

Integrating LLMs: Intelligence is tricky

FOSDEM 2024



... Large Language Model (LLM) is a program that can process and respond with natural language



Paraphrased definition of large language model (LLM)
from [Mozilla AI Guide](#)

Getting started

Terminology overload

“Tokenization, Embedding, Positional Encoding, Encoder, Decoder, Attention Mechanism, Prompts, Tokens, Embeddings, Hidden states, Query, Attention scores, Retrieval-Augmented Generation (RAG), Top-K, Fine-tuning, Chain-of-thought...”

A lot of different technology

Fast changing space

LLMs have utility

Create 3 related keywords to the following text:
"Free and open software is a great thing."

Sure, here are three related keywords for the given text:

1. "Open source"
2. "Software liberty"
3. "Digital freedom"

Take the following data and translate it to a JSON object:

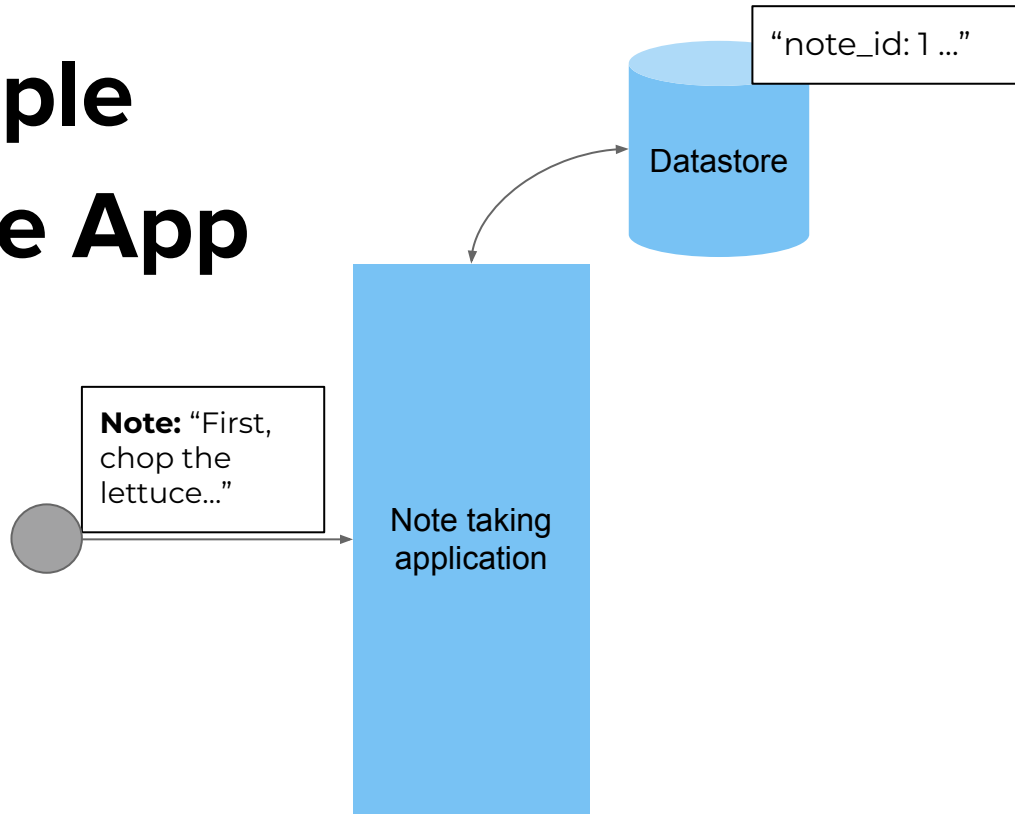
"ingredients:
- Lettuce x1
- Tomatoes x3
- Cucumber x2"

Sure! Here's a possible JSON object that represents the data:

