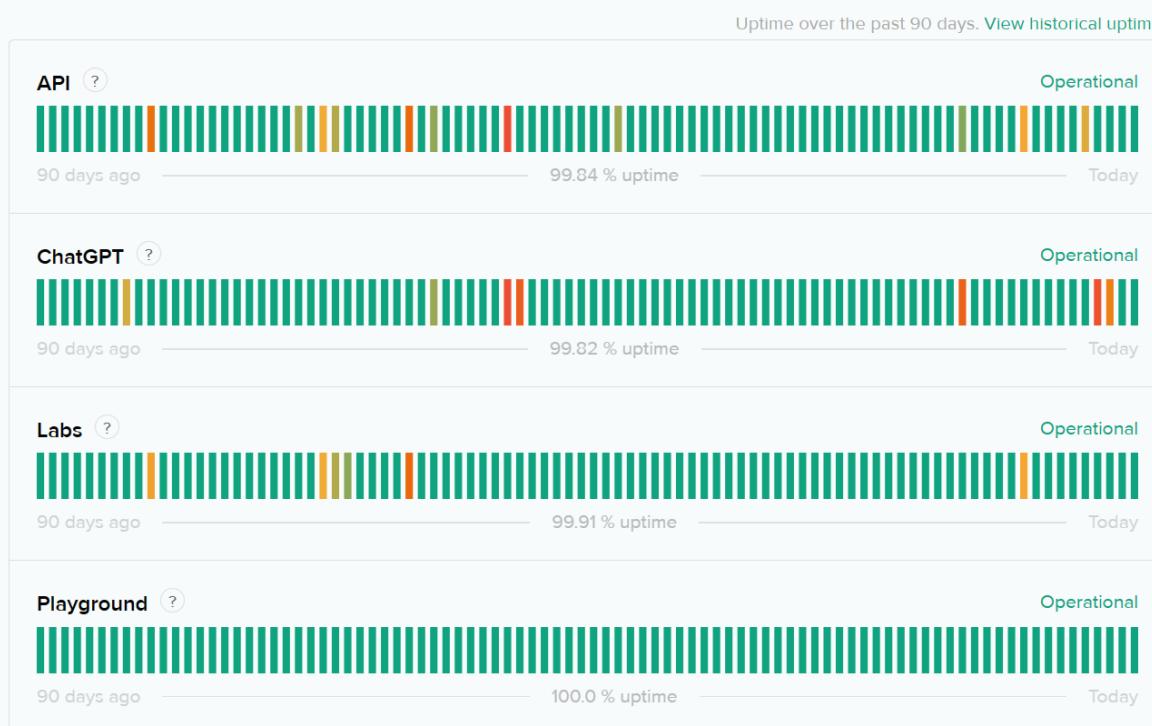
Regenerate response

ChatGPT Mar 14 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.



SUBSCRIBE TO UPDATES

All Systems Operational



Elevated error rate for ChatGPT

Incident Report for OpenAI

Resolved

This incident has been resolved.

Posted 2 days ago. Jul 12, 2023 - 09:20 PDT

Monitoring

A fix has been implemented and we are monitoring the results.

Posted 2 days ago. Jul 12, 2023 - 09:10 PDT

Update

Some users are seeing a "Failed to get service status" error.

We are investigating

Posted 2 days ago. Jul 12, 2023 - 08:37 PDT

Update

We are continuing to work on the problem and things are improving. Logins are starting to work.

Posted 2 days ago. Jul 12, 2023 - 08:14 PDT

Investigating

ChatGPT is currently unavailable for most users, and may be slow to load. We've disabled login while we recover the service.

Posted 2 days ago. Jul 12, 2023 - 07:43 PDT

This incident affected: ChatGPT.



Beispiel I

- Schreiben erstellen lassen
 - Einladung schreiben





Beispiel II

- Ideen sammeln
 - Welche Möglichkeiten gibt es im Unternehmen Energie zu sparen?





Beispiel III

- Marktüberblick
 - Welche Tools zur Erstellung von Dashboards sind in Unternehmen in Deutschland verbreitet?





Beispiel IV

- Marketing
 - Schreibe einen Text nach dem AIDA Prinzip in 3 Varianten um

*Wir befähigen Menschen überall auf der Welt, durch Training zu Erfolg zu kommen.
Das Unternehmen brainymotion versteht sich als Dienstleistungsanbieter von praxisnahen und innovativen Aus- und Weiterbildungen mit der Möglichkeit zum Erwerb von Voll- und Teil-Qualifizierungen auf quartärer Bildungsebene. Die Kernkompetenzen liegen dabei in den Bereichen IT (technisch und Endanwender) sowie Soft- und Business Skills. Um diesem Anspruch gerecht zu werden, sollen die folgenden Leitsätze allen Mitarbeitern, einschließlich der freiberuflichen Dozenten als Orientierung für die Erfüllung ihrer Aufgaben und als Grundlage für das tägliche Handeln dienen.*



Beispiel IV

- Werbetexte für Blogeinträge
 - Schreibe einen kreativen Werbetext für einen Blogeintrag
„Microsoft Office Schulungen“





Beispiel VI

- Anleitungen mit Beispielen erstellen lassen
 - Wie erstelle ich in Excel eine sVerweis?
 - Erstelle mir ein Makro um in Excel-Tabellen leere Zellen zu löschen
 - Use Cases





Konklusion

- Welche Vorteile bietet ChatGPT für Unternehmen?
 - Produktivitätssteigerung
 - Mitarbeiterentlastung
 - Effizientere Kommunikation
 - Zeitersparnis
 - Verbesserung der Arbeitsabläufe
 - Verbesserung der Datenanalyse
 - Kosteneinsparungen



Übung: PowerShell Wrapper für OpenAI Models

siehe github.com/www42/openAI

```
$Key = 'sk-xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx'

$headers = @{
    "Authorization" = "Bearer $Key"
    "Content-Type" = "application/json"
}

$body = ConvertTo-Json @{
    "model" = 'text-davinci-003'
    "prompt" = "This is a test"
}

Invoke-WebRequest ` 
    -Uri "https://api.openai.com/v1/completions" ` 
    -Method POST ` 
    -Headers $headers ` 
    -Body $body
```

Tag 2 Guten Morgen!

OpenAI und ChatGPT:
Neuer Support in der IT!

OpenAI API

How does an Azure API work?

1

The client is any software application that runs on your phone, computer, or other smart devices. When you use the client, it sends a request.



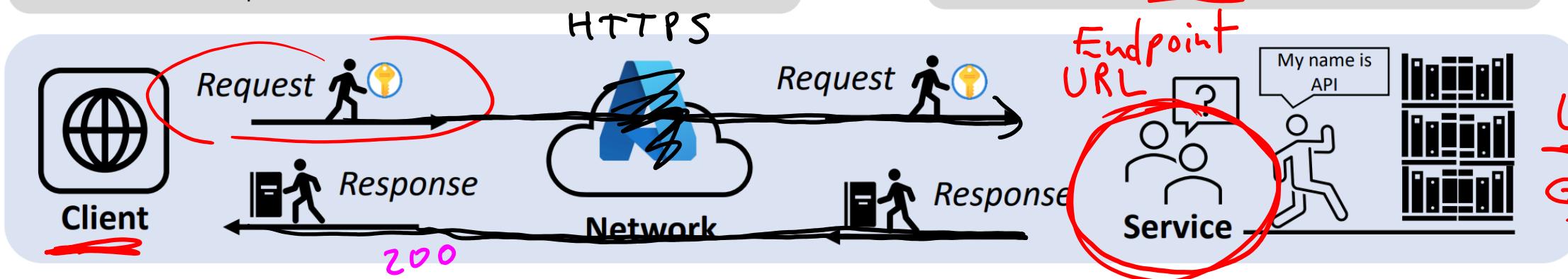
Imagine a visitor who walks to a librarian's desk holding a key. The key is like a pass at the library, allowing the visitor to borrow a book. The visitor wants to make a request such as "I want Book A" or "I want a prediction of the weather".

