

Oops!

Welcome back! Our systems are currently free, and we're excited to continue to develop retrieval part of the RAG. Please relax and enjoy the content.



Homayoun Afshari



Arash Daneshvar



Hossein Khodadadi

You had hit the free plan limit for GT3.

You can use GT3 or upgrade to GT3-4o to continue this chat.
Your limit is lifted since it's January 9th.

Continue

Get plus



ChatGT3



Explore GT3s

December

Project objectives



Research question



Main task



Datasets



Embedders and VectorDBs



Similarity and performance



Next steps



Conclusion



November

Project definition



Project objectives




Upgrade plan


What can I help with?


Can you randomly provide a Persian poem from Saadi?
Include its meaning if possible. |



 Create Image

 Code

 Summarize

 Get advice

More





ChatGT3 ▾



ChatGT3



Explore GT3s

Today

Introduction: RAG & poetry ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...

Similarity and performance ...

Next steps ...

Conclusion ...

Can you randomly provide a Persian poem from Saadi? Include its meaning if possible.

Now you can view two options: text-based and visualized responses.



Response 1

Here is a Persian poem from **Saadi**:

”به هوش بودم از اول که دل به کس نسپارم
شما یل تو بدیدم نه صبر ماند و نه هوشم“

The translation is as follows:

“I was mindful from the beginning, determined not
to give my heart to anyone. But when I saw you, I
lost all my strength and sense.”

Enjoy the harmonious blend of nature and wisdom!



Response 2

There's also a **Visualization**:



Message ChatGT3



Upgrade plan

ChatGT3 can make mistakes. Check important info.





Today

Introduction: RAG & poetry ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...

Similarity and performance ...

Next steps ...

Conclusion ...



GT3s

Discover and create custom versions of ChatGPT that combine instructions, extra knowledge, and any combination of skills.



Search GT3s

Top Picks

Writing

Productivity

Research & Analysis

Education

Lifestyle

Programming

Featured

Curated top picks from this week



GT3's Retrieval System

Enhanced Retrieval, Delivering Smarter Responses.

By gt3.ai



Code Copilot

Code smarter, build faster - with the expertise of 10 programmers.

By promptspellsmith.com





ChatGT3



Explore GT3s



GT3's Retrieval System

Available Information



English Literature.pdf
PDF



Italian Literature.html
HTML



Persian Literature.pdf
PDF



Indian Literature.docx
Microsoft Word



French Literature.txt
Plain Text



German Literature.pdf
PDF



Russian Literature.pdf
PDF



Dutch Literature.docx



Upgrade plan



GT3's Retrieval System

Enhanced Retrieval, Delivering Smarter