



Join the **LA-AI** Lower Alabama Artificial Intelligence **Meetup**



Keynote Talk:
Dominik Fretz

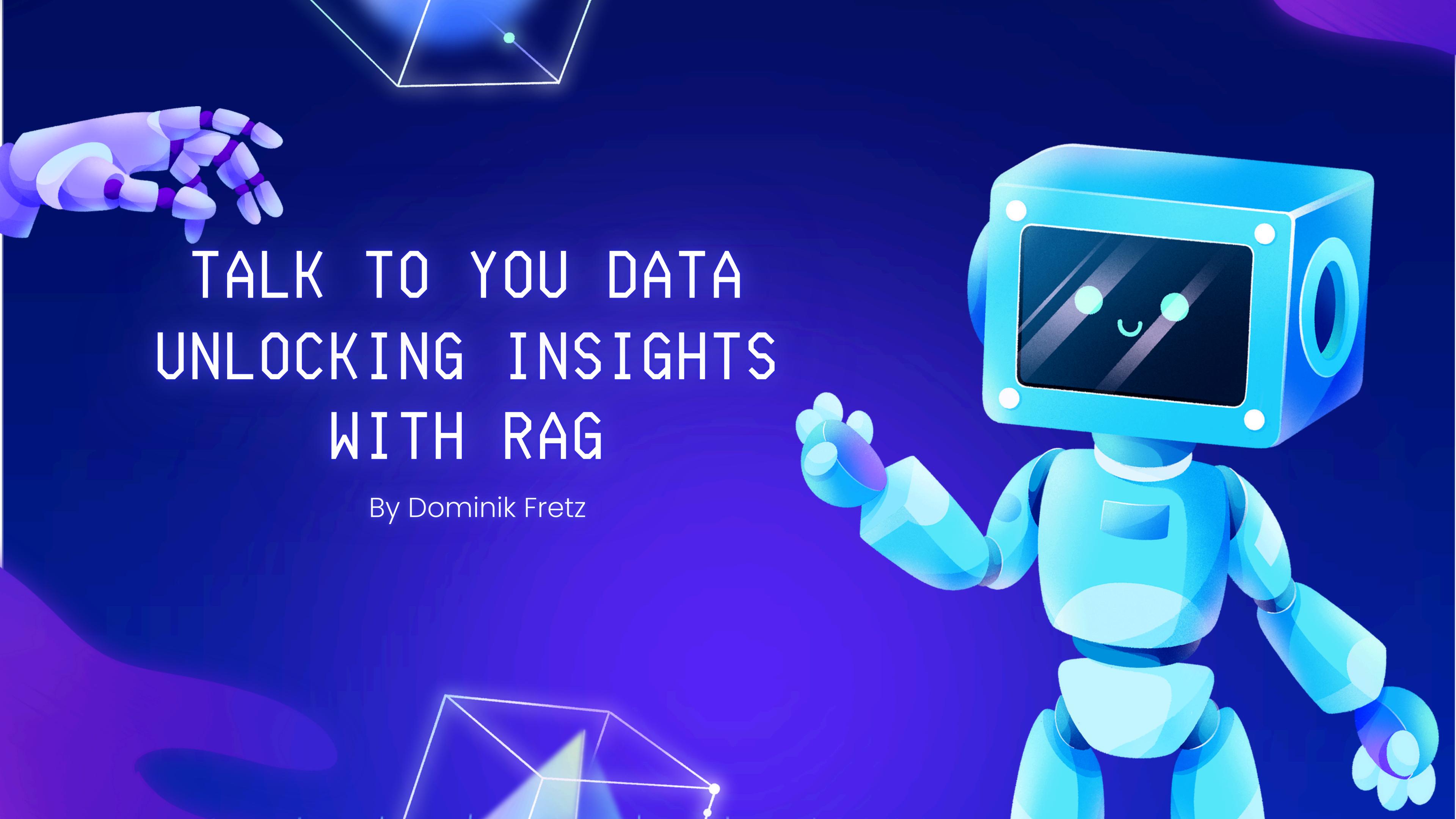
“Talk to Your Data: Unlocking
Insights with RAG”

Friday
July 19th
3:30 PM



36 No. Section St,
Fairhope, AL 36532 (2nd Fl)

-
- Engaging discussions on AI Use Cases
 - Demos of current AI applications
 - Network with fellow AI enthusiasts and industry leaders
 - Enjoy free refreshments!



