

Ah, That's Why!

The tech behind AI chatbots

Tilman Scheel & ASK

Lawyer, art historian, lecturer on AI

IT for film festivals and libraries since 2004

CEO of ava-library.com and **AboutSomethinK.org**



ASK is the first AI – Chatbot for public libraries in the world. **ASK** is in production for all Berlin libraries since June 2024 on voebb.de



Collaborations

The screenshot shows a library website with a chatbot test interface. At the top, there is a navigation bar with links for 'ANMELDEN', 'MEIN KONTO', 'SITZUNG BEENDEN', 'English', 'DGS', 'Leichte Sprache', 'Häufige Fragen', and 'Kontakt'. Below this is a banner for 'Verbund der Öffentlichen Bibliotheken Berlins' with the slogan 'WIR BRINGEN MEDIEN IN BEWEGUNG'. The main content area features the 'de Bib' logo and 'testbib Zonedorp'. A navigation menu includes 'Home', 'Nieuws', 'Activiteiten', 'Collectie', 'Doelgroepen', and 'Chatbot-test'. A search bar is present with the text 'Überall suchen' and a 'Suchen' button. A chatbot test box is highlighted with a green border, containing the text 'de Bib tipt' and 'Test tot 15 december'. Below the chatbot test box, the URL 'https://zonedorp.bibliotheek.be/test-chatbot' is displayed. To the right, there is a 'Druckversion' button and a 'URL: voebb.ava.watch' link.

