

AI Assistant for system monitoring using LangChain

PyMongo, 10 August 2023

Large Language Model

Natural Language Processing

backbone

Large Language Model (LLM)

Text



Vectorization

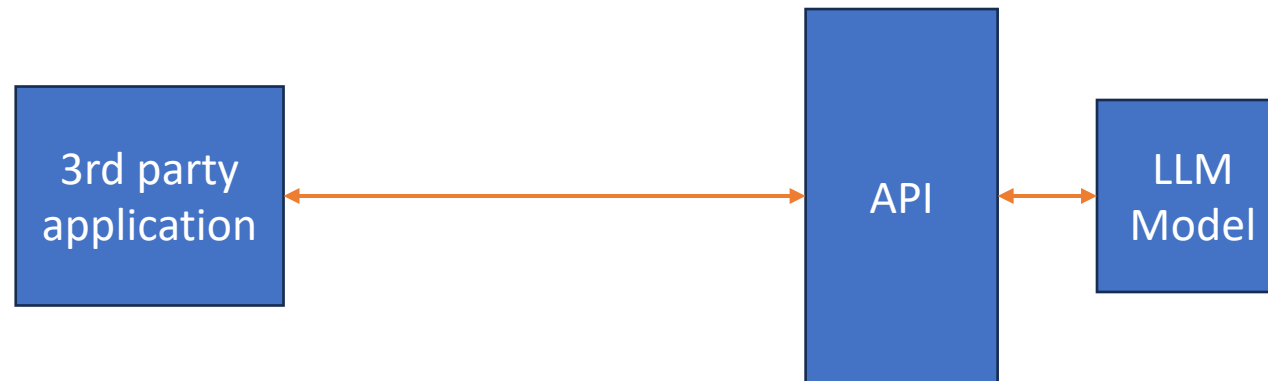
LLM

1010101101

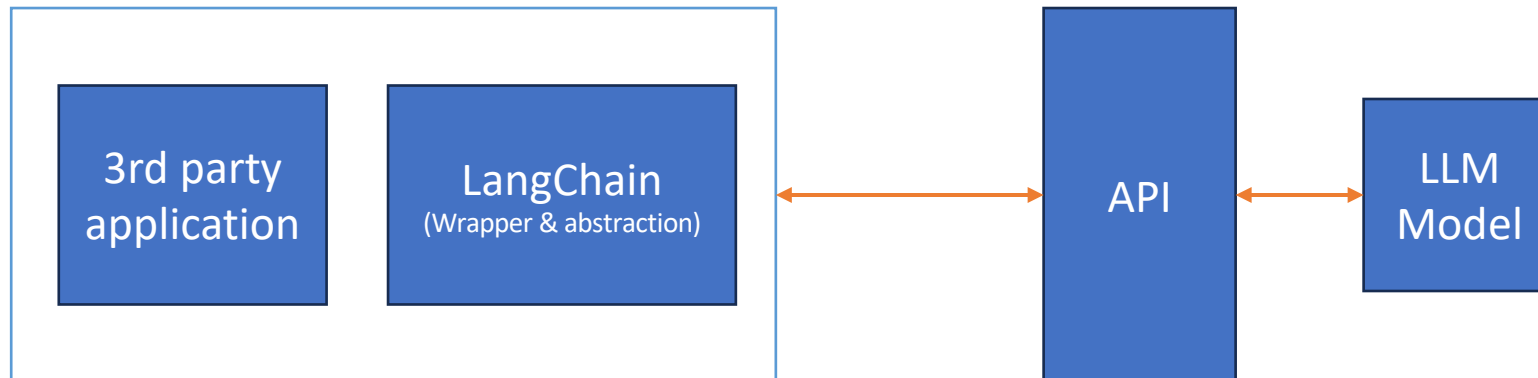
Large Language Model on Production

- GPT – OpenAI
- Galactica
- Falcon 40B
- BERT
- Others

LLM Open API



LLM Open API



Langchain

Langchain

```

▶ from langchain.llms import OpenAI

  llm = OpenAI(openai_api_key=OPENAI_API_KEY)

```

OpenAI API

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.