2

Requests are sent through a network to an endpoint. The Application Programming Interface (API) at that endpoint performs a service to fulfill the request and send back a response.



The endpoint is like the library information desk. An API is the librarian who gets a request "I want Book A", goes to get Book A.



3

An AI-informed API can fulfill requests such as predictions and send back a response. The client receives the response.



Azure AI services consist of a trained machine learning models that behave like a function – accepting one or more input values and generating a predicted output based on probability.



An AI-informed librarian could get a request "I want a prediction of the weather" and goes to get a prediction that the weather will be sunny. The visitor takes that response back to the client.

Verb

List models

GET https://api.openai.com/v1/models

Lists the currently available models, and provides basic information about each one such as the owner and availability.

Endpoint

Create chat completion

POST https://api.openai.com/v1/chat/completions

Creates a model response for the given chat conversation.

Create image

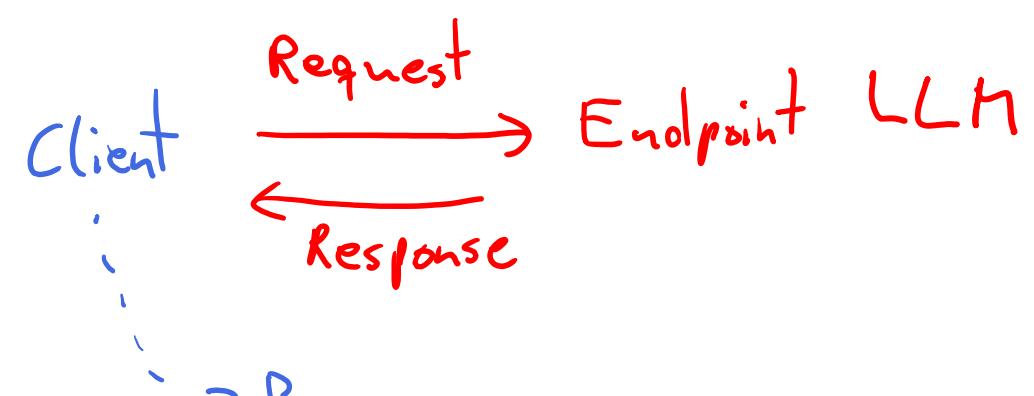
POST https://api.openai.com/v1/images/generations

Creates an image given a prompt.

Create embeddings

POST https://api.openai.com/v1/embeddings

Creates an embedding vector representing the input text.



z.B.

Curl

Invoke - Rest Method (Power
shell)
(jeste Sprache)

Postman Tool

Models

Overview

The OpenAI API is powered by a diverse set of models with different capabilities and price points. You can also make customizations to our models for your specific use case with [fine-tuning](#).

MODEL	DESCRIPTION
GPT-4 and GPT-4 Turbo	A set of models that improve on GPT-3.5 and can understand as well as generate natural language or code
GPT-3.5 Turbo	A set of models that improve on GPT-3.5 and can understand as well as generate natural language or code
DALL-E	A model that can generate and edit images given a natural language prompt
TTS	A set of models that can convert text into natural sounding spoken audio
Whisper	A model that can convert audio into text
Embeddings	A set of models that can convert text into a numerical form
Moderation	A fine-tuned model that can detect whether text may be sensitive or unsafe
GPT base	A set of models without instruction following that can understand as well as generate natural language or code
Deprecated	A full list of models that have been deprecated along with the suggested replacement

GPT-3.5 Turbo

GPT-3.5 Turbo models can understand and generate natural language or code and have been optimized for chat using the [Chat Completions API](#) but work well for non-chat tasks as well.

MODEL	DESCRIPTION	CONTEXT WINDOW	TRAINING DATA
gpt-3.5-turbo-0125	<p>New Updated GPT 3.5 Turbo</p> <p>The latest GPT-3.5 Turbo model with higher accuracy at responding in requested formats and a fix for a bug which caused a text encoding issue for non-English language function calls. Returns a maximum of 4,096 output tokens. Learn more.</p>	16,385 tokens	Up to Sep 2021
gpt-3.5-turbo	Currently points to gpt-3.5-turbo-0125.	16,385 tokens	Up to Sep 2021
gpt-3.5-turbo-1106	GPT-3.5 Turbo model with improved instruction following, JSON mode, reproducible outputs, parallel function calling, and more. Returns a maximum of 4,096 output tokens. Learn more .	16,385 tokens	Up to Sep 2021
gpt-3.5-turbo-instruct	Similar capabilities as GPT-3 era models. Compatible with legacy Completions endpoint and not Chat Completions.	4,096 tokens	Up to Sep 2021



Wie kann ChatGPT die moderne IT unterstützen?



Einführung

- ChatGPT, entwickelt von OpenAI, ist ein fortschrittliches Sprachmodell, das in der Lage ist, natürliche und interaktive Gespräche zu führen.
- Durch seine Fähigkeit, menschenähnliche Texte zu generieren, kann ChatGPT die moderne IT in verschiedenen Bereichen unterstützen.



Kundensupport und Helpdesk

- ChatGPT kann als virtueller Assistent eingesetzt werden, um Kunden bei technischen Fragen und Problemen zu unterstützen.
- Durch die Bereitstellung von genauen und relevanten Informationen kann ChatGPT die Effizienz des Kundensupports verbessern.



Fehlerbehebung und Diagnose

- Bei der Fehlerbehebung und Diagnose von IT-Problemen kann ChatGPT als Wissensbasis dienen.
- Es kann Fragen zu gängigen Problemen beantworten, Lösungen vorschlagen und den Benutzern bei der Behebung von Schwierigkeiten helfen.



Code-Snippets und Programmierung

- Entwickler können ChatGPT verwenden, um nach Code-Snippets, Beispielen oder Programmierhilfe zu fragen.
- Es kann dabei helfen, schnelle Lösungen und Syntaxvorschläge zu erhalten, um die Programmierarbeit effizienter zu gestalten.



Wissensmanagement

- ChatGPT kann als leistungsfähiges Werkzeug für das Wissensmanagement in IT-Unternehmen dienen.
- Es kann verwendet werden, um firmeninternes Wissen zu konsolidieren, FAQs zu erstellen und den Zugriff auf wichtige Informationen zu erleichtern.



Datenauswertung und -analyse

- ChatGPT kann bei der Datenauswertung und -analyse unterstützen.
- Es kann Fragen zu Dateninterpretation, Visualisierungsmethoden und statistischen Analysen beantworten und wertvolle Einblicke liefern.