datronicsoft

koha

axiell

a|S|tec|
angewandte Systemtechnik eG

ASK

Terminology & Challenges

Recommendation bot

similarity search, probability

Service bot

precision and timeliness, hallucinations

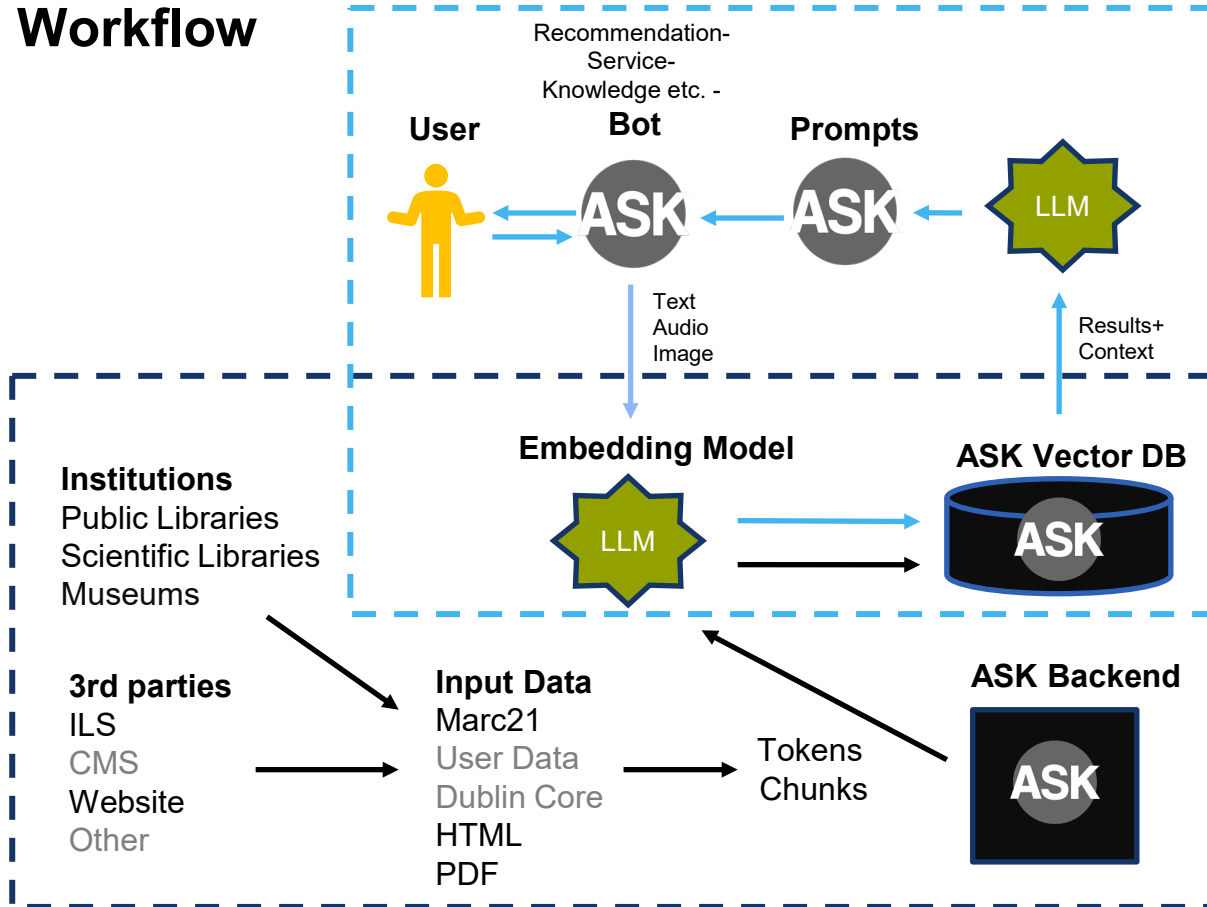
Knowledge bot

format chaos, data protection

Media bot

amount of data, copyright, references

RAG Workflow



Conversations

Terms:

- Tokens
- Chunking
- Dimension
- Embedding
- RAG
- graph RAG
- Hybrid Search
- Context
- Hallucinations
- Prompting
- Metrics

Metadata



Alternatives: RAG vs Fine Tuning



Prompt: A librarian holds a book and appears to teach a child. Style of Abd al-Samad and Mir Sayyid Ali.



Prompt: Take the same image and let the child hold the book

Trade-Offs & Peculiarities

(in tech ...)

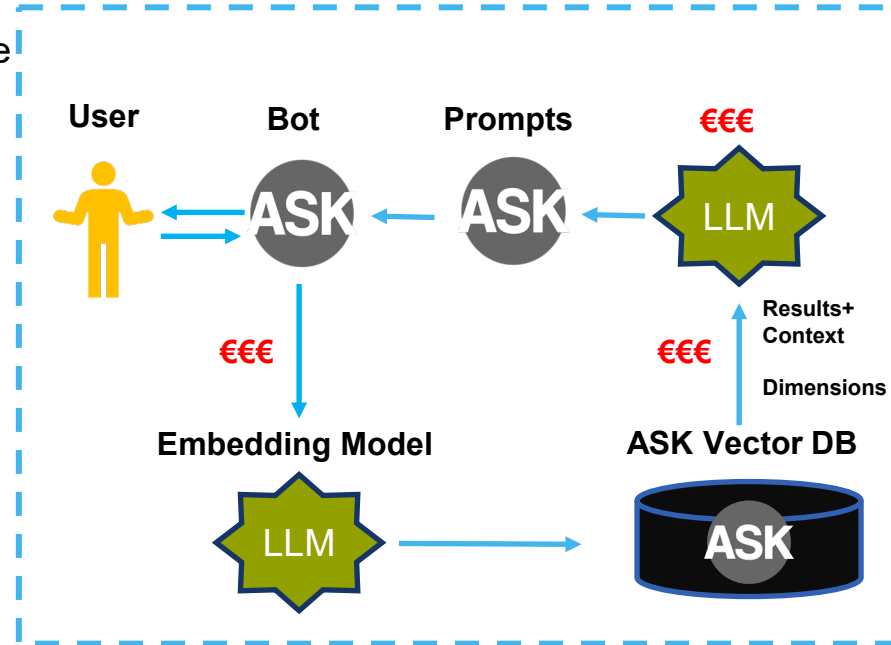
Trade-Offs

Size of the LLM vs cost vs speed vs usability of the results (not: precision!)