Example request

curl   Copy

```

1 curl https://api.openai.com/v1/models \
2   -H "Authorization: Bearer $OPENAI_API_KEY"




```

Retrieve model

GET `https://api.openai.com/v1/models/{model}`

Retrieves a model instance, providing basic information about the model such as the owner and permissioning.

Example request

text-davinci-003  curl   Copy

```

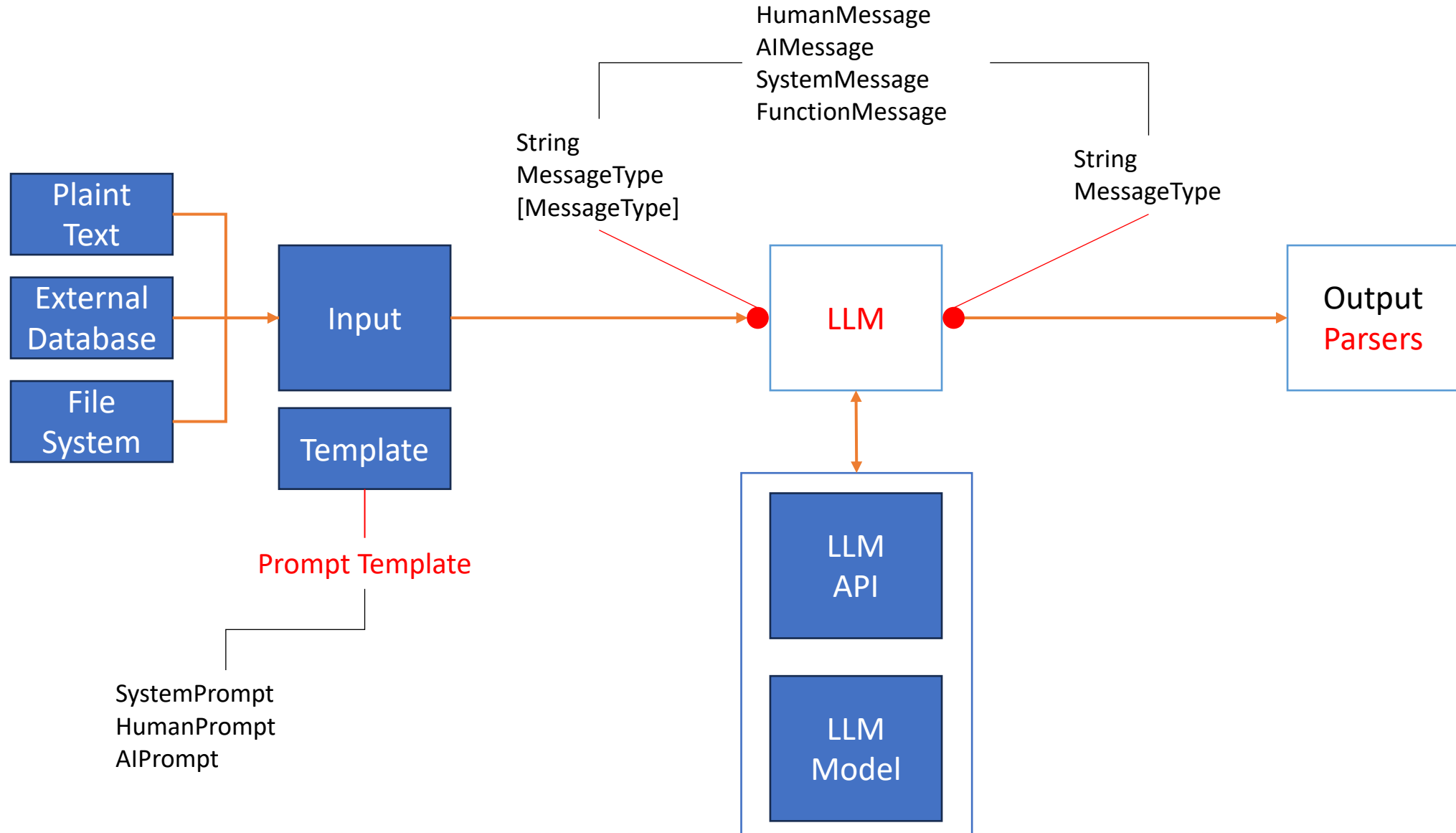
1 curl https://api.openai.com/v1/models/text-davinci-003 \
2   -H "Authorization: Bearer $OPENAI_API_KEY"