TALK TO YOUR DATA UNLOCKING INSIGHTS WITH RAG

By Dominik Fretz

WE START WITH AN LLM



Trained by OpenAI,
Meta, Anthropic, etc



Trained on general
knowledge data



Able to answer your
questions

-
'The Internet'

What is the most prominent landmark in france?



The most prominent landmark in France is the Eiffel Tower, located in Paris. It is an iconic symbol of France and a major tourist attraction, recognized worldwide for its unique iron lattice structure.



LIMITATIONS



Doesn't know about
'your' data



Makes stuff up
aka Hallucinations

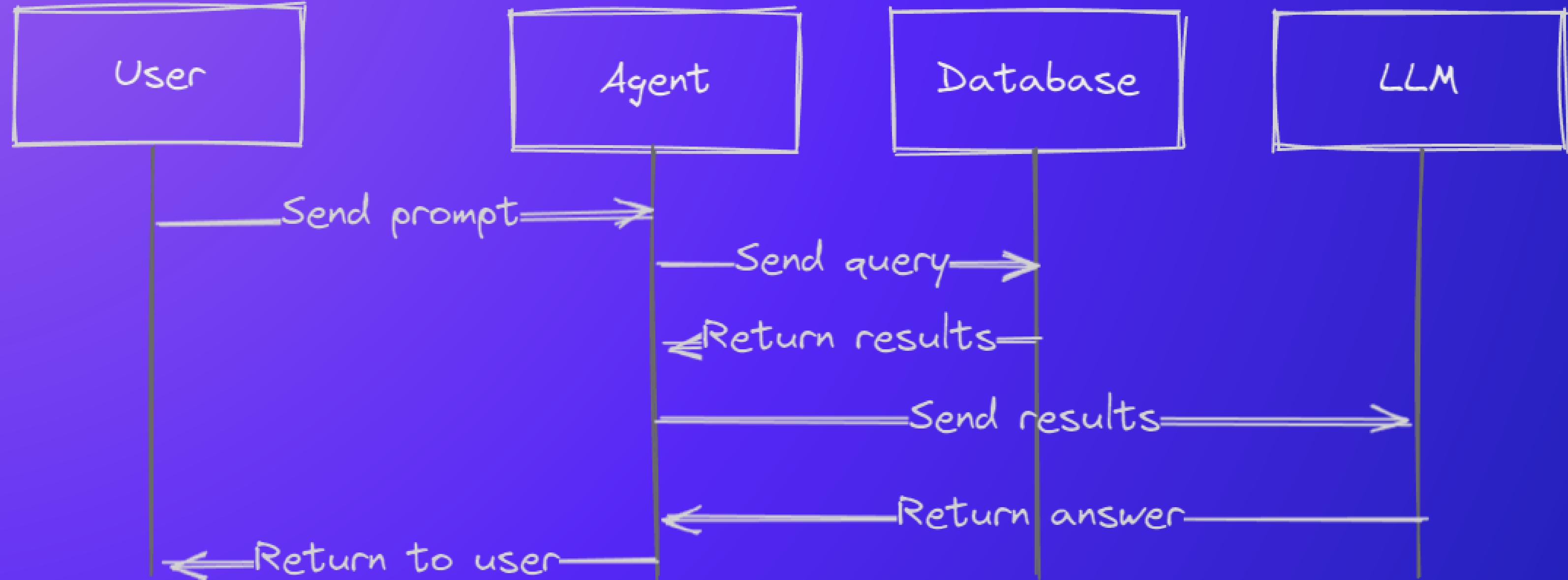


No up to date
informations

RAG TO THE RESCUE

- RAG – Retrieval Augmented Generation
- Combines data retrieval with the capabilities of LLMs
- Allows the LLM to read your data and generate responses based on that



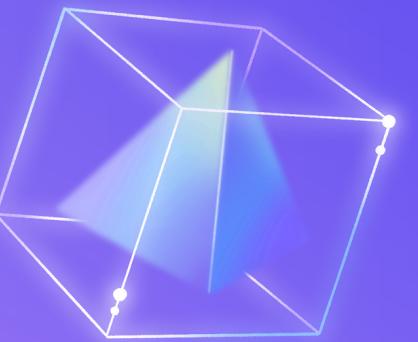


COMPONENTS



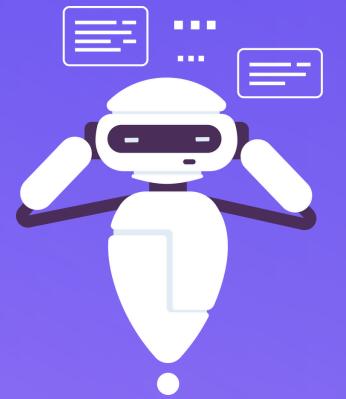
EMBEDDINGS

Text sources are converted to embeddings - a representation that makes it easier for retrieval