Controls & data protection vs cost and effort (self-hosting - API usage)

Jack of all trades vs AI “team”

Prompting & Context vs cost vs speed vs usability of the results



Trade Offs - Size of Prompts

```
HTTP 200 OK
Allow: GET, PUT, PATCH, DELETE, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

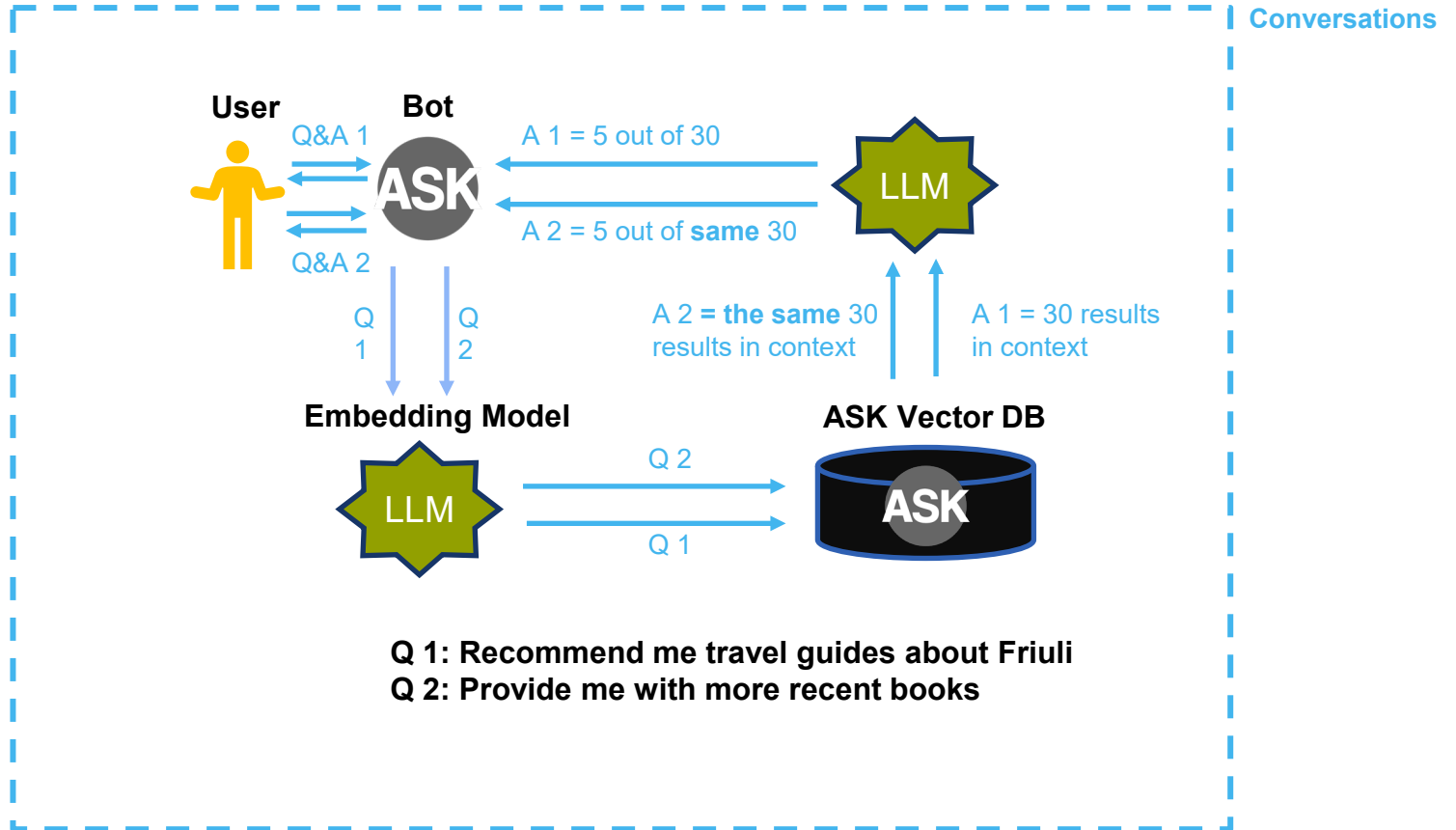
{
  "id": "bf04723e-8ee3-46ee-92c8-2cda7a3e20bb",
  "response": "Adolf Hitler was the leader of the National Socialist German Workers' Party (NSDAP) and served as Chancellor of Germany from 1933 to 1945. He is best known for his role in instigating World War II and for the Holocaust, which led to the systematic extermination of six million Jews and millions of others deemed undesirable by the Nazi regime. Hitler's aggressive expansionist policies and totalitarian rule had a profound impact on Europe and the world, leading to widespread destruction and loss of life. His ideology was rooted in extreme nationalism, anti-Semitism, and the belief in Aryan racial superiority.

If you're looking for books or films about Hitler or related topics, I can help you find some recommendations. Would you like that?",
  "state": {
    "messages": [
      {
        "role": "assistant",
        "content": "Hallo! Ich bin dein hilfreicher KI-Empfehlungsassistent für den Verbund der Öffentlichen Bibliotheken Berlins. Ich erweitere die klassische Suchfunktion, indem ich dir gezielte Empfehlungen aus unserem umfangreichen Katalog anbiete.

Stell dir vor, du planst eine Reise und möchtest die besten Reiseführer für dein Ziel finden, aber auch spannende Romane, die dort spielen, um dich in die Kultur einzutauchen. Ich kann dir helfen, eine Liste von Reiseführern und Romanen zu erstellen, die perfekt auf dein Reiseziel abgestimmt sind.

Was für eine Art von Empfehlung suchst du heute?"
      },
      {
        "role": "user",
        "content": "Tell me something about Hitler"
      }
    ]
  }
}
```

