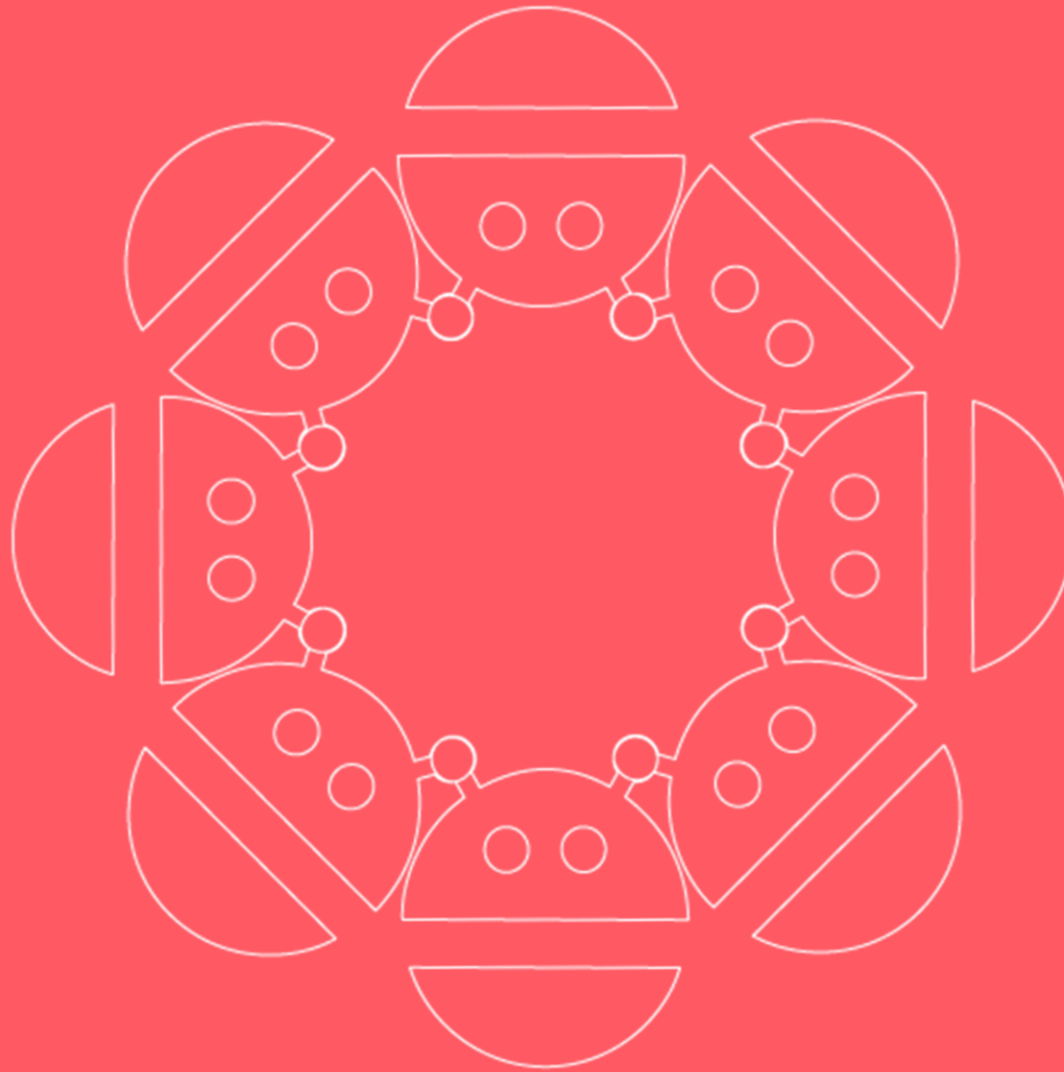


WienGPT



Stadt
Wien



GPT

Generative Pre-trained Transformer

- Sie sind in der Lage, aus Startwörtern, dem „Prompt“ (einer Frage), Sätze zu vervollständigen und Konversationen zu **generieren** (oder einfacher: zu antworten).
- Große Modelle werden sehr aufwendig über Wochen mit großen Texteingaben (z.B. u.a. die gesamte Literatur der Menschheit) **vortrainiert**.
- **Transformer** (neuronales Netz) sind ein mathematisches System, das durch die Analyse von Daten Fähigkeiten erlernt..