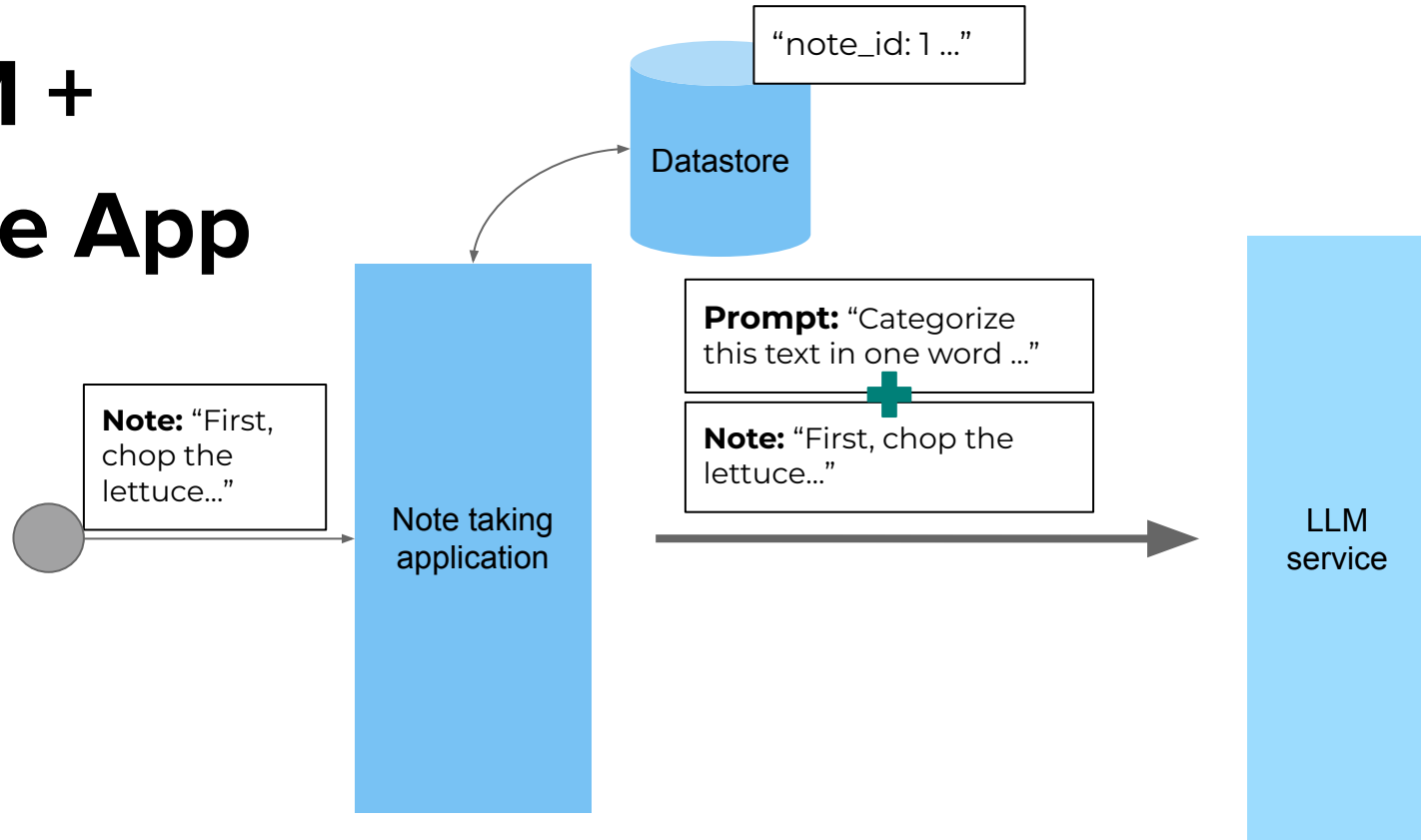
```
{
  "ingredients": [
    {
      "name": "Lettuce",
      "quantity": 1
    },
    {
      "name": "Tomatoes",
      "quantity": 3
    },
    {
      "name": "Cucumber",
      "quantity": 2
    }
  ]
}
```

Model: [mistralai/Mixtral-8x7B-Instruct-v0.1](#)