VECTOR DB

The place to store the data. It's not exactly a normal database



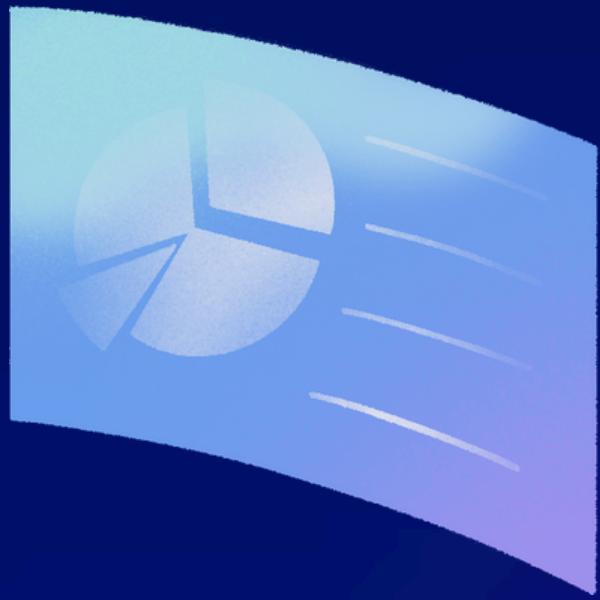
RETRIEVER

Code that is responsible for searching the database and retrieving text



LLM

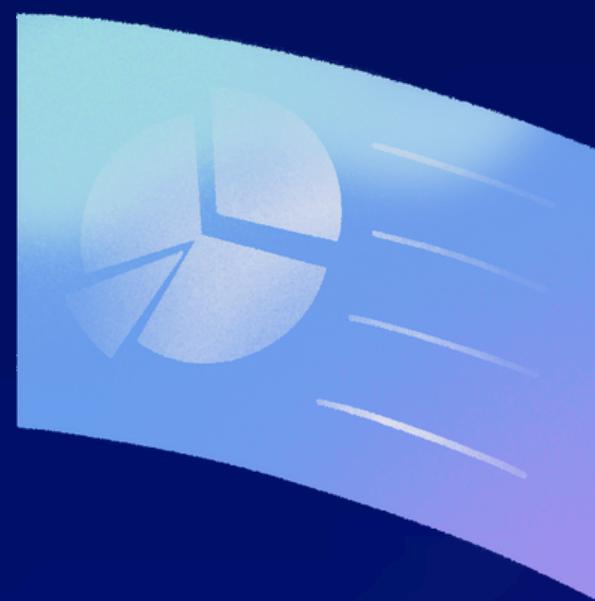
The language model that answers your questions



EMBEDDINGS

Traditionally, we would store text documents in a database and then do full text search across that.

For large data sources that can be very slow



EMBEDDINGS

Similarity search is also very complex.

eg:

You are a library and store all your book's texts in a database so that a user can find books.

The user searches for:

