



Cyber [security] affects everything, but
AI will affect [the future]
even more fundamentally.

Benjamin Netanyahu

Ein paar Worte zu mir



Sebastian Schedy

Ultramarathon + Katzen-Aufpasser

Nebenberuflich

MBA, Uni Liechtenstein

Hauptberuflich

IT Solutions Engineer, LGT

Davor

Data Analysis Manager, tk Presta

Techniker Maschinenbau, Lämpfle







Das Projekt

ChatGPT wird bekannt

Idee **ChatGPT** für uns nutzbar machen

unsere Daten

No Cloud = Only Sunshine

Alte Technik

n-Gram Vorhersage

des nächsten Buchstabens

der nächsten Buchstabenkette

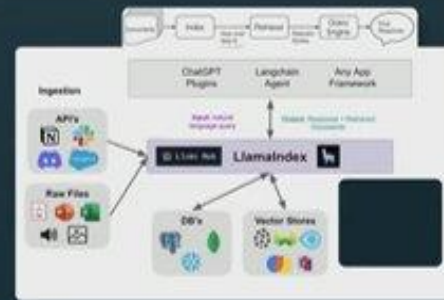
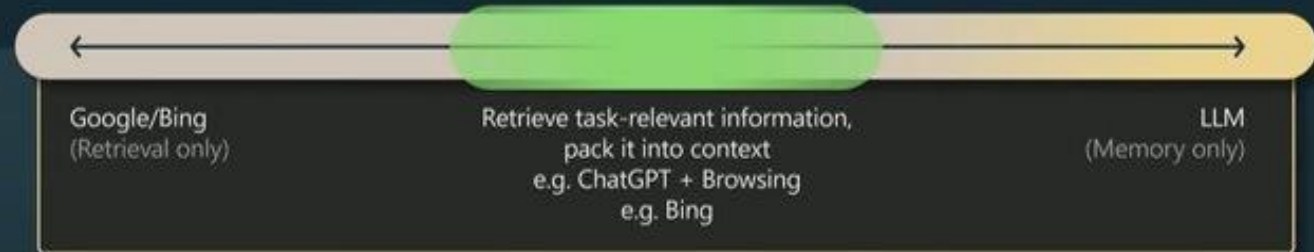


Inspiration



Retrieval-Augmented LLMs

Load related context/information into "working memory" context window



Emerging recipe:

- Break up relevant documents into chunks
- Use embedding APIs to index chunks into a vector store
- Given a test-time query, retrieve related information
- Organize the information into the prompt

Andrej Karpathy, State of the GPT

Neue Technologie

Transformer-Architektur

Sätze, die Sinn ergeben

oft sogar richtig, meistens überzeugend



Kombination Vektorsuche + Sprachmodelle

Überschreiten der Schwelle

Ausrüstung

2x Open Source LLM

1x Datenbank

1x Ähnlichkeitssuche

1x **Vision**





Prüfungen, Verbündete und Feinde



Ressourcen schonen

Anweisungen folgend

Tischmanieren

Verlässlich



Sponsor

IT

Topmanagement



Ähnliche Projekte

Meine Freizeit

Zeitdruck



Entwicklung – Hochskalieren

1. **Chat** auf Laptop
2. Chat mit pdf
3. Erster Prototyp
4. technische Umstellung (LLM Dateiformat)
5. Upgrade von Experiment zu App
6. neue Chatoberfläche
7. Code-Cleanup
8. technische Umstellung (LLM Dateiformat)
9. Wechsel von Laptop zu Linux Server
10. Code-Upgrade