```

Langchain Component

- LLM
- Prompt Templates
- Output Parsers

Langchain Component



LangChain - LLM

- HumanMessage
- AIMessage
- SystemMessage
- FunctionMessage

LangChain – Prompt Template

- What is Prompt
- SystemPrompt
- HumanPrompt
- AIPrompt

Without prompt

“Please give me login component in HTML.”

With prompt

“Please give me login component.”

HTML should utilize semantic tag with
CSS Tailwind, and responsive must be considered.

This is prompt



Output Parsers

- BaseOutputParser
 - Overwrite parse()
- Chain everything

Data Connection

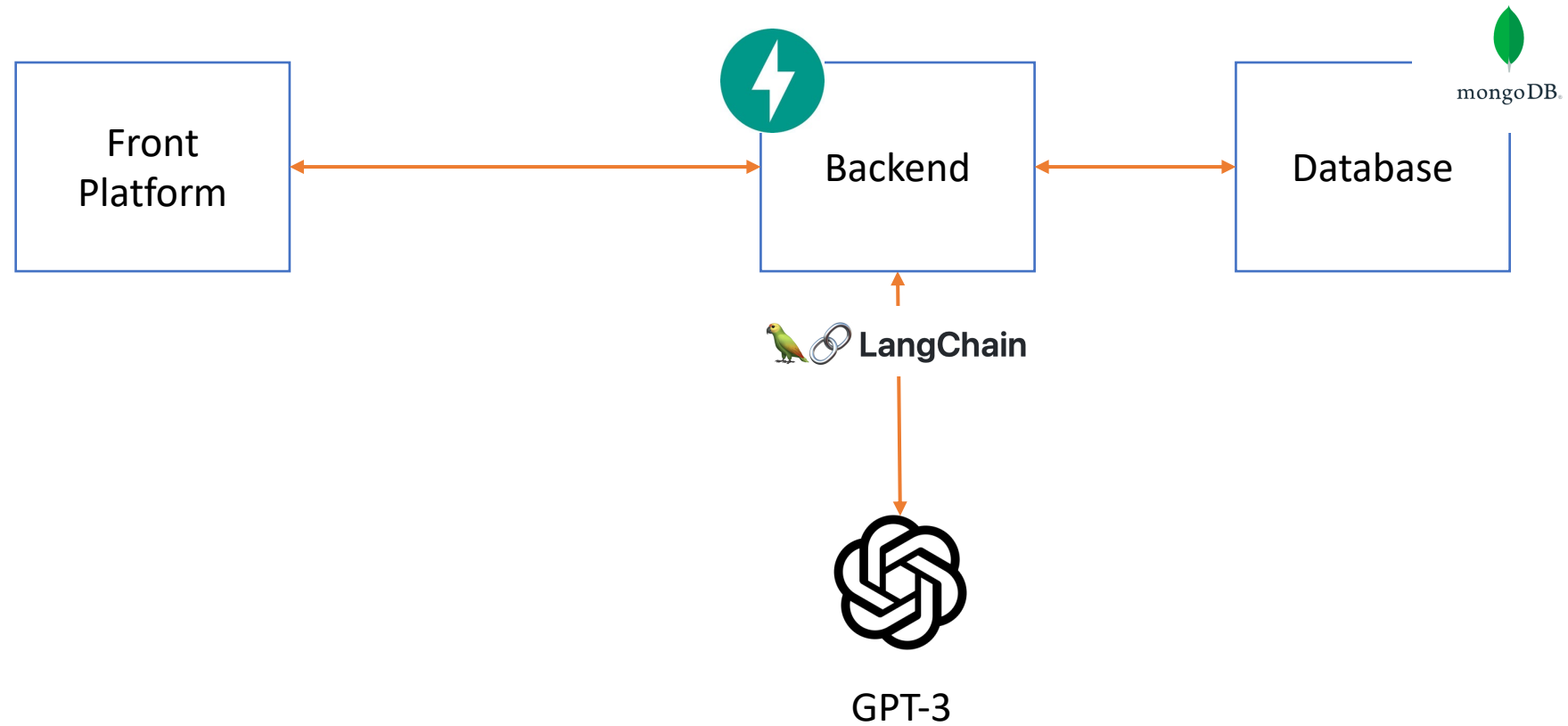
- Connect with Mongo Atlas
- Manual search