I LOVE FOOTBALL

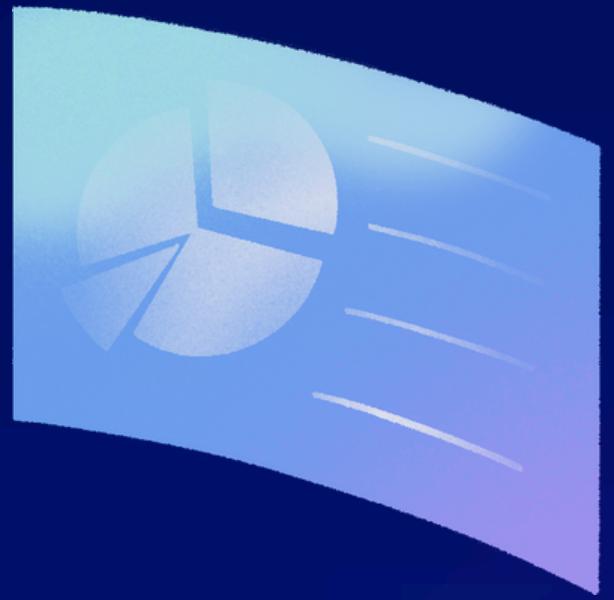


(VECTOR) EMBEDDINGS

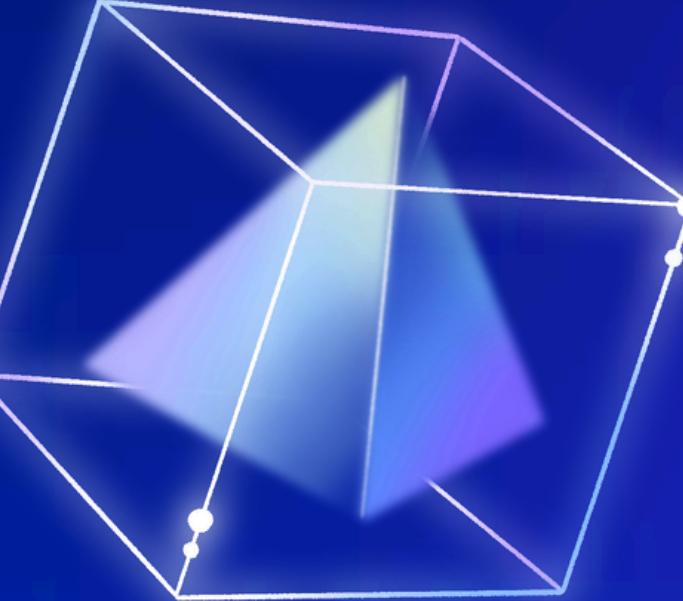
Vector Embedding are numerical values that represent information in multi dimensional space

???