IT-Sicherheit und Datenschutz

- ChatGPT kann Unternehmen dabei unterstützen, Sicherheitslücken zu erkennen und bewährte Verfahren im Bereich IT-Sicherheit und Datenschutz zu implementieren.
- Es kann Schulungen und Richtlinien für Mitarbeiter bereitstellen und bei der Sensibilisierung für Sicherheitsrisiken helfen.



Automatisierung und Prozessoptimierung

- Durch den Einsatz von ChatGPT können Unternehmen Prozesse automatisieren und optimieren.
- Es kann wiederkehrende Aufgaben übernehmen, wie z. B. das Generieren von Berichten, das Verfolgen von Projekten oder das Erstellen von Dokumentationen.



Forschung und Entwicklung

- In der IT-Forschung und -Entwicklung kann ChatGPT als kreativer Assistent dienen.
- Es kann Ideen für neue Technologien, Algorithmen oder Lösungsansätze liefern und bei der Entwicklung innovativer IT-Konzepte unterstützen.

Zusammenfassung

- ChatGPT bietet vielfältige Möglichkeiten, die moderne IT zu unterstützen.
- Von Kundensupport bis hin zur Prozessautomatisierung kann es Unternehmen helfen, effizienter zu arbeiten und bessere IT-Dienstleistungen anzubieten.

OpenAI und Microsoft



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



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

GPT-3.5 and GPT-4

Text

ChatGPT

Conversation

Codex

Code

DALL·E 2

Images

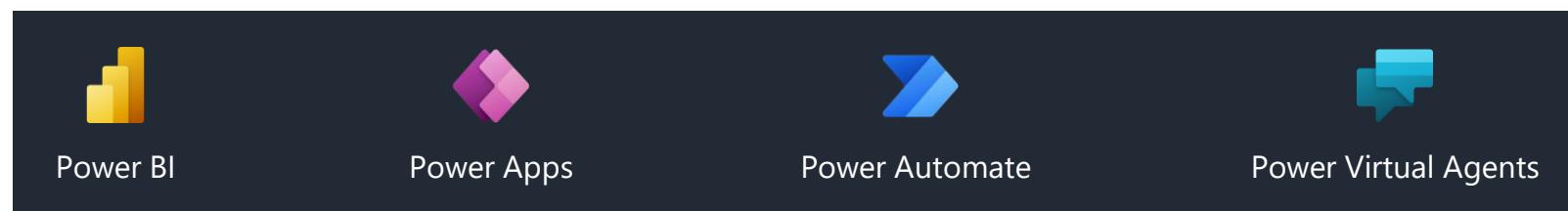
What is Azure OpenAI Service?

Applications



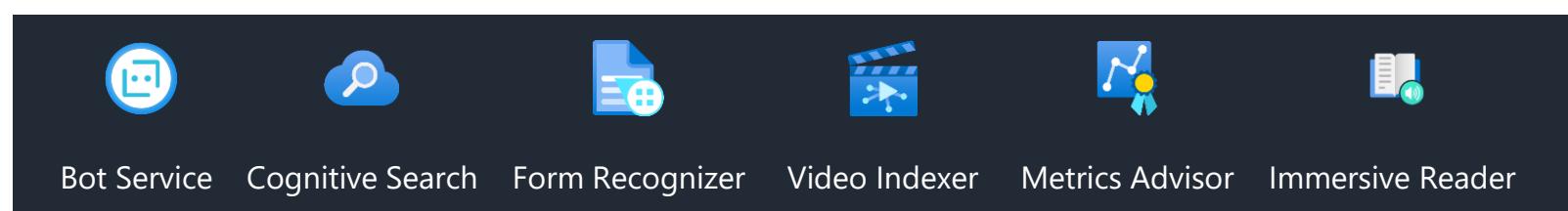
Application Platform

AI Builder



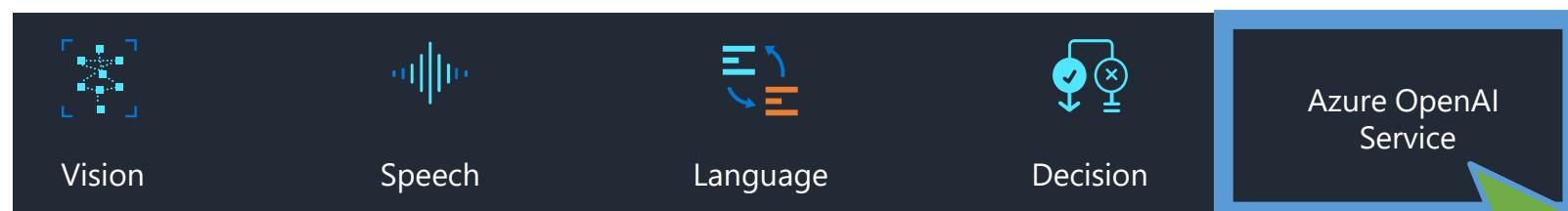
Scenario-Based Services

Applied AI Services



Customizable AI Models

Cognitive Services



ML Platform



Cutting-edge generative AI models from OpenAI
+
Scalability, interoperability, and data protection of Azure



Azure OpenAI model capabilities

- The service include multiple models, optimized for specific tasks
- Models generate responses based on natural language *prompts*

	Language Generation	Code Generation	Image Generation
Prompt:	Write a haiku about marmalade	Write a Python function to add two numbers	Paint a pink fox in a field in the style of Monet
Output:	<i>Orange sunrise, sweet Spread on toast with morning tea A marmalade treat</i>	def add_two_numbers(a, b): return a + b	 A vibrant, colorful illustration of a fox walking through a field of pink and yellow flowers. The fox is pinkish-purple with a bushy tail, and the background is a soft-focus green and yellow.

Azure OpenAI Service Capabilities

Creative Ideation	Subject Research
Essay outlines	Extracting insights
Summarizing text	Code generation
Virtual Assistants	Language Translation
Comments from code	Classifying text
	Answering questions
	Poem creation
	Image Generation
	Code refactoring