Großes Sprachmodell

(large language model, LLM)

Ein Sprachmodell berechnet aus vielen Texten die Wahrscheinlichkeiten, wie ein menschlicher Satz aufgebaut ist.

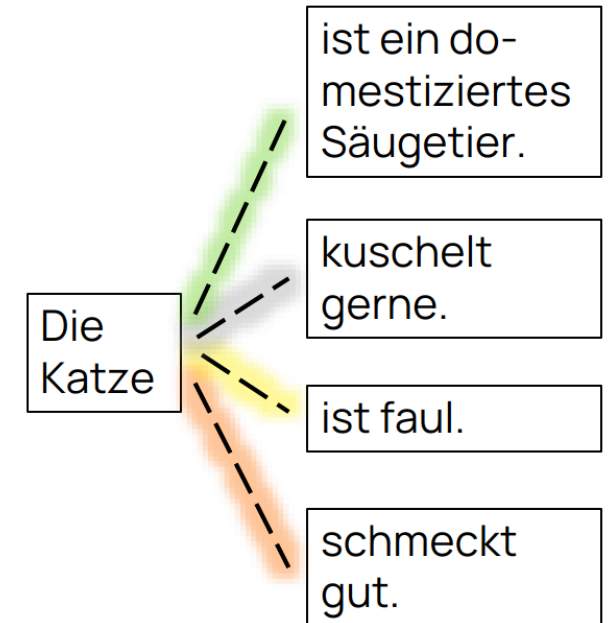
=Ein Sprachmodell das Worte nach Wahrscheinlichkeit vorhersagt

- ChatGPT
 - 430 Mrd. Wörter aus dem Internet (webcrawler)
 - 67 Mrd. Wörter aus Büchern
 - 3 Mrd. Wörter aus Wikipedia
- grundsätzlich „unsupervised learning“
(lernt selbständig ohne menschliche Kontrolle)

Reinforced Learning

bestärkendes Lernen oder verstärkendes Lernen

- Idee: Das Verhalten von GPT wird belohnt oder bestraft.
- Am Anfang bewerten Menschen die Antworten von GPT. (sehr zeitaufwendig)
- Aus dieser Bewertung wird ein Ethik-Modell trainiert, das die menschliche Bewertung vorhersagt.
- Mit diesem Ethik-Modell werden dann alle weiteren Antworten von GPT automatisch bewertet.
- Und damit die Wahrscheinlichkeit der Antworten verändert.




Grenzen/ Probleme bei der Nutzung von GPT

- „halluziniert“, erfindet also sehr überzeugend Antworten, vor allem wenn es keine gibt.
- übernimmt alle Inhalte des Textes (Input) – einschließlich des Bias, also aller Vorurteile.
- Sprachmodell kein Wissensmodell -> inhaltliche Fehler
- Textbasis vor September 2021
- Keine Verbindung zum Internet
- Urheberrechtsfragen
- Datenschutz