Data Connection - bonus

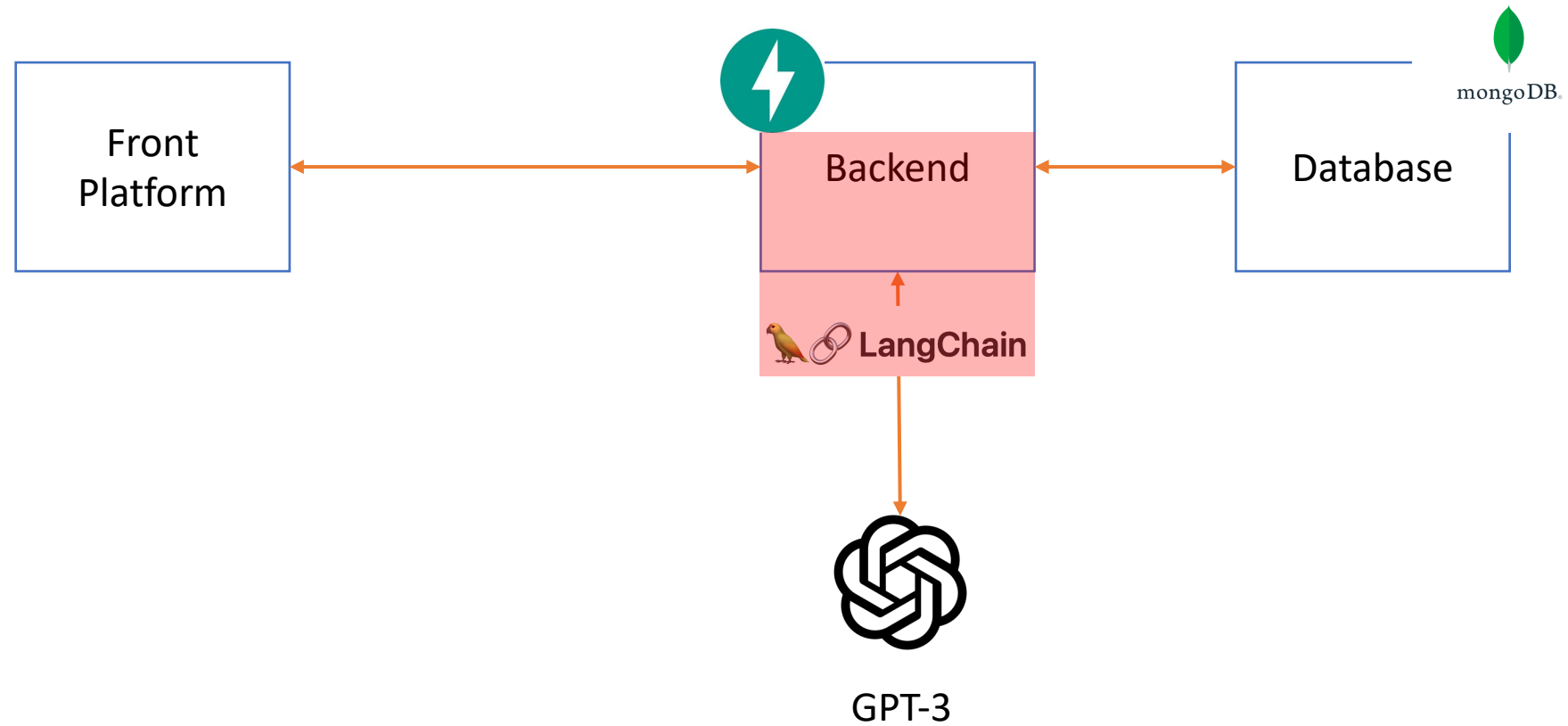
- VectorSearch with MongoDB

Demo

Demo



Demo



Walkthrough

Google Colab

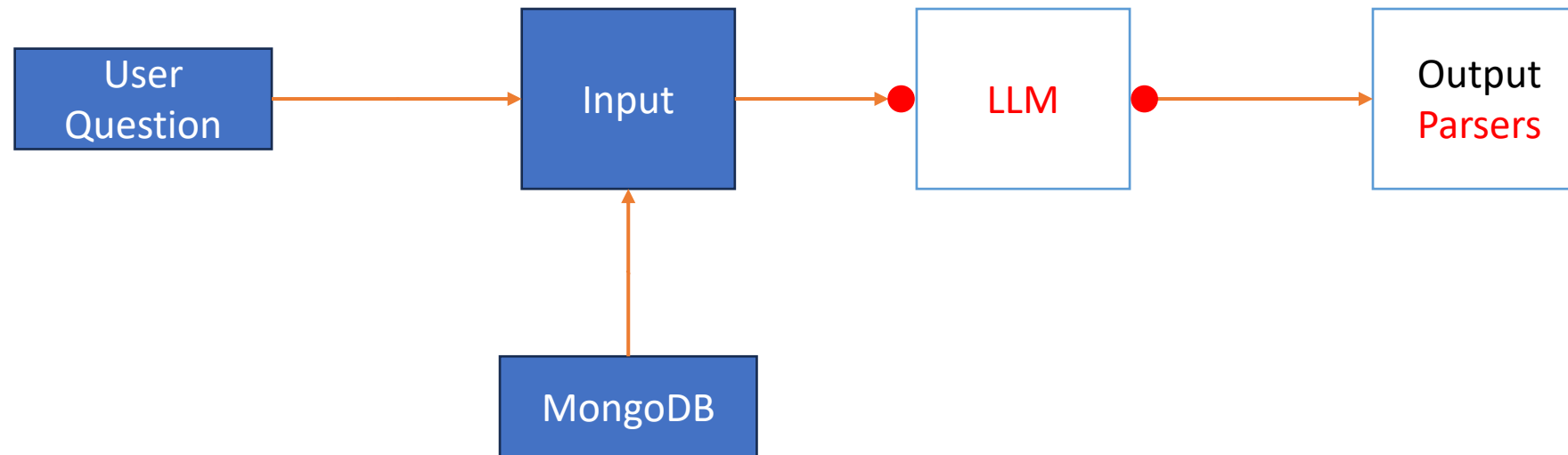