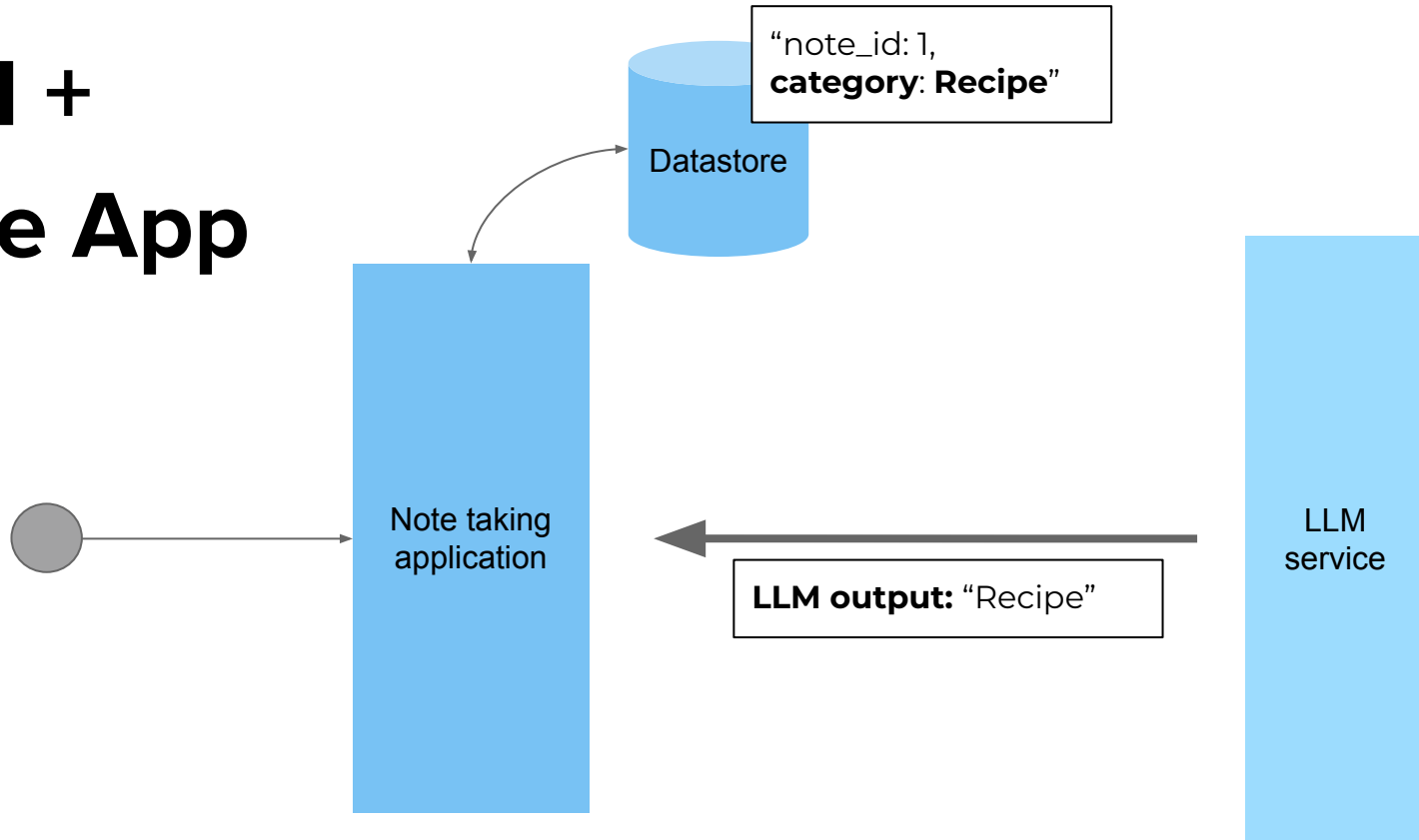
Simple Note App



LLM + Note App



LLM + Note App



Reality check

Hallucinations

- Fact or not?

Inconsistent format

- JSON response 

Performance

- GPUs are expensive

Text count limitations

- Increased complexity

Education/Documentation

- What prompts work?