Trade Off – Context Size



Probability vs. Precision / Predictability

The cat chases

- Samurai 5%
- Mouse 70%
- Squirrel 20%
- Temple 5%
- Fuji 0%



Prompt: Cat thinking about chasing five very different items, including a mouse, hokusai style

Similarity Search

Question:

Can you recommend books in the style of „A Tramp Abroad?“ by Mark Twain

Problem:

The most similar books are the various editions of „A Tramp Abroad“

Solution:

Exclude Function



Prompt: an image with 4 boys, looking and dressing similar because they are quadruplets. Michelangelo style

„Niche Data“

Questions:

Do you have books of Mosolff?

Problem:

- LLMs store semantics in a context.
- They store vectors, not letter sequences.
- Proper names have no meaning, and if they are very rare, there is no context in which they could be embedded.

Solution:

- Hybrid Search
- Time



Prompt: design an image of a crazy scientist who is writing a large volume of a book, Sumi-e style

Response Engine

Question:

Until when can I return my book?

Problem:

Precision and timeliness

Solutions:

Prompts & Links



Prompt: a khmer sculpture representing a know-it-all

Response Engine – Live

Prompt für den Chatbot *

Information About VÖBB Libraries

- Annual membership typically costs 10 €, free for persons under 18.
- No library card is needed to use the desks.
- Accounts can be managed online at [Mein Konto](adisintern:*SBK).
- Users can register in person or online at [Online-Anmeldung](/ausweis).
- Fees can be paid online or in person.
- For contact info, opening hours, library cafés, accessibility, and locations, refer users to [Kontakt & Standorte](adisintern:*SW320).
- An overview of all VÖBB's online media is available at [Digitale Angebote](adisintern:*SW2). These include Onleihe, Genios, OverDrive, Pressreader, AVA, Filmfreund, Tiaerbooks, Brockhaus, Munzinaer, Duden, Freeeal Music, Naxos Music, and phase6

Der Prompt der zur Instruktion des Chatbots genutzt wird. Wird als Jinja2 Template ausgewertet.

Sprache des Nutzers erkennen

Erkennt die Sprache der Nachrichten des Nutzers. Stellt die erkannte Sprache als Template Variable `detected_language` bereit

Structured Data vs. Natural Language

Question:

I'm looking for health journals

Problem:

Metadata is not always conclusive
(according to the understanding of an LLM).

Possible solutions:

- Prompt the specific metadata
 - Customize importer
 - Fine-tuning
- Research Project: How to prepare structured metadata (chunking, „naturalizing“, weighing...)