<https://shorturl.at/dlTW7>

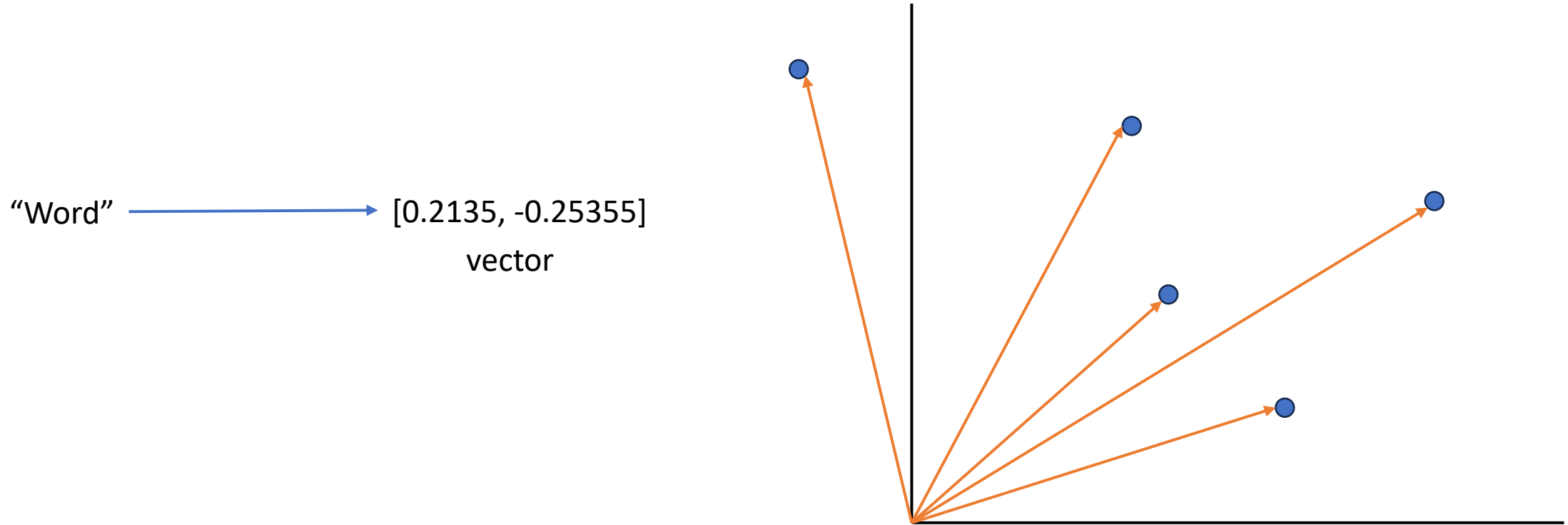
Github

<https://shorturl.at/nvwD5>

Simple assistant for system monitoring.