Using prompts to get completions from models

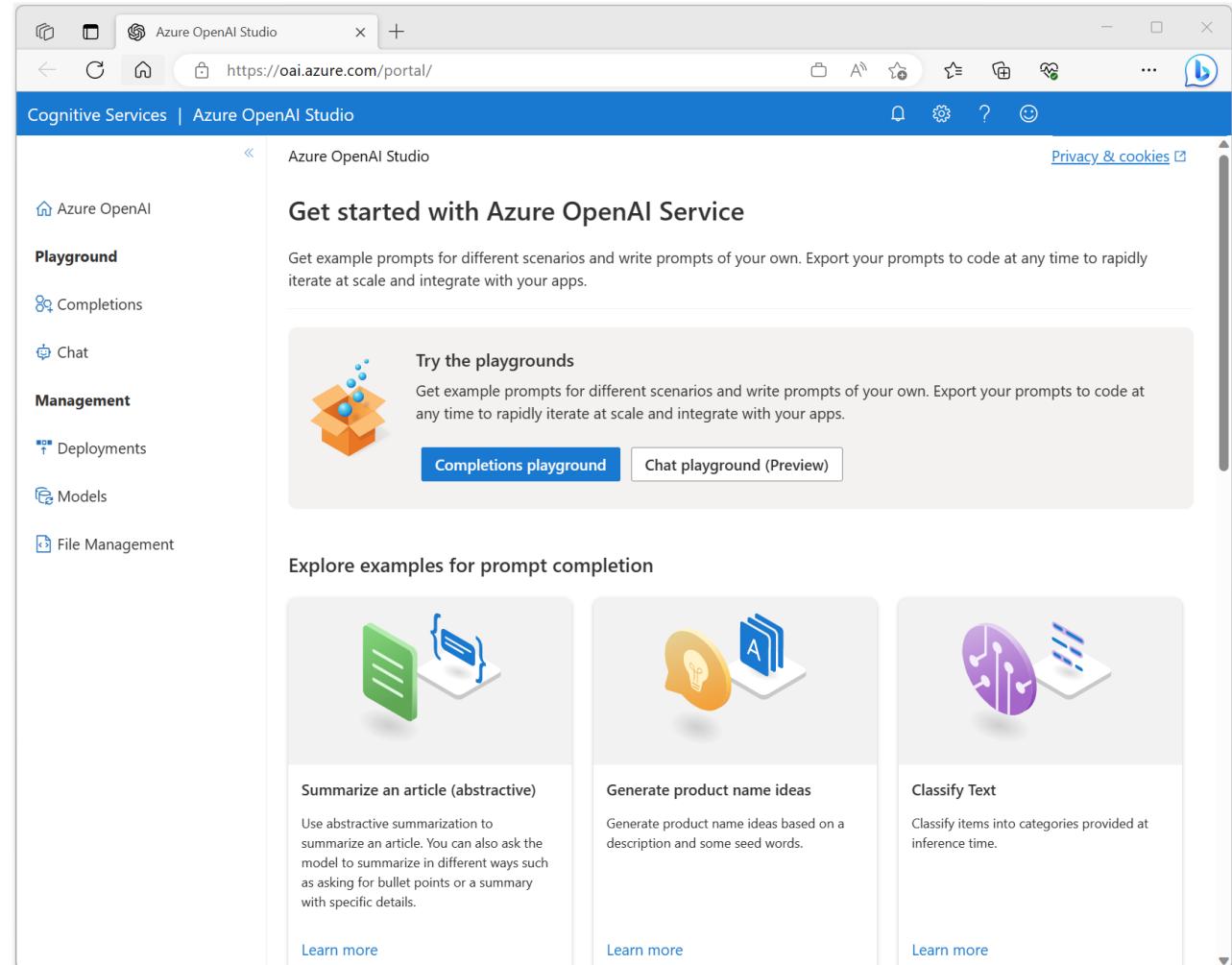
Task	Prompt	Completion
Classifying content	Tweet: I enjoyed the training course. Sentiment:	Positive
Generating new content	Write a poem about databases	Databases, oh databases, You keep our information safe, From the small to the large, You store our data in a place.
Transformation/Translation	English: Hello French:	Bonjour
Summarization	Scotland is [long description of Scotland...] Summarize the previous text	Scotland is [summarized description...]
Continuation	One way to grow tomatoes is to	start with seeds...
Question answering	How many moons does Earth have?	Earth has one moon.
Chat	<i>Setup, followed by messages...</i>	<i>A sequence of relevant responses</i>

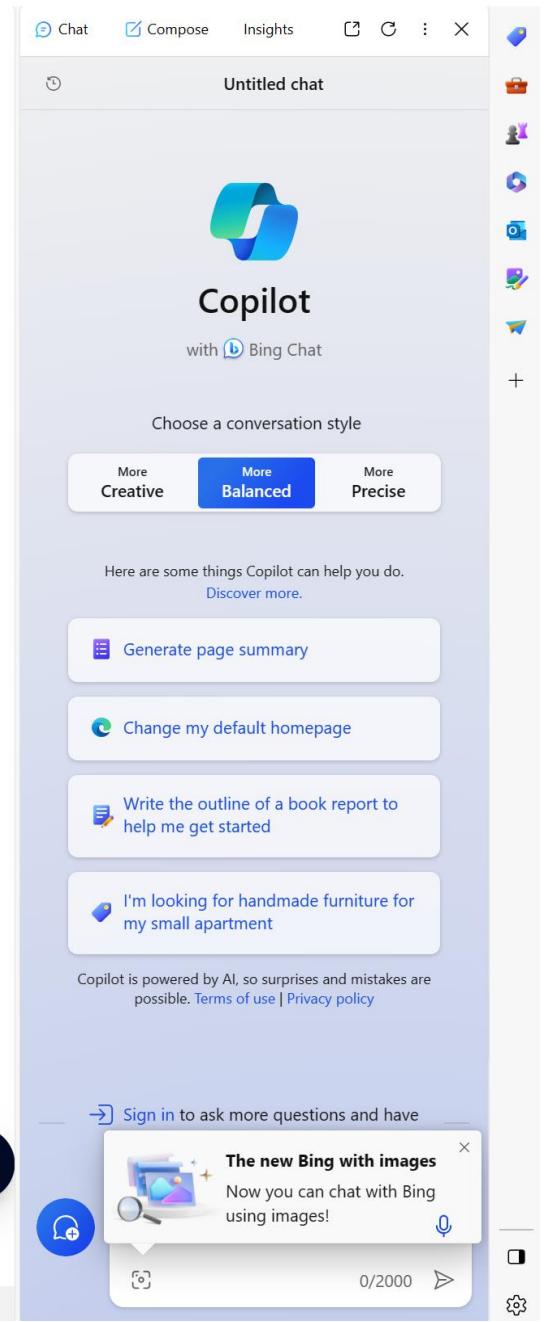
| Azure OpenAI | GPT-3 Family of Models

Model	Request	Description, performance, cost	Use cases
Davinci	4,000 tokens	Most capable GPT-3 model. Can do any task the other models can do, often with <i>higher quality, longer output</i> and <i>better instruction-following</i> .	Complex intent, cause and effect, summarization for audience
Curie	2048 tokens	Very capable , but <i>faster</i> and <i>lower cost</i> than Davinci.	Language translation, complex classification, text sentiment, summarization
Babbage	2048 tokens	Capable of straightforward tasks, <i>very fast</i> , and <i>lower cost</i> .	Moderate classification, semantic search classification
Ada	2048 tokens	Capable of very simple tasks, usually the <i>fastest</i> model in the GPT-3 series, and <u><i>lowest cost</i></u> .	Parsing text, simple classification, address correction, keywords

Azure OpenAI Studio

- Web portal for working with Azure OpenAI models:
<https://oai.azure.com/>
- View and deploy base models
- Manage fine tuning and data files for custom models
- Test models in visual playgrounds:
 - **Completions** (GPT-3 and earlier models)
 - **Chat** (GPT-3.5-Turbo and later models)





Entfesseln Sie Ihre Produktivität und Kreativität mit  Copilot

Microsoft Copilot in Edge, Ihr KI-Begleiter für das Web, hilft Ihnen, mehr zu erstellen, zu finden und zu tun, als Sie für möglich gehalten haben.