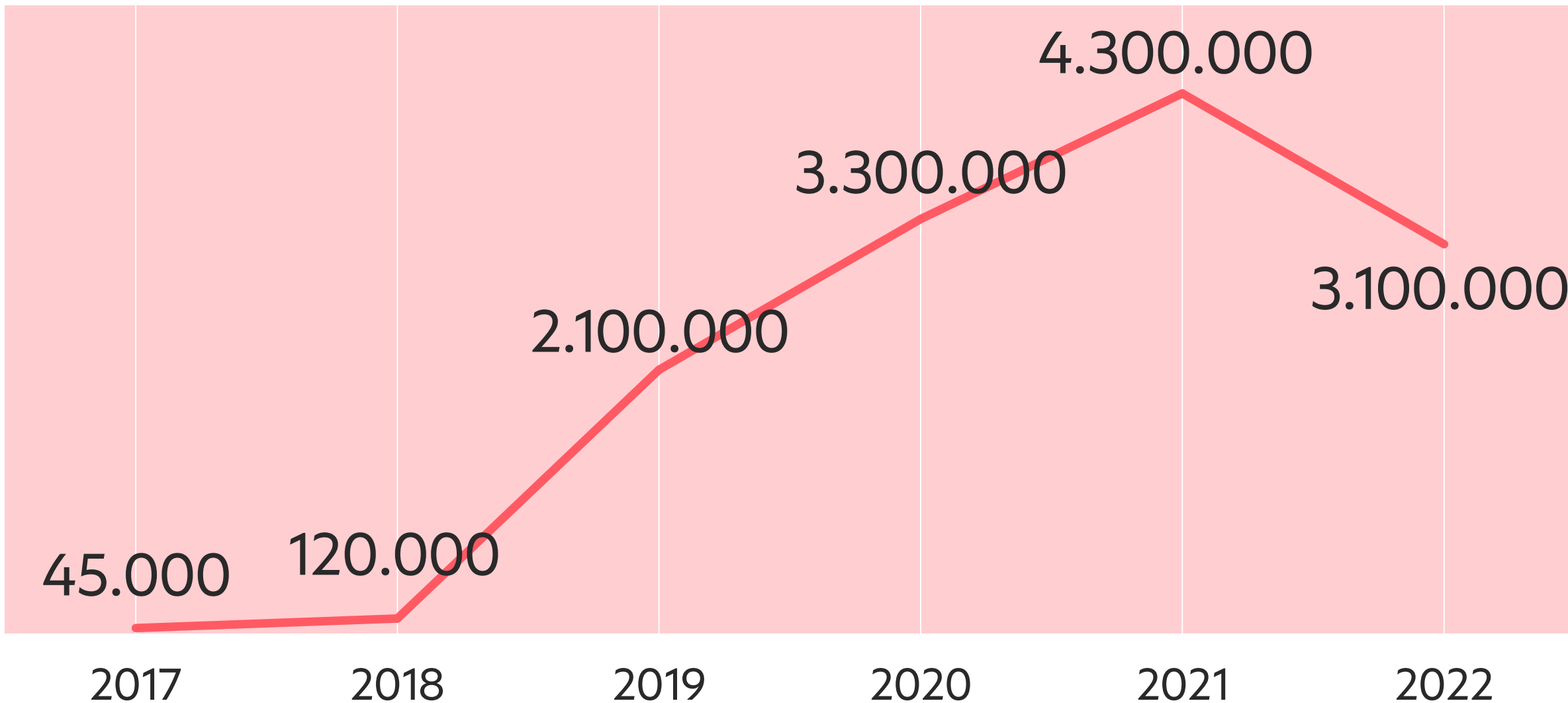
ChatGPT vs. ChatBot

	ChatGPT (LLM)	ChatBot (NLU)
Umfang	<ul style="list-style-type: none">• nahezu unbegrenzt• kennt einen großen Teil des menschlichen „Wissen“	<ul style="list-style-type: none">• in der Regel sehr spezialisiert• auf Stichworte und Themen beschränkt
Sprache	<ul style="list-style-type: none">• spricht fast alle Sprachen auf hohem Niveau• auch Programmiersprachen	<ul style="list-style-type: none">• hat nur allgemeine Sprechregeln
Vorteile	<ul style="list-style-type: none">• hat immer eine Antwort• weiß fast alles• folgt ethischen Schranken	<ul style="list-style-type: none">• verlässliche und genaue Antworten,• geringe Betriebskosten
Nachteile	<ul style="list-style-type: none">• Extrem hohe Trainingskosten• hohe Betriebskosten• erfindet Antworten (halluziniert)	<ul style="list-style-type: none">• Antworten eingeschränkt• oft nicht viel besser als eine Suchmaschine• Schwierigkeiten mit richtiger Konversation

ChatBots sind wieder leiwand

... waren sie aber eh immer 

Fragen pro Jahr



13.0000.0000

FRAGEN

Die Menschen lernen zu fragen

Die Antworten waren immer schon da 

Noch mehr Antworten für Wien

WienGPT macht's möglich

Sprachmodell (LLM) erweitern bzw. trainieren

- Hohe Kosten und Wartelisten
- Wissen wird nur ergänzt, die Vorurteile (Bias) des Modells bleiben
- Richtiger Zeitpunkt schwierig durch Schnelligkeit der Entwicklung

Smart Prompting

- Sprachmodell (LLM) wird nicht erweitert, es wird nur als Tool verwendet.
- Wissen und Inhalte werden in einer eigenen Datenbank durchsuchbar gemacht.
- Datenbank findet zur Frage den relevanten Inhalt (Textpassage).
- Die Textpassage wird zusammen mit der Frage dem Sprachmodell übergeben und gibt die Antwort.