Prompt: Two women discussing with each other. One in a baroque costume the other dressed strictly in a kimono, looking a bit in a bad mood. Rubens style.

Structured Data vs. Natural Language - Live

The title is Bewegungstherapie und Gesundheitssport B & G ; offizielles Organ des Deutschen Verbandes für Gesundheitssport und Sporttherapie e.V. (DVGS).

....

A variation of the title is "B & G".

A variation of the title is "B und G".

The link id is 990125318930206441.

This catalog entry refers to an **electronic resource**.

This catalog entry has been produced in Germany.

The physical description is **Online-Ressource**.

The content type is text.

The media type is computer.

The carrier type is online resource.

[Link](<https://doi.org/10.1055/s-00000119>).

[Link](<https://ezb.ur.de/?2141948-6>).

Index terms are **Zeitschrift**.

Vision – Do science, spread knowledge

VÖBB KI Labor

gpt-4o

Deutsch (de) ▾

Verfügbare Sprachmodelle

gpt-4

gpt-4o

gpt-4o-mini

o1-mini

o1-preview

Abmelden

Bot vor 1 Minute

Natürlich! Ich stehe bereit für Ihre Fragen oder Aufgaben. Was möchten Sie testen?

vor weniger als 1 Minute

Welche Prompts soll ich für eine Ähnlichkeitssuche nutzen= ?

Bot vor weniger als 1 Minute

Für eine Ähnlichkeitssuche, bei der Sie nach Inhalten suchen, die einem bestimmten Thema oder einer bestimmten Anfrage ähneln, können Sie verschiedene Arten von Prompts verwenden. Hier sind einige Ansätze, die Sie in Betracht ziehen können:

1. Themenbasierte Prompts:

- "Finde Artikel, die sich mit [Thema] befassen."
- "Suche nach Inhalten, die ähnlich sind wie [Thema/Artikel]."

2. Fragenbasierte Prompts:

- "Welche anderen Inhalte behandeln ähnliche Fragen wie [Frage]?"
- "Welche Artikel beantworten ähnliche Fragen wie [Frage]?"

Neue Konversation

Nachricht eingeben

Absenden

Neue Zeile: Umschalt + Enter

Vision – Networks of knowledge and culture


Select a role

Tourist

Child


Scholar



Artwork History

 Marble statue of a youthful Hercules

[+ Start New Chat](#)

☰

 **Marble statue of a youthful Hercules -**


Welcome to The MET! Enter the ID of a particular work. The bot will take it from there. For those who want to know more, we will suggest literature from the Watson library.



Please choose your role and get started.

Tourist


Child

Scholar

 **Marble statue of a youthful Hercules -**

The "Marble Statue of a Youthful Hercules" is a Roman sculpture from the Early Imperial period, specifically the Flavian era (69–96 CE). This impressive marble piece stands over 97 inches tall and was originally part of the Giustiniani Collection in Rome before becoming part of the Metropolitan Museum of Art in 1903, thanks to the generosity of Mrs. Frederick F. Thompson. The statue underwent significant restorations in the early 17th century, with

Ask anything 

Homework

The VÖBB chatbot's metaprompt is freely available at <https://github.com/voebb-dev/voebb-chatbot>

Test the VÖBB chatbot and identify problems, malfunctions, weaknesses.

Is there a fundamental weakness here (see above), or is there a chance to do it better using prompting?

If the latter: try to adapt the meta prompt to fix the problems.

In another session, we will test selected approaches live in the bot. You can register here:

12 March 5pm

Please register here:

https://us06web.zoom.us/meeting/register/VXI_1nTIT4W5BEb_sXKJPSQ



Answer everything vs. Not really everything

Question:

I'm looking a biography of Hitler

Problem:

Prompting prohibits recommending books by Hitler.

This restricts the recommendations in the context of National Socialism too much.

Solutions:

Finely granulated prompting,

Now what?

Play, Test, Learn, Share

DO!

www.aboutsomethink.org/pricing

Thank you !

tilman.scheel@aboutsomethink.org

<https://www.linkedin.com/in/tilman-scheel-2221297/>

Survey please!

<https://aboutsomethink.org/webinars>