Friction points

- Bugs, security

There are a lot of choices but...



Proprietary & closed ecosystems

Model dependant prompts

Prompt: "Define the word "open" in latin"

Responses CLEAR CONVERSATION

USER define the word "open" in latin

AI Apertus - open, unfolded, exposed to view, laid bare, uncovered, accessible, exposed, free from hindrance, easy, plain, clear, manifest, evident, public, liberal, generous.

👍 🗨

Model: [Bison@002](#) via Vertex

User: Define the word "open" in latin

Llama: . nobody is able to do it, you can try to find the word on internet but it will be a translation of open in english and not in latin.

Answer: Open

Model: [Llava 1.5](#) running locally with [llamafire](#)

Chatbot

The Latin word for "open" is "apertus".

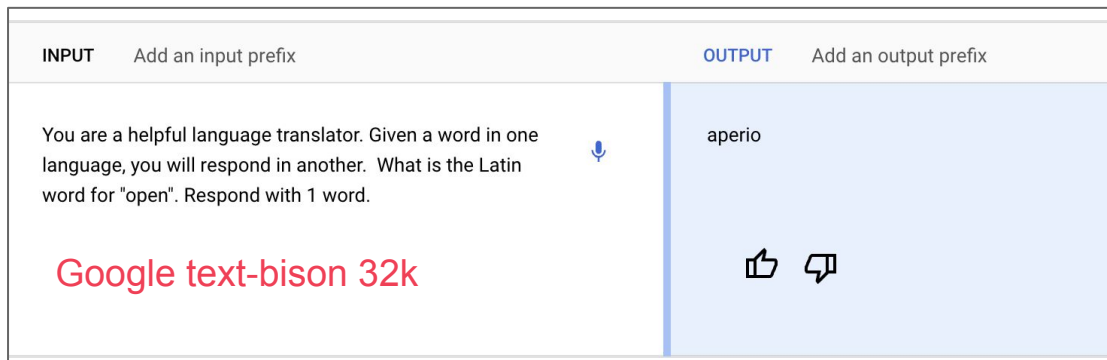
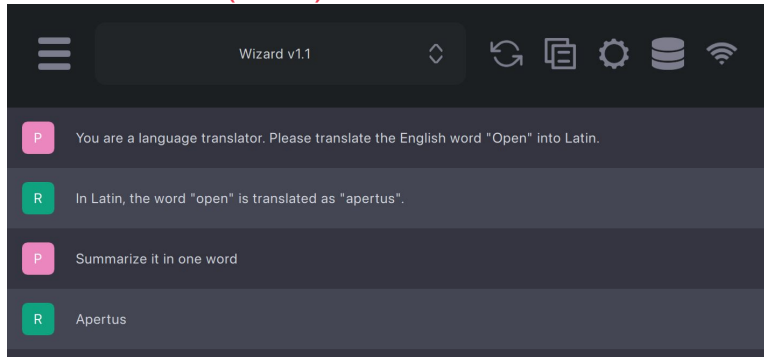
Here are some other words that use this root:

- Aperire (to open)
- Apertura (opening)
- Aprire (to open)
- Exigere (to demand or require, often used in the sense of "opening up")

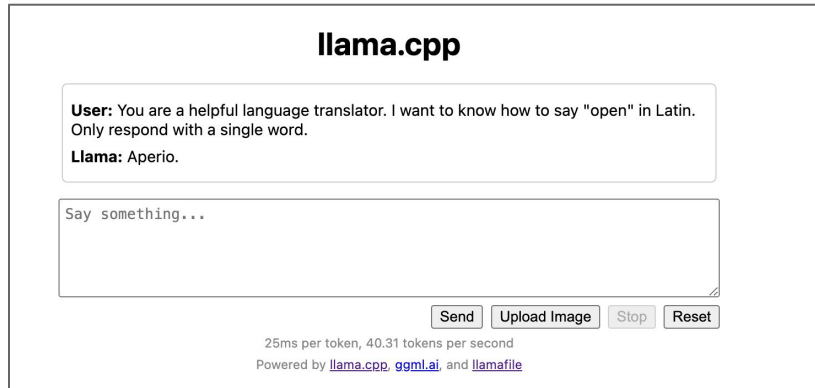
Note that while "apertus" itself does not have a direct English translation, many words derived from it do. For example, "aperture" is an opening or gap in something, such as a window or a hole in a wall; and "expose", which can mean to make something visible or accessible by removing obstructions, has a similar meaning to "aperire" or "apertus".