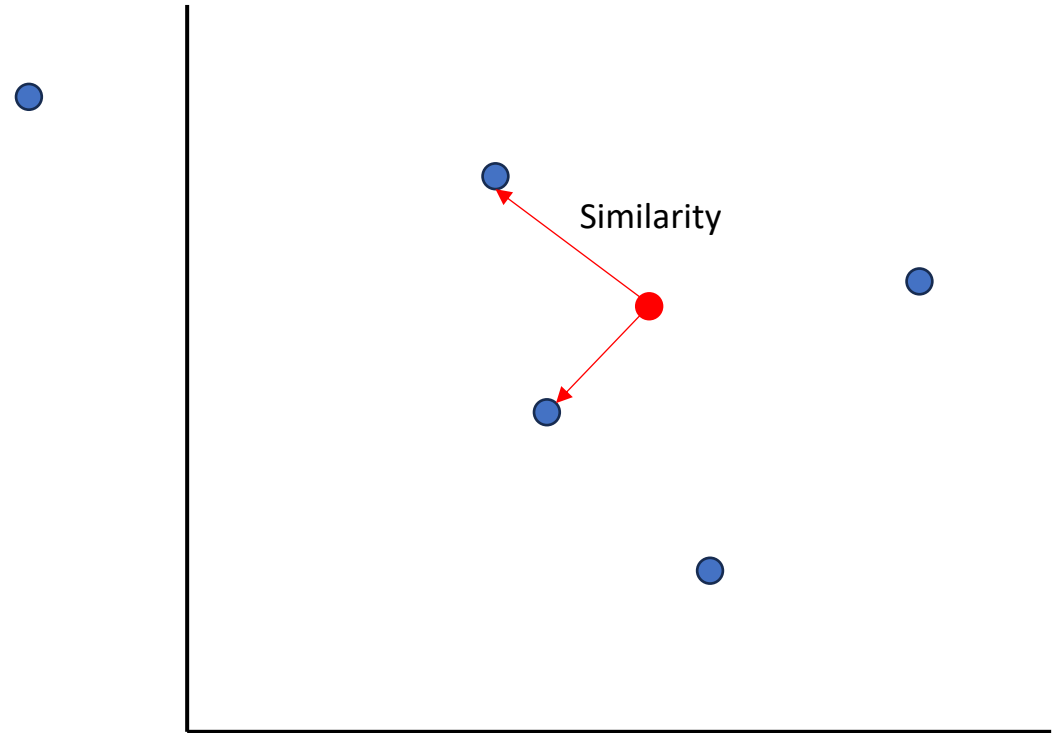


Vector Search with MongoDB



Vector Search with MongoDB

“Word”  $[0.2135, -0.25355]$
vector



Vector Search with MongoDB

```
{
  "time": "2023-08-08T10:00:00Z",
  "request_route": "/api/users",
  "request_header": { "Authorization": "Bearer token123" },
  "request_body": { "username": "user123", "email": "user@example.com" },
  "request_return_status": 201,
  "text": "Created a new user",
  "embedding": [ ...
],
```

```
{
  "time": "2023-08-08T10:00:00Z",
  "request_route": "/api/users",
  "request_header": { "Authorization": "Bearer token123" },
  "request_body": { "username": "user123", "email": "user@example.com" },
  "request_return_status": 201,
  "text": "Created a new user",
  "embedding": [
    0.00624163765340448, 0.005578401879851044, -0.009574232478285028,
    0.017887660963206882, -0.010158667187346027, 0.001917473127957561,