Servus! Wie kann ich helfen?

Einige dieser Antworten werden automatisch von einer künstlichen Intelligenz erstellt und können Fehler enthalten.

Was sind die Voraussetzungen für den Energiebonus?

Was kostet parken?

Welche Unterlagen brauche ich für eine Förderung einer Markise?

Welche Gebäude bekommen eine Förderung für eine Markise?

Wer bekommt eine Förderung für eine Markise?

Referenz finden

1. Schritt

WienBot Antwort zur Frage liefert die Referenz auf eine wien.gv.at Seite

Die Stadt Wien vergibt Förderungen für die Installation von Rollläden, Außenjalousien und Markisen. 🌞

Voraussetzungen & erforderliche Dokumente:

- Gebäude: min. 10 Jahre alt
- Zustimmung der Hauseigentümer*innen
- Bewilligung innerhalb von Schutzzonen
- Fassaden-Markisen mit Qualitätsnachweis
- Rechnung bei Antragstellung: max. 6 Monate alt

🔴 Höhe der Förderung:
50 % der Kosten, max. 1.500 €

🚫 Keine Förderung bei Eigenheim, Kleingarten-Wohnhaus, Reihenhaus und Geschäftslokal.

📄 Antragsformular und weitere Infos:
www.wien.gv.at/gogv/l9vasonnenschutz

Kontext extrahieren

2. Schritt

Webseite laden und Text extrahieren.

The screenshot shows a web page from the City of Vienna (Stadt Wien) regarding the application for a grant for the installation of an external sun protection system. The page is titled "Montage eines außenliegenden Sonnenschutzes - Förderungsantrag". The main content area is highlighted with a red border. The page includes a navigation menu at the top, a sidebar with a search bar and various filters, and a main content area with the following sections:

- Wichtig:** A notice regarding the application deadline and the need to submit a completed application form.
- Informationen:** Details about the grant, including the maximum amount of 10,000 EUR and the requirement to submit a completed application form.
- Voraussetzungen:** A list of conditions for the grant, including the applicant's status, the type of property, and the location of the sun protection system.
- Fristen und Termine:** Information about the application deadline, which is 31. December 2023.
- Zuständige Stelle:** The responsible authority, the Department for Urban Planning and Construction, is mentioned.
- Verfahrensablauf:** A description of the application process, from submission to decision.
- Erforderliche Unterlagen:** A list of required documents, including a completed application form, a site plan, and a technical drawing.
- Kosten und Zahlung:** Information about the grant amount and the payment process.
- Formular:** A link to the application form.
- Zusätzliche Informationen:** Additional information about the grant and the application process.



WienGPT

Frage stellen

3. Schritt

Sprachmodell (GPT):

Kontext [Text von der Webseite]

Beantworte die Frage nur anhand des Textes oberhalb:

„Wer bekommt eine Förderung für eine Markise?“

Antwort geben

4. Schritt

Antwort von Sprachmodell mit:

- Hinweis auf automatisierte Erstellung durch KI
- Quellenangabe
- Weiterführende Themen (Quicklinks) vom WienBot

Wer bekommt eine Förderung für eine Markise?

Mieter*innen von Wohnungen, Eigentümer*innen von Wohnungen und Eigentümer und Eigentümer von Mehrfamilienhäusern oder deren Hausverwaltungen können eine Förderung erhalten.

Diese Antwort wurde automatisch von einer KI mit folgenden Quellen beantwortet:

WienBot,

<https://www.wien.gv.at/gogv/l9vasonnenschutz>

Gebäudebegrünung

Smart Prompting

Konzept

1. **Suche (Embeddings)**
Findet die relevante Textpassage (Webseite) zu Frage

2.