Schaffen Antworten erhalten Planung

Erstellen Sie eine Collage aus Himmelskörpern in Kombination mit Musiknoten



Probieren Sie es in Copilot aus

Erfahren Sie, wie Sie das Beste aus Copilot herausholen können:

Zusammenfassen und Suchen

Erstellen von Bildern

[Verwalten Sie Ihren Browser](#)

Schreiben wie ein Profi

 Klicken Sie auf das ⓘ, um Ihre Stimme zu verwenden!

 Willkommen bei Microsoft Edge

Microsoft Copilot in Edge

Weitere neue Funktionen

Die Inhalte auf dieser Seite wurden möglicherweise maschinell übersetzt.

 Ihre Datenschutzoptionen Datenschutz Cookies verwalten © Microsoft 2023



Vector Search

**"LOREM IPSUM
DOLOR SIT AMET,
CONSECTETUR..."**

DATA

**embedding
model**

-0.001
0.006
-0.014
...
-248.6

vector



Retrieval Augmented Generation (RAG)



ML Applications | Open Source

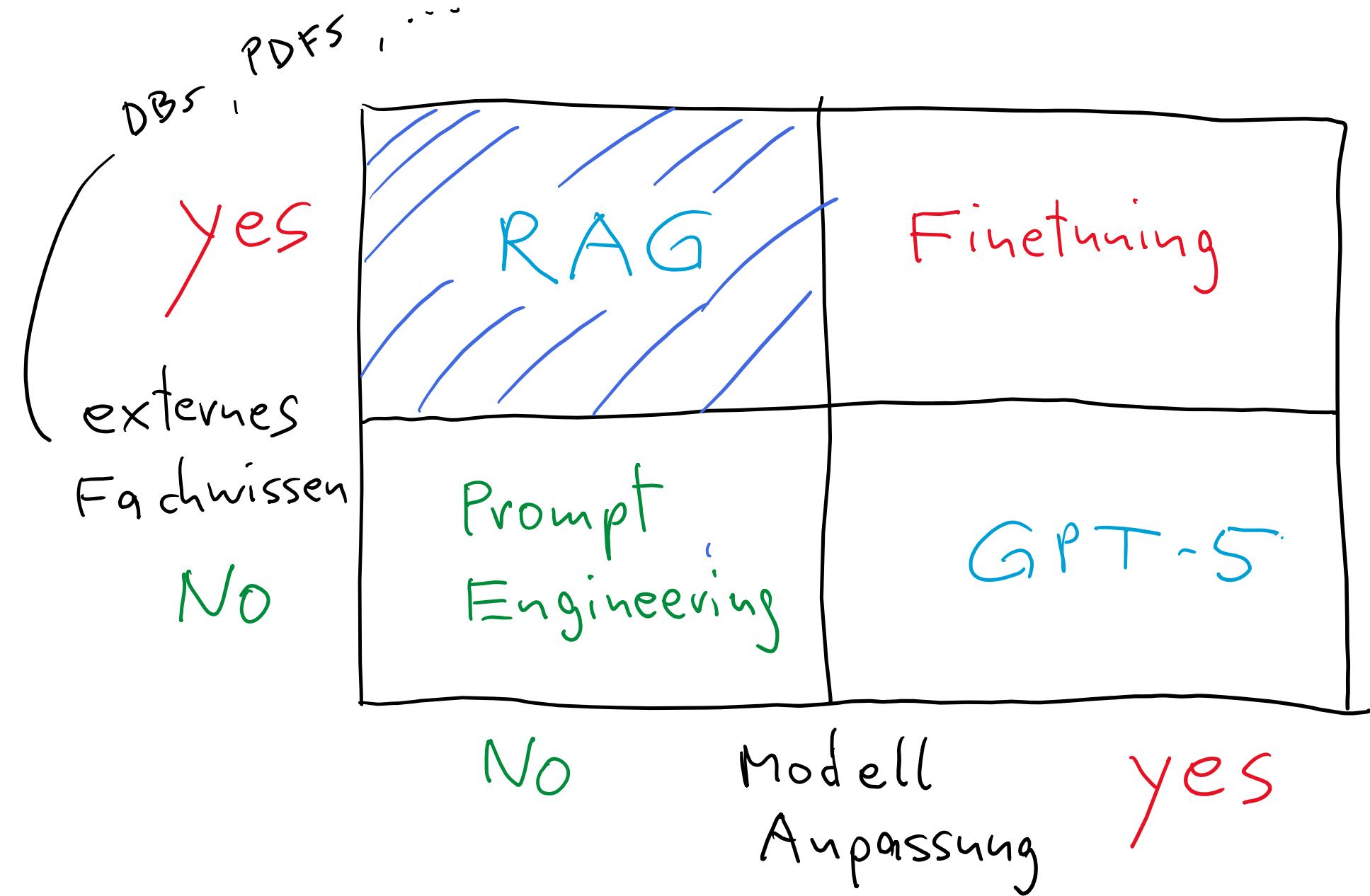
Retrieval Augmented Generation: Streamlining the creation of intelligent natural language processing models