    -0.005910019766627762, -0.009068596971071348, -0.0193060652229098,
    -0.02467761710243589, 0.014919516404991062, 0.008825629816187343,
    -0.007873459600129312, -0.0159439199988446, -0.002237599076413812,
    0.008615495506852628, 0.02610915487288756, -0.04315627673611046,
  ]
}
```

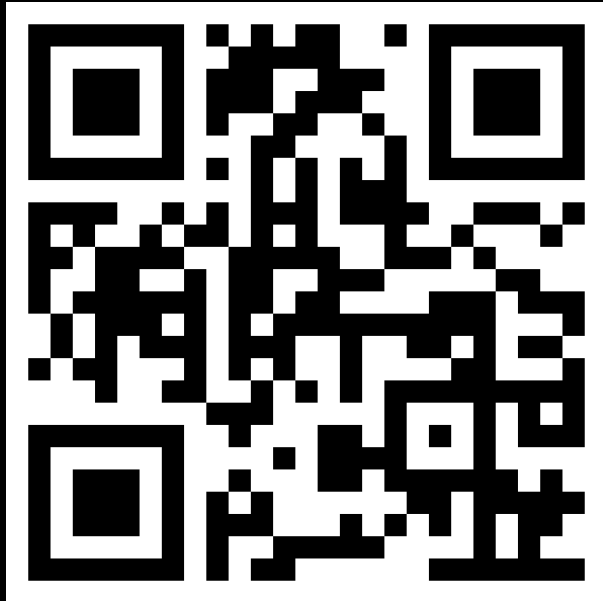
Vector Search with MongoDB

MongoDB Vector Search

```
{
  "mappings": {
    "dynamic": true,
    "fields": {
      "embedding": {
        "dimensions": 1536,
        "similarity": "cosine",
        "type": "knnVector"
      }
    }
  }
}
```

APPENDIX

Submit your topic here



th.pycon.org

PyCon Thailand 2023

Date:

- Conference Days: 15 – 16 December 2023
- Workshop Day: 14 December 2023

Venue:

- Conference: Avani Sukhumvit Bangkok
- Workshop: Microsoft (Thailand)



Submit your topic here



th.pycon.org

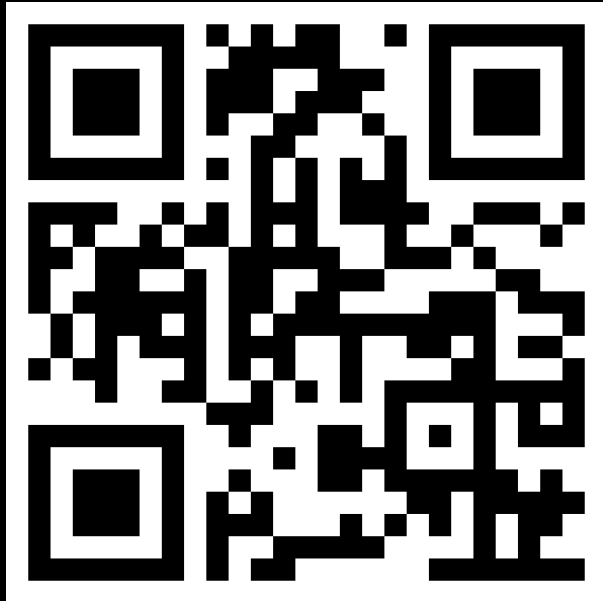
Call for Proposals is now open!