The screenshot shows the website of the City of Vienna (Stadt Wien). At the top, there are language options: english, bosanski, hrvatski, srpski, and türkçe. Below that is a navigation bar with categories: Themen, Virtuelles Amt, Stadtplan, Video, and S., Margareten. A secondary navigation bar lists various service areas: Persönliche Dokumente, Gesellschaft & Soziales, Gesundheit & Pflege, Freizeit & Sport, Kultur & Bildung, Finanzielles & Förderungen, Wirtschaft, and Bauen & Wohnen. The main content area shows a breadcrumb trail: wien.at > Virtuelles Amt > Bauen & Wohnen > Wohnen. A sidebar on the left lists sub-topics: Bauen, Energiesparen, Grund und Kleingarten, Mietrecht, Stadtgestaltung, and Wohnen (which is selected). The main content area displays the title 'Montage eines außenliegenden Sonnenschutzes - Förderungsantrag' and a button 'Seite vorlesen'. A prominent yellow box contains the following text: 'Bitte beachten Sie: Auf Grund der derzeitigen COVID-19 Situation ist ein eingeschränkter KundInnenverkehr ausschließlich mit Online-Terminreservierung oder telefonischer Terminvereinbarung unter Berücksichtigung von Schutzmaßnahmen möglich. Online-Terminreservierung Bitte beachten Sie die Schutzmaßnahmen (FFP2-Schutzmasken-Pflicht in Amtsgebäuden, Mindestabstand 2 Meter, Hygienemaßnahmen). Bitte bringen Sie Anträge und Unterlagen weiterhin per Post ein: Wohnbauförderung und Schlichtungsstelle für wohnrechtliche Angelegenheiten 20., Maria-Restituta-Platz 1, per Fax: +43 1 4000-997 48 79 oder per E-Mail: wv@ma50.wien.gv.at ein oder nutzen Sie das vorhandene Online-Formular-Service. Die Stadt Wien fördert die nachträgliche Montage von außenliegenden Rollläden, Jalousien und Fas-

3. **GPT (LLM):**
Text von der Webseite
Beantworte die Frage nur anhand des Textes oberhalb: „Wer bekommt eine Förderung für eine Markise?“

4.

Wer bekommt eine Förderung für eine Markise?

Mieter*innen von Wohnungen, Eigentümer*innen von Wohnungen und Eigentümer und Eigentümer von Mehrfamilienhäusern oder deren Hausverwaltungen können eine Förderung erhalten.

Diese Antwort wurde automatisch von einer KI mit folgenden Quellen beantwortet:
WienBot,
<https://www.wien.gv.at/gogv/l9vasonnenschutz>

Gebäudebegrünung

Suche mit GPT

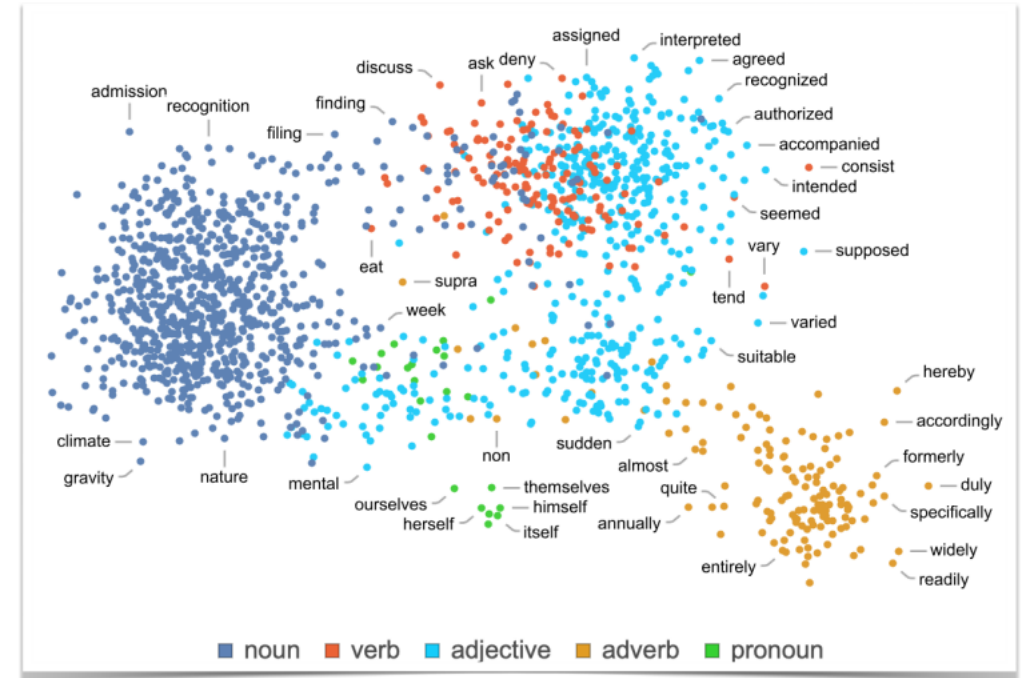
"Nebenprodukt" von großen Sprachmodellen