September 28, 2020

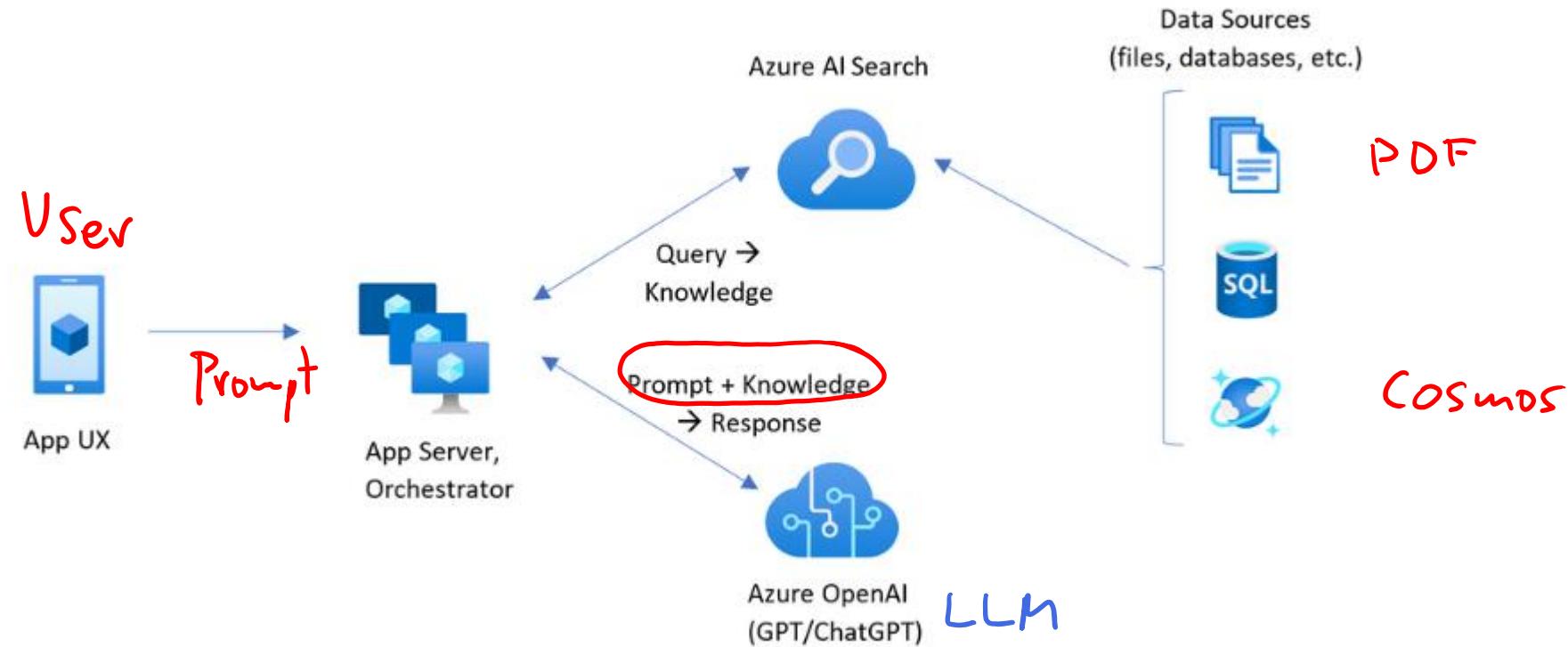


Teaching computers to understand how humans write and speak, known as natural language processing (NLP), is one of the oldest challenges in AI research. There has been a marked change in approach over the past two years, however. Where research once focused on developing specific frameworks for specific tasks, today powerful general-purpose language models can be fine-tuned for a wide variety of different tasks.

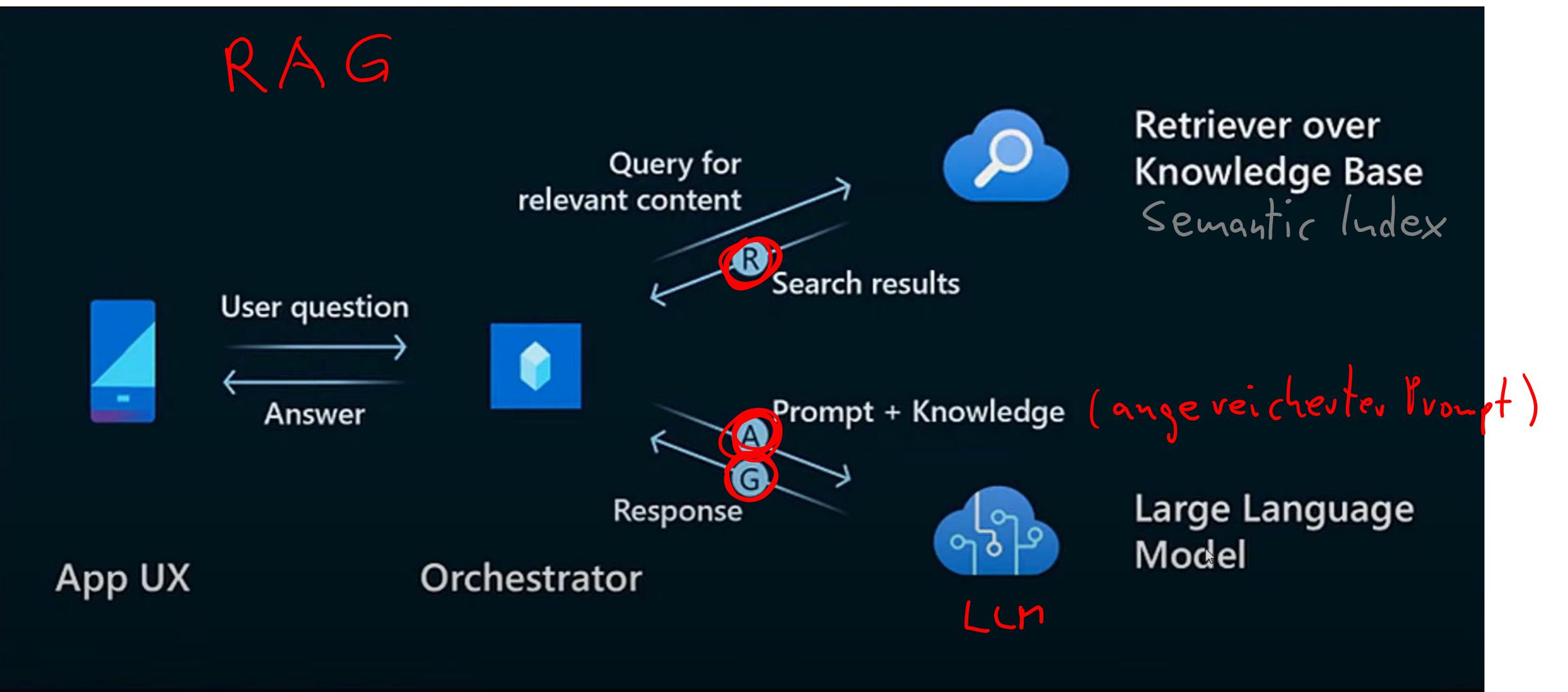
[eneration-of-intelligent-natural-language-processing-models/](#)



How does RAG work?



RAG



The AI-Powered Organization

The world is changing

Industries are transforming rapidly

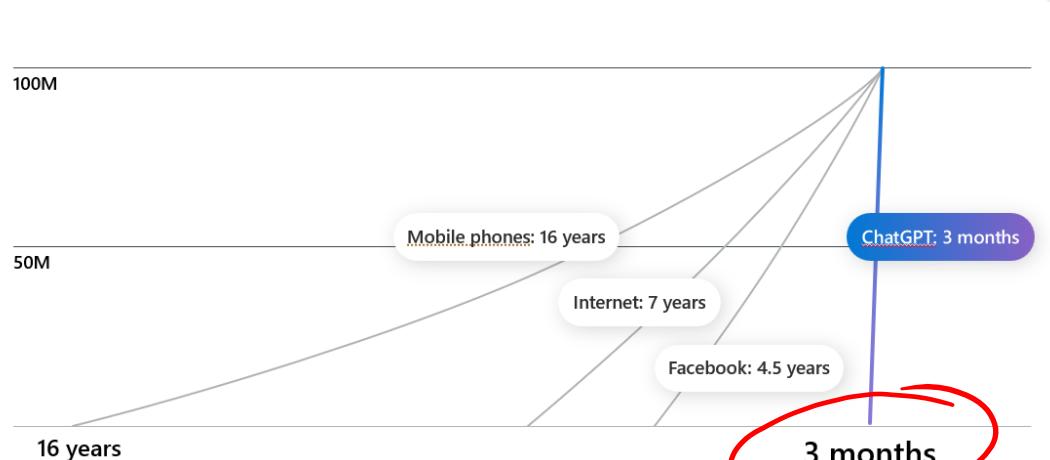
Drivers for economic growth are evolving