Deadline: 16 August 2023

Topic:

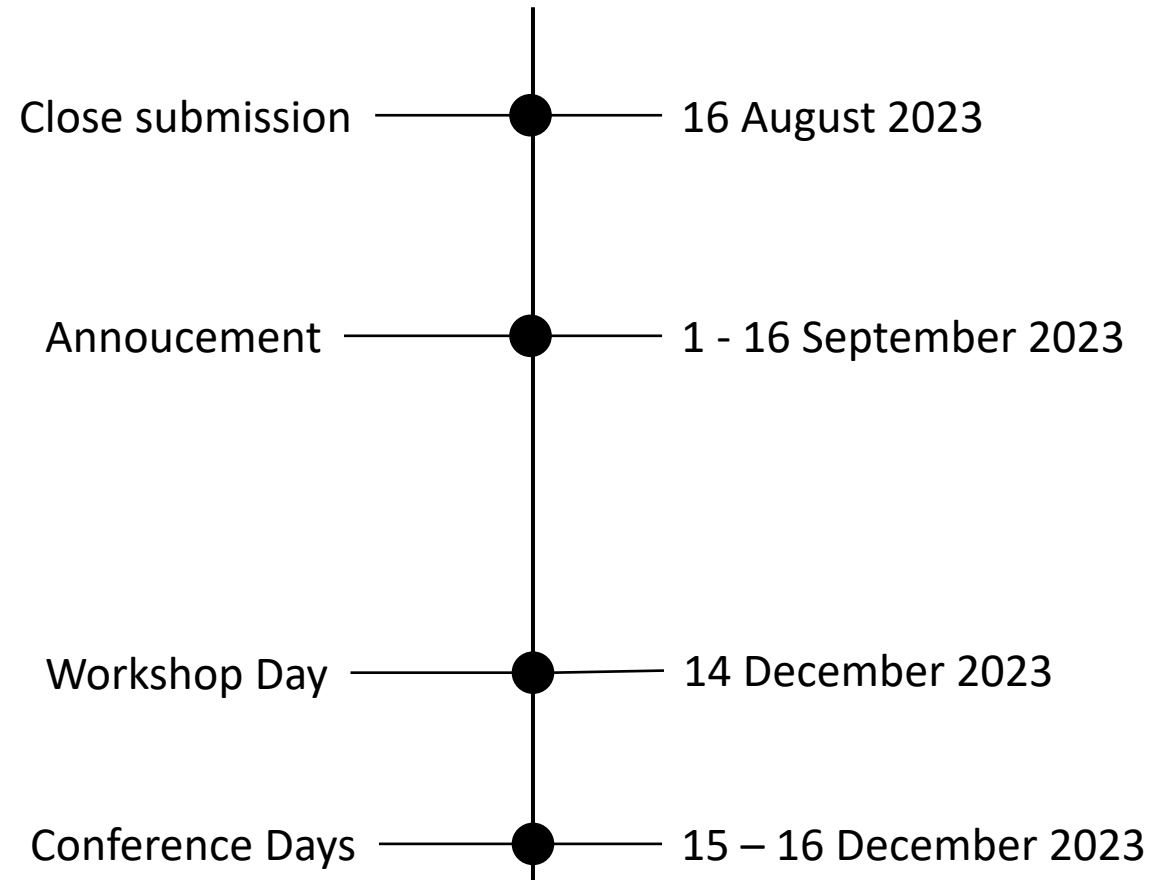
- Web Development.
- AI / Machine Learning.
- GUI / Game Development.
- Testing and Development Tool Chains.
- Network Programming.
- Other (see more in papercall's description).

Submit your topic here



th.pycon.org

Call for Proposals Timeline



Submit your topic here



th.pycon.org

Benefit for PyCon TH speaker.

- Free ticket to conference / after party.
- Internal network party of speakers (tentative).
- Contribution to community.

More information



th.pycon.org