Model: [Llama-2 7B Chat](#) via Hugging Face

wizard v1.1 (local)

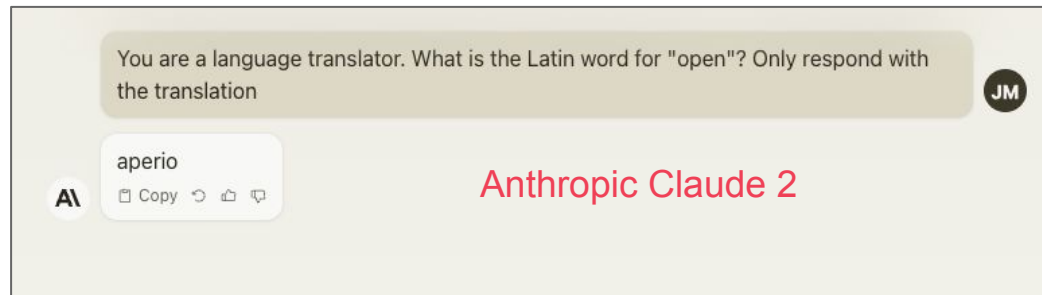


Google text-bison 32k

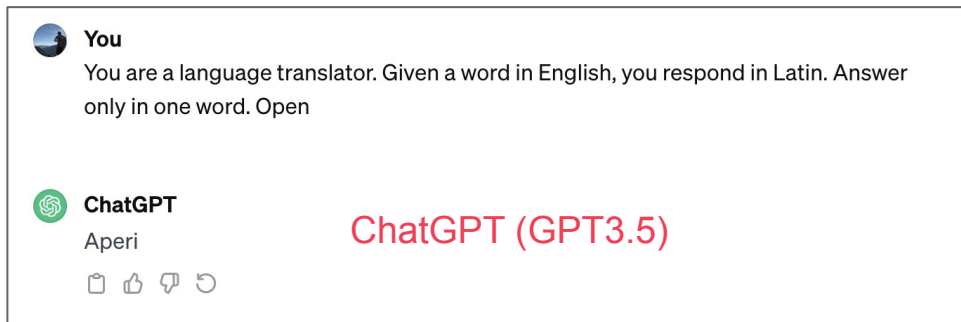


25ms per token, 40.31 tokens per second
Powered by [llama.cpp](#), [ggml.ai](#), and [llamafire](#)

Model: **Mistral 7B** running locally with [llamafire](#)



Anthropic Claude 2



ChatGPT (GPT3.5)

Create, share and test automated AI scripts

[Start Exploring](#)

Supercharge your chatbots by creating editable, reusable scripts with the best prompts possible for any model.



An open prompt engineering community by default

Avoid model or vendor lock-in by sharing and creating in the open.

[+ Create Formula](#)


Keyword Selector

 javaun

Get better keywords from unstructured text data

[Run Formula](#)

JSON Formatter

 mozgzh

Get better JSON consistency with this set of prompts

[Run Formula](#)

Gift generator

 javaun

Get an idea for a great gift for someone special

[Run Formula](#)

Ultimate Travel Packing List

 mozJava

Pack like a pro for your next trip!

[Run Formula](#)


Inspirational Thoughts

 MozillaRaj

This recipe will search for influencers about a topic and then find recent snipp...

[Run Formula](#)


Recipe Ideas

 mozgzh

Use all the ingredients in the fridge, or plan a great feast! This is the place ...

[Run Formula](#)

Alternative Ending

 mozMatt

Type in a popular movie and have the AI give you the REAL ending

[Run Formula](#)

Q & A