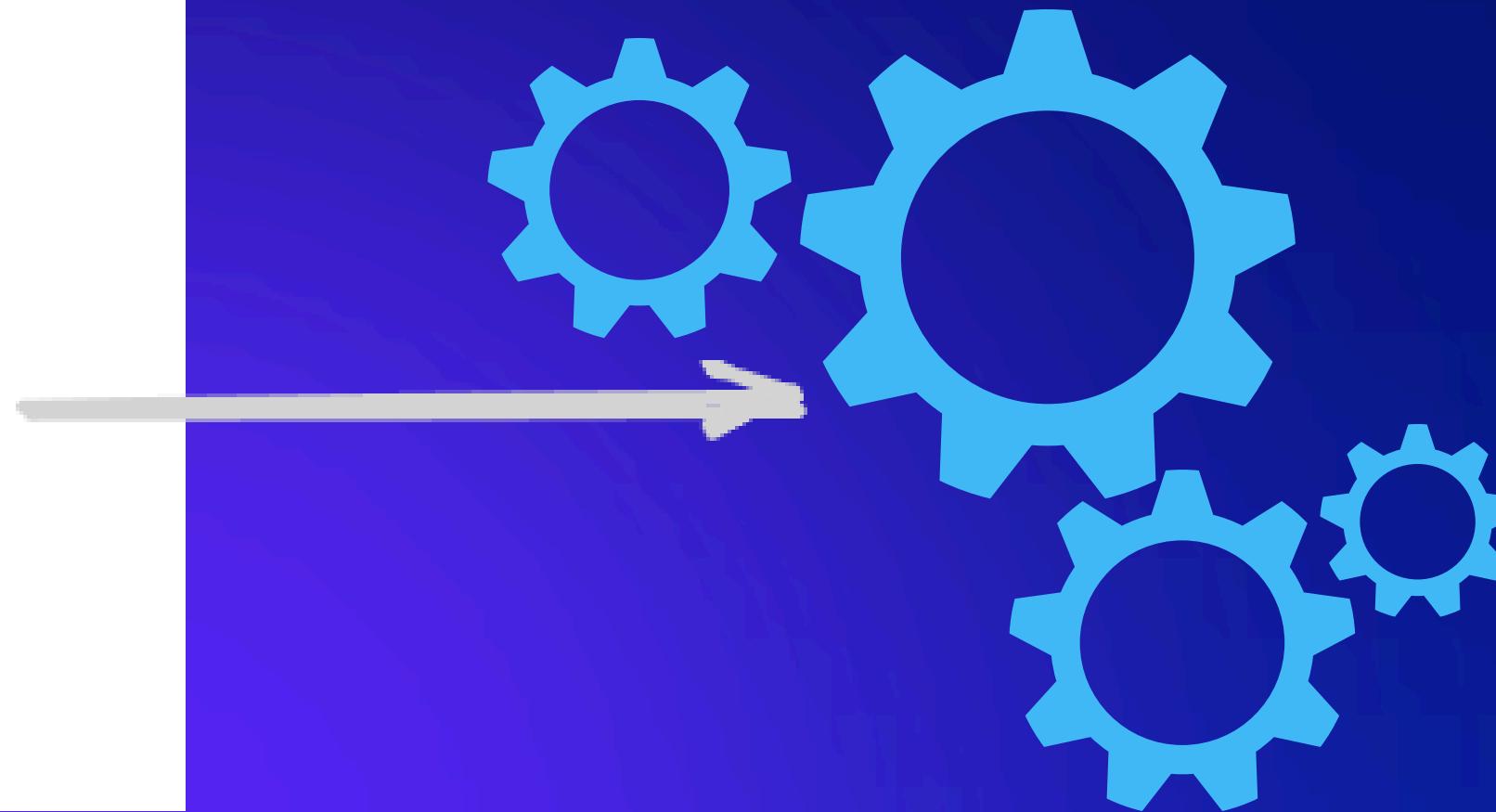


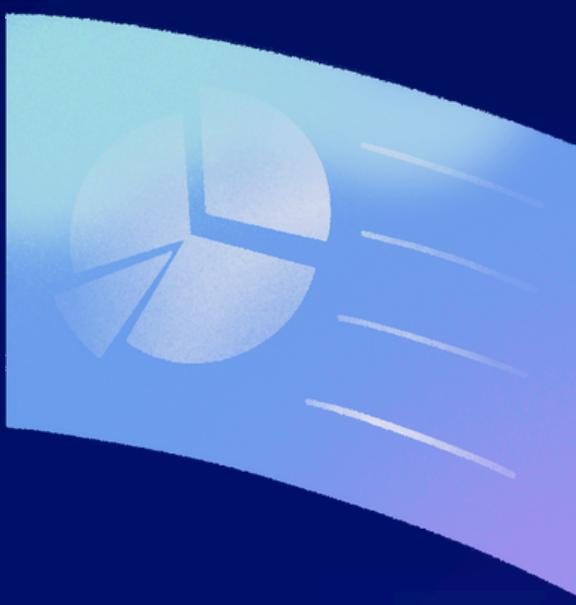


VECTOR EMBEDDINGS

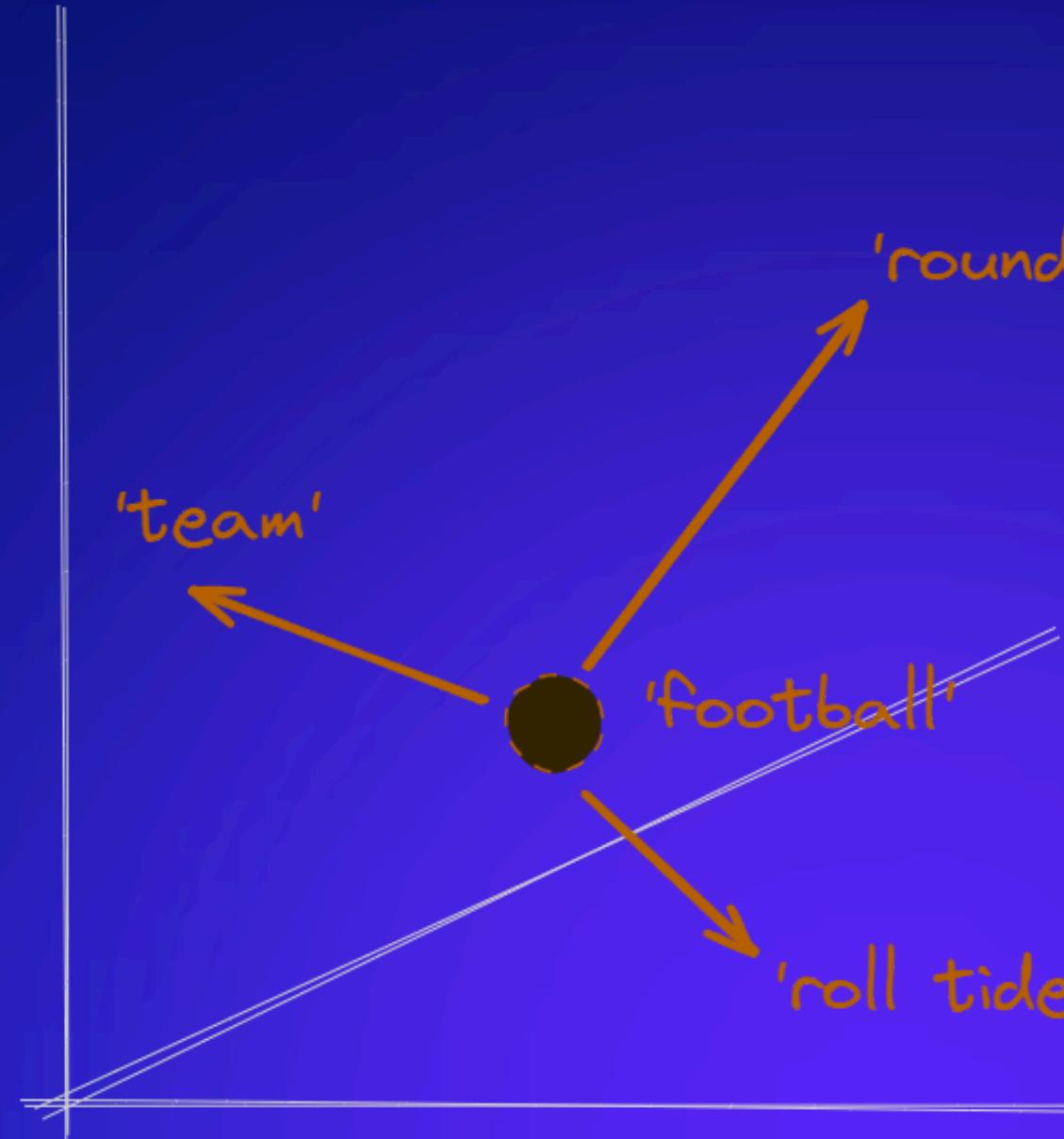


2023
OFFICIAL PLAYING RULES
OF THE