Hallway Testing

Übersetzen für das Sprachmodell

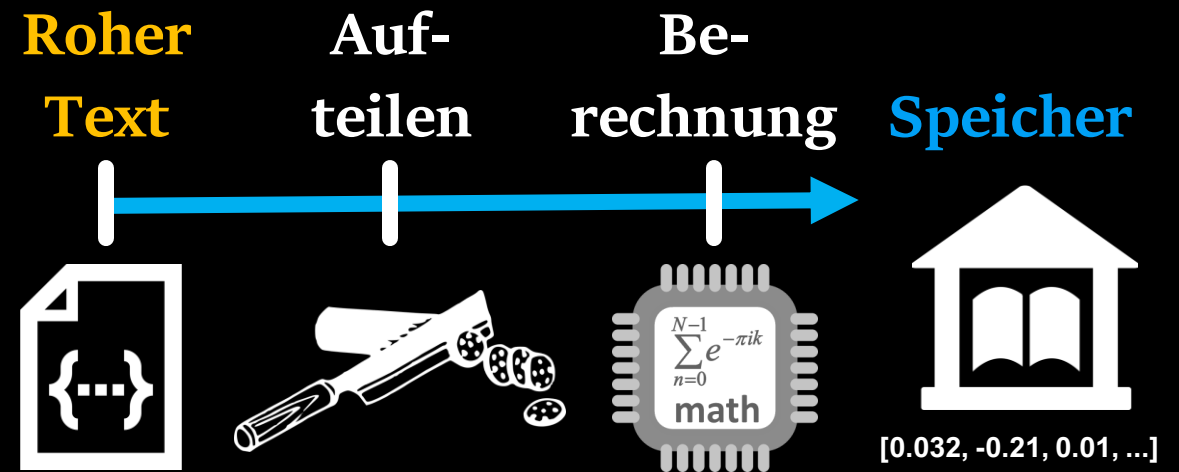
Roher Text

in 500-Zeichen Stücke

>110'000 Stücke

~ 33 Stunden Berechnung

1 GB Daten

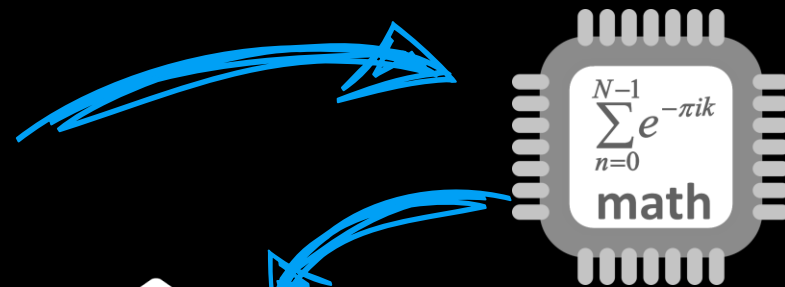


In der Maschine

Start



INA Lager sind nicht laut, dafür teuer.

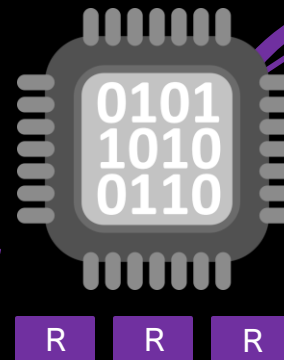


Frage wird umgerechnet



Berechnete Frage

Ähnlichkeits-
suche
-
3 Stücke +
Referenzen



Daten rein Frage Antworten

Willkommen beim ChatBot

Welche Lager sind leise? 

Eine leise Option sind Lager von INA.
Quelle: [Wiki Seite 3](#)





Belohnung und der Weg zurück

Vorstellung

Innovation Fair

interne Demos

Präsentation

The screenshot shows a chatbot interface with a white background and a light blue header. The header contains the text "Welcome" followed by a redacted area. Below the header, the chatbot's response to the user's question "Heyyy, can you tell me how I start a quality complaint in MHP CCX?" is displayed. The response includes a detailed list of steps for filing a complaint and a source link. The user's follow-up question "Where is MHP CCX?" is shown in a text input field at the bottom, with the chatbot's response indicating that it cannot answer due to a lack of context. The interface includes a "Send" button and a footer with a disclaimer.

Welcome [redacted]

Heyyy, can you tell me how I start a quality complaint in MHP CCX?

Bot: Sure! To start a quality complaint in MHP CCX, follow these steps: 1. Log in to your account on the MHP website. 2. Go to the "Complaint Management" section of the main menu. 3. Click on the "New Complaint" button. 4. Choose the type of complaint you want to file (e.g. Customer complaint, Quality complaint, etc.). 5. Fill out the complaint form with as much detail as possible, including the reason for the complaint and any relevant information. 6. If necessary, add attachments to support your complaint. 7. Once you've completed all the required fields, click "Save" to submit your complaint. That's it! Your complaint will now be processed in MHP CCX.

Sources:

- Title: MHP CCX (Complaint Cockpit Extension), Source: [https://\[redacted\]](https://[redacted])


[redacted]

Where is MHP CCX?

Bot: I'm sorry, but I cannot answer the question as there is no information provided in the context about MHP CCX. Could you please provide more context or information so that I can assist you better?

Sources:

[redacted]

Where is MHP CCX? 

Questions and answers are recorded to help our services.
The Chat Bot does not always tell you what you expect. This is not on purpose.

Herausforderungen der Technologie

Wahrheit

schlau genug für Antworten, nicht für Diskussionen

nicht deterministisch

Tischgespräche

Alignment/Ausrichtung der neuen Mitdenker

Der Bananentest

**Heute habe ich vier Bananen. Gestern ass ich zwei.
Wie viele Bananen habe ich jetzt?**

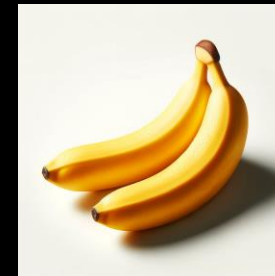
Today I have four bananas. Yesterday I ate two. How many bananas do I have now?

GPT3.5



ChatGPT

You currently have two bananas.



GPT4



ChatGPT

If you have four bananas today and you ate two yesterday, then you still have four bananas today. Eating bananas yesterday doesn't affect the number of bananas you have today, assuming you've already accounted for those you ate. So you have four bananas now.





Was seit dem passierte

ChatGPT verändert weiterhin die Welt

One-Trick-Pony? Nein

viele kleine Schritte



No Cloud - Only Sunshine?

→ einfacher als je zuvor

*The road to wisdom?
Well, it's plain
and simple to express:*

*Err
and err
and err again
but less
and less
and less.*

– Piet Hein

schedy+uni@posteo.de



© J. J. J.