Nutzen der Vektor-Darstellung, die LLMs nutzen, als Suchdatenbank

- Ersetzt klassische, "stichwortbasierte" Suche (Solr, Elasticsearch, etc.)
- Beantwortung von langen Eingaben / ausformulierten Fragen möglich

Durchsuchbare Inhalte werden als Vektor in einem "Semantischen Raum" gespeichert und durchsucht

- so genannte "Embeddings"
- ▶ zB. OpenAI: ~480 Dimensionen



Herausforderungen

Für den Produktiveinsatz

OpenAI (API)

Datenschutzniveau

- OpenAI APIs außerhalb der EU
- Alternative über Microsoft AZURE Cloud

Privatsphäre

- Fragen werden nicht fürs Training verwendet
- Alle Fragen werden nach 30 Tagen gelöscht

Performance

- Lange Antwortzeiten (wird besser)

Kosten

- Verrechnet wird nach Token (1-3 Zeichen) für Eingabe (Frage) UND Ausgabe (Antwort)
 - Preis variiert noch Modell (Fähigkeiten) und Größe des Kontext (Maximale Anzahl Token)
 - Da bei jeder Frage die relevante Textpassage (Kontext) mitgegeben wird, relativ hohe Kosten.
 - Deutsche Sprache verbraucht mehr Token als Englisch
- **Open Source Modell als Alternative?**

Urheberrecht

- Das aktuelle **Urheberrecht, gilt nur für** von **Menschen** erstelle Texte.
 - Von einer **KI erstellte Texte sind daher nicht urheberrechtlich geschützt.**
 - **Ausnahme:** wenn bei der Erstellung „**der Einfluss des Menschen maßgeblich war**“.
- **Das geltende Urheberrecht ist derzeit nicht in der Lage, die Rechtsfragen und „Verteilungskonflikte“ bei KI-basierten Werken zufriedenstellend zu regeln.**

Problemstellungen

- Halluzinieren / Erfundene Antworten
- Prompt-Injection
- übernimmt alle Inhalte des Textes (Input) – einschließlich des Bias, also aller Vorurteile.
- Verlässlichkeit und Konsistenz
- Verantwortungsvolle KI - Filter

Lösungsansätze

- Inhalt für die Antwort als Kontext mitgeben (Smart Prompting)
- Prompt Engineering: Eine konkrete Fragestellung (Prompt) soll falsche Antworten möglichst verhindern.
- Sprachmodell konfigurieren (System Prompt)
- Suchoptimierung: Überprüfen, wie relevant die übergebene Textpassage ist.
- Überprüfen, ob sich die Antwort wirklich auf den übergebenen Text bezieht. z.B. die Frage mehrfach beantworten lassen und dann vergleichen. (Few-Shot)
- „Temperatur“ Einstellung entscheidet wie kreativ die KI ist.

Wie geht es weiter?

Evaluierung verschiedener Modelle

- Suche (Embeddings, Vektor-Datenbanks)
- Sprachmodelle:
 - MS OpenAI
 - Aleph Alpha
 - Google
 - Open Source