NATIONAL FOOTBALL LEAGUE





VECTOR EMBEDDINGS

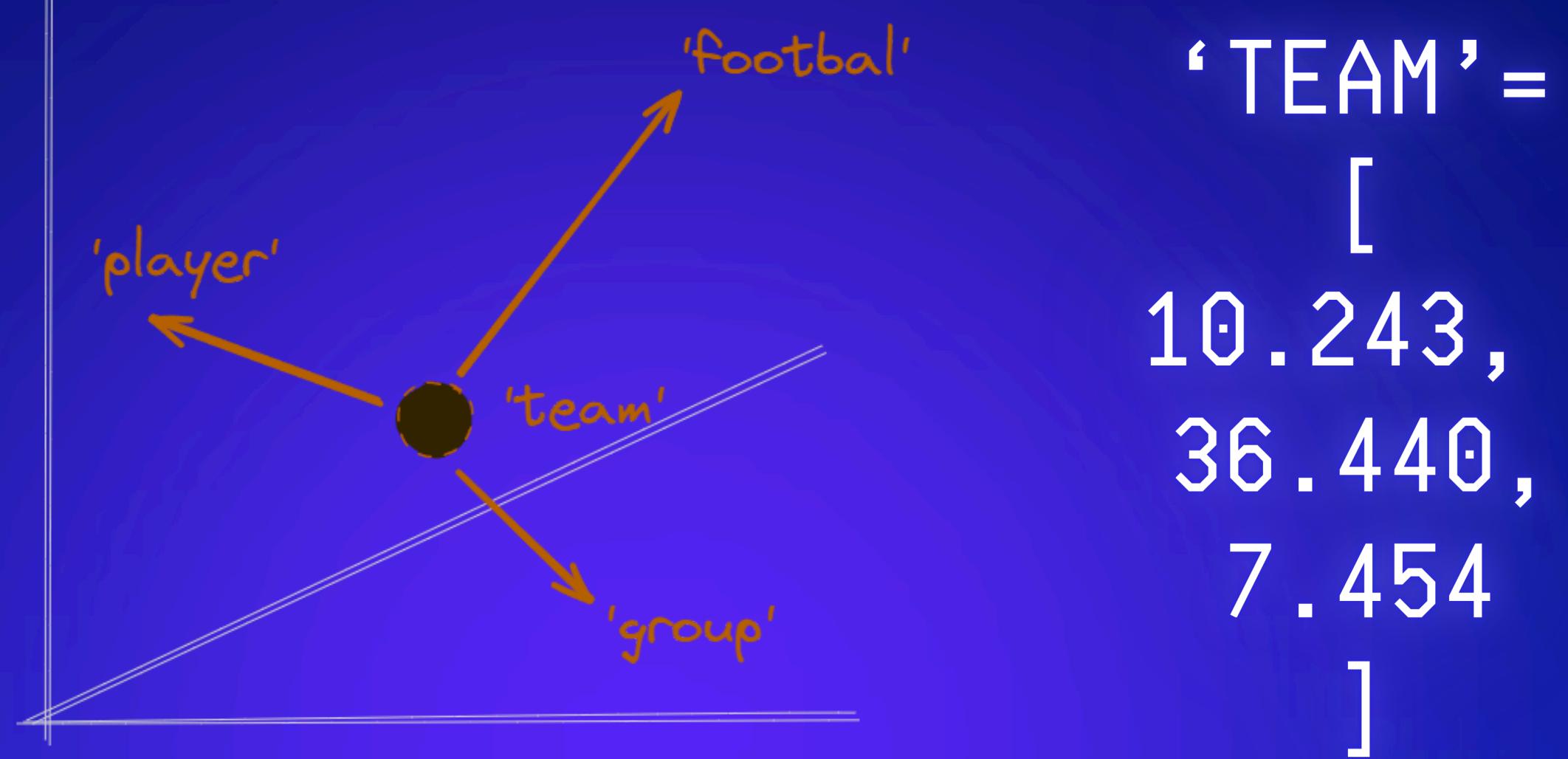


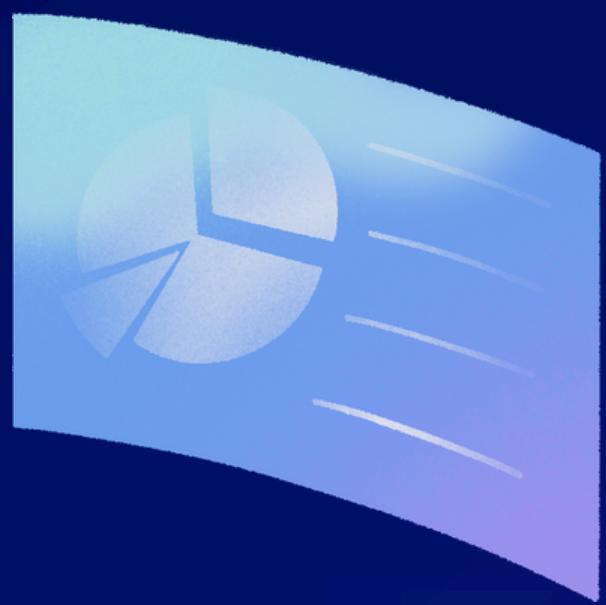
'FOOTBALL' =
[
0.421,
36.297,
99.007
]





VECTOR EMBEDDINGS





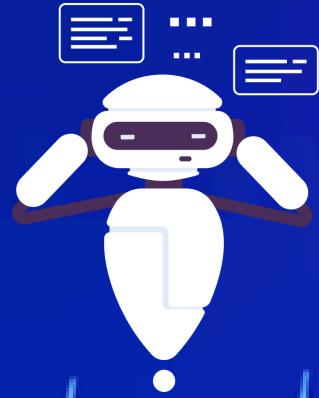
VECTOR DATABASE

‘FOOTBALL’=[0.421, 36.297, 99.007]

‘TEAM’=[10.243, 36.440, 7.454]

‘ROLL TIDE’=[97.456, 17.122, 66.486]