Technology is at an inflection point

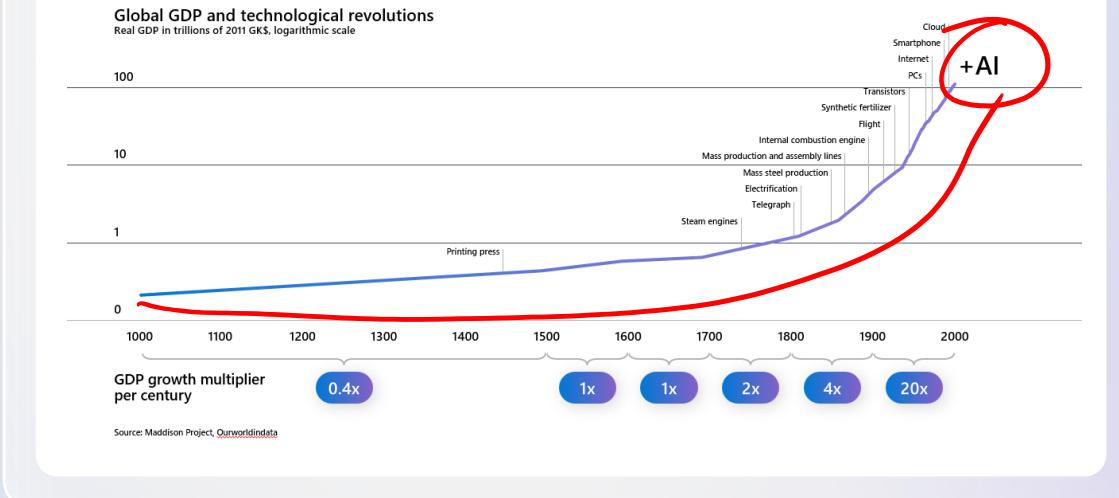
AI technology and impact

The AI technology is here

Time to reach 100M users



And the impact is **real**



Key Microsoft AI breakthroughs

- 2016 • Object and speech recognition
Human parity
- 2018 • Reading comprehension and machine translation
Human parity
- General language understanding
Human parity
- 2020 • Turing-NLG language models
- First AI supercomputer
- Exclusive license for OpenAI GPT-3 models
- Image captioning
Human parity
- 2021 • Natural Language Understanding
Human parity
- Commonsense Question Answering
Human parity
- Azure OpenAI Service preview
- 2022 • GitHub copilot general availability

\$ 10^{10}



"We formed our partnership with OpenAI around a shared ambition to responsibly advance cutting-edge AI research and democratize AI as a new technology platform," said Satya Nadella, Chairman and CEO, Microsoft. "In this next phase of our partnership, developers and organizations across industries will have access to the best AI infrastructure, models, and toolchain with Azure to build and run their applications."



Supercomputing
at scale



New AI
experiences



Provision of cloud
services

Generative AI capabilities

Google : Gemini
video

Extracting
insights

Creative
ideation

Subject
research

Dialog
agents

Essay
outlines

Summarizing
text

Virtual
assistants

Classifying
text

Code
generation

Answering
questions

Image
generation

Language
translation

Poem
creation

Code
refactoring

Writing
assistance

Comments
from code

Semantic
search

Responsible AI



Microsoft 365 Copilot

28€ / User / Month

Microsoft 365 Copilot



Grounded in **your business data**

Microsoft 365 Copilot has real-time access to both your content and context in Microsoft Graph.



Comprehensive **security, compliance, and privacy**

Copilot inherits your security, compliance, and privacy policies set up in Microsoft 365.



Architected to protect data

Your data never leaves its secure partition, and it is never used for training purposes.



Integrated into the **apps you use every day**

Word, Excel, PowerPoint, Outlook, Teams, and more.



Individual user and admin **always in control**

User decides what to use, modify, or discard.



Designed to **learn new skills**

As Copilot learns about processes, it can perform more sophisticated tasks and queries.



Microsoft 365 Copilot

Natural Language



Large Language
Models

LLM

+



Microsoft Graph
- Your Data -

Externes
Fachwissen



Microsoft 365
Apps

Powerpoint
Teams
Excel
...



Microsoft 365 Copilot

Embedded across Microsoft 365 apps



Open Source
LangChain
Pinecone

Embedding Vector



Microsoft 365 Copilot

Graph API

Email
Postfach
Semantic Index

Vector DB



Enable fast semantic search
across **billions of vectors**



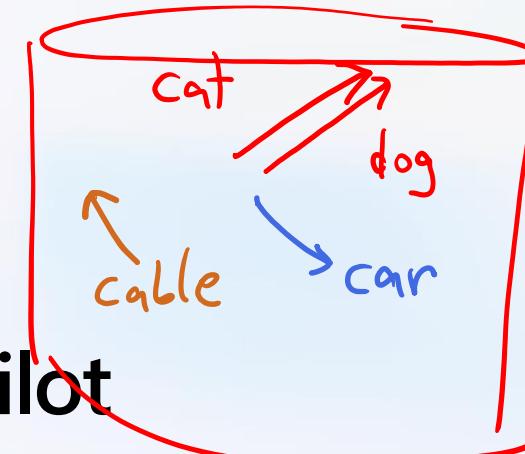
Tenant-wide and user-scoped
vector indices from Graph



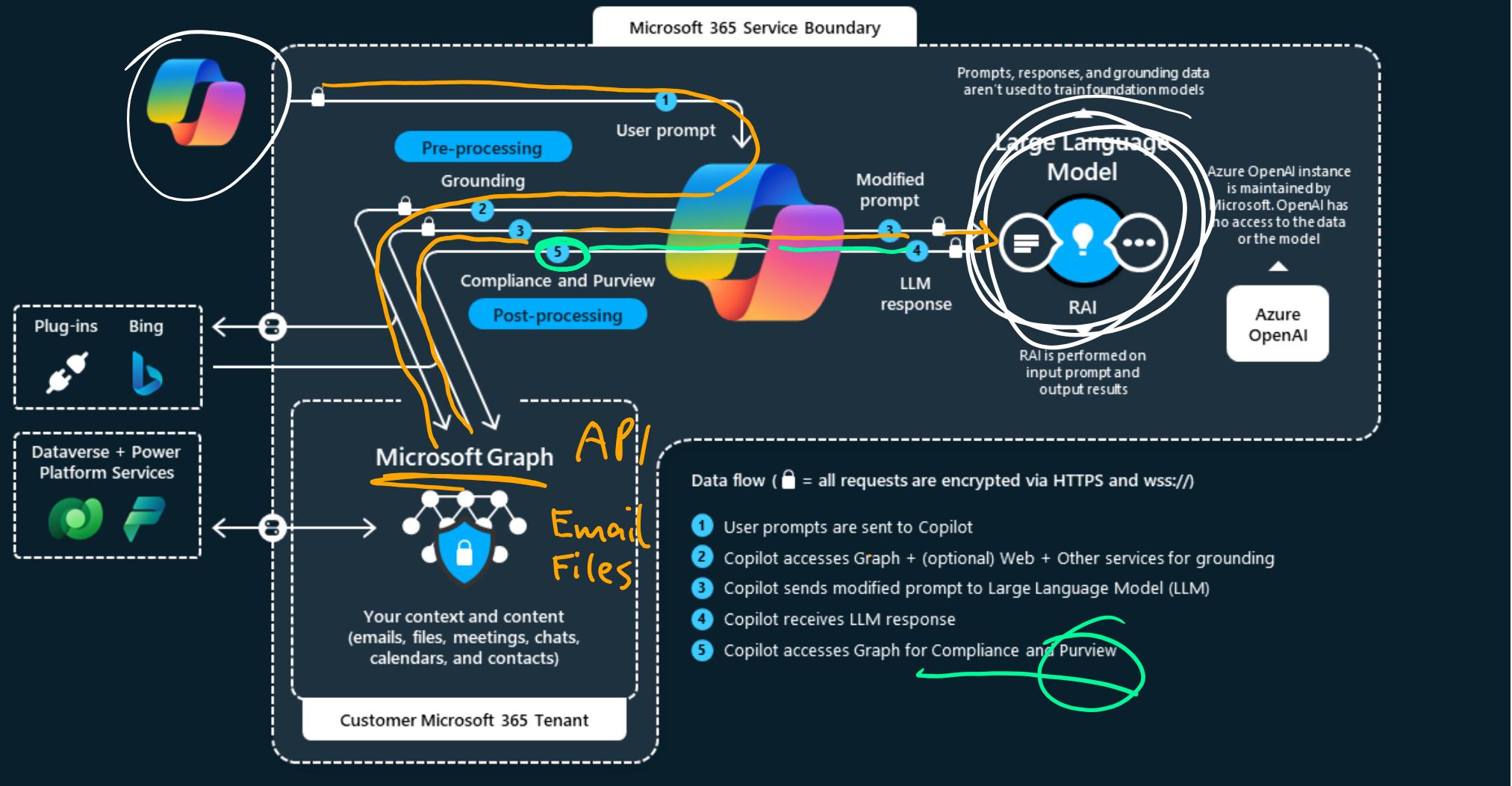
Identify semantics and similarities
around content and users



Bring your data into Semantic
Index with **Graph Connectors**



Microsoft Copilot for Microsoft 365 architecture



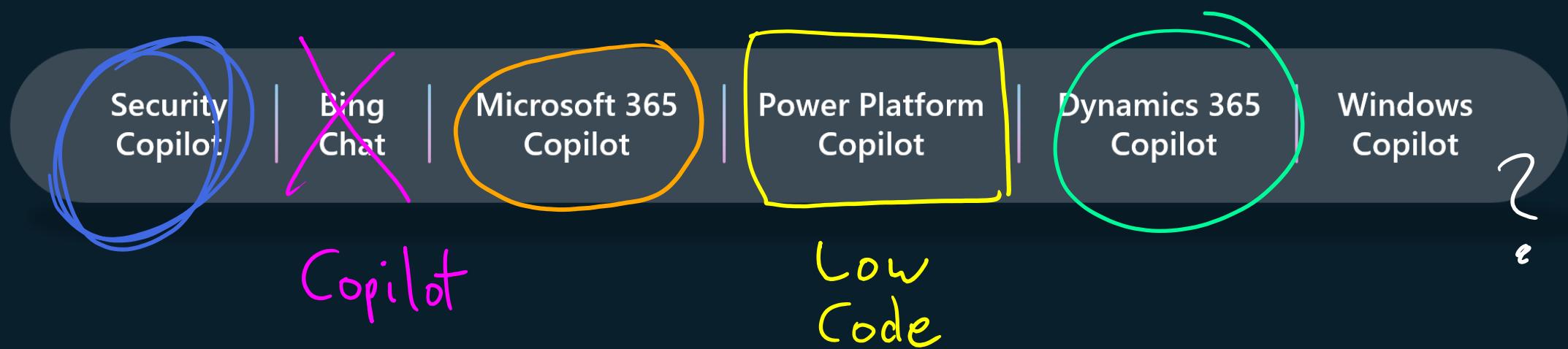


How it works:

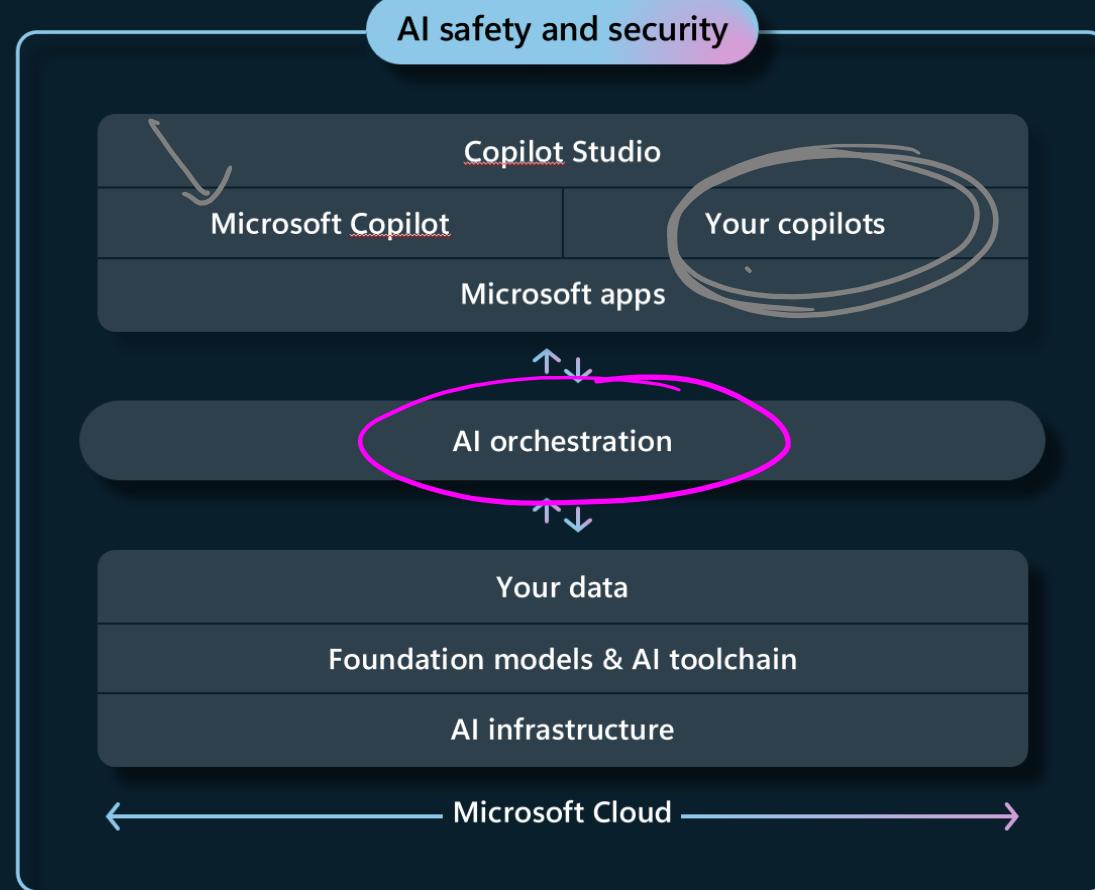
- Copilot receives an input prompt from a user in an app, such as Word or PowerPoint.
- Copilot then pre-processes the input prompt through an approach called grounding, which improves the specificity of the prompt, to help you get answers that are relevant and actionable to your specific task. The prompt can include text from input files or other content discovered by Copilot, and Copilot sends this prompt to the LLM for processing. Copilot only accesses data that an individual user has existing access to, based on, for example, existing Microsoft 365 role-based access controls.
- Copilot takes the response from the LLM and post-processes it. This post-processing includes other grounding calls to Microsoft Graph, responsible AI checks, security, compliance and privacy reviews, and command generation.
- Copilot returns the response to the app, where the user can review and assess the response.

Microsoft's Copilots

(Logic App)



Copilot Stack



Chat-GPT

Eine Bereicherung für die täglichen Aufgaben