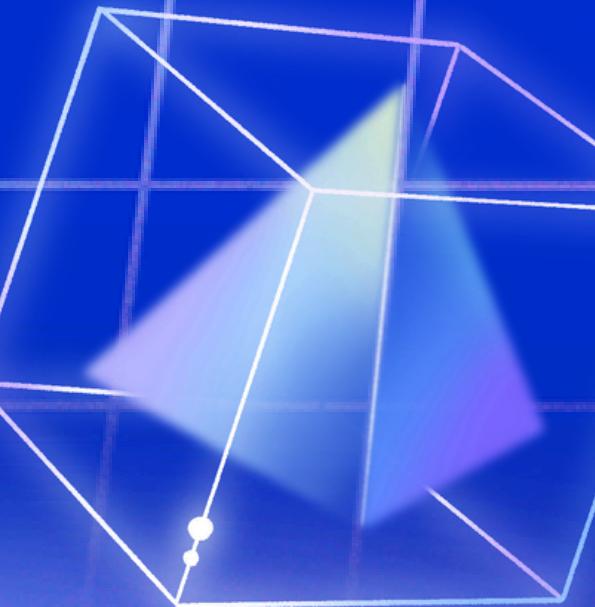


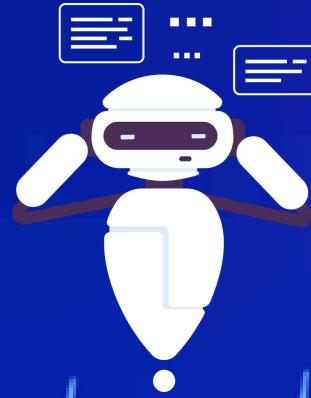


RETRIEVER

USER: I LOVE FOOTBALL

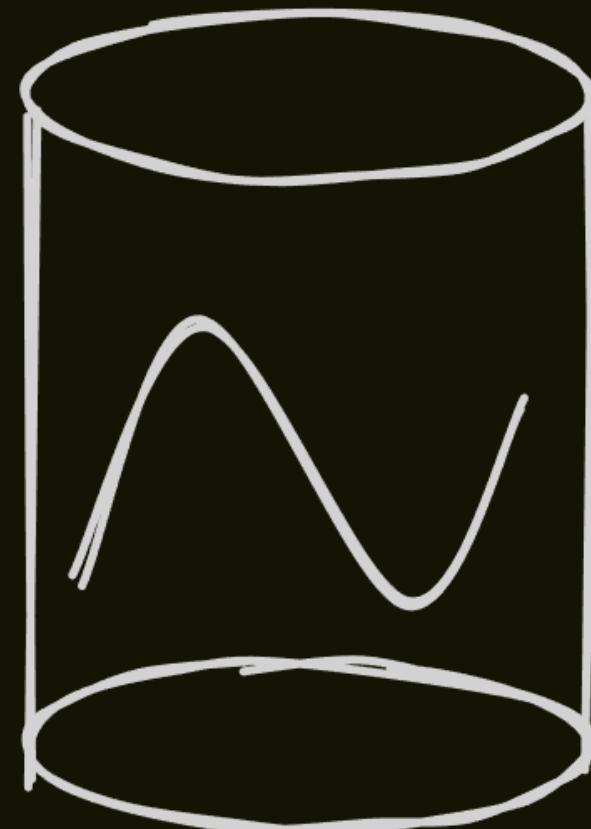
'I LOVE FOOTBALL' =
[15.111, 35.002, 77.251]



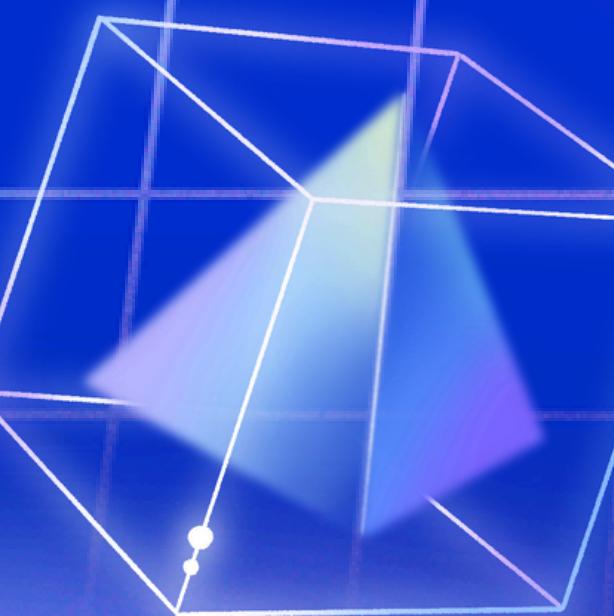


RETRIEVER

[15.111,35.002,77.251]



[10.243,36.440,7.454]
[0.421, 36.297,99.007]
[97.456,17.122,66.486]





LLM

[15.111,35.002,77.251] = 'I love football'

=

[10.243,36.440,7.454] = 'team building in football'

[0.421, 36.297,99.007] =
'OFFICIAL PLAYING RULES
OF THE NATIONAL FOOTBALL LEAGUE'

[97.456,17.122,66.486] = 'roll tide - an Alabama football story'



LLM

You are a helpful assistant that can help users find books they might be interested in.

Here are possible books in our library:

- team building in football
- OFFICIAL PLAYING RULES
OF THE NATIONAL FOOTBALL LEAGUE
- roll tide - an Alabama football story

Please answer the user with book recommendations for their question:

I love football



LLM

Here are some recommendations to satiate your love of football:

- roll tide - an Alabama football story

Description: Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla a eros ipsum. Curabitur gravida urna in felis consequat.

- OFFICIAL PLAYING RULES

OF THE NATIONAL FOOTBALL LEAGUE

Description: In convallis elit ut dui vulputate, id gravida turpis gravida. Curabitur porttitor quis sapien sed ollis.

- team building in football

Description: Pellentesque pellentesque, lectus vitae euismod condimentum, leo leo volutpat metus, eget ultricies nulla massa non dolor.



LETS
LOOK AT
SOME
CODE

THANK YOU!

FIND ME ON:

- LinkedIn
- LA-AI discord
- info@codewithpassion.com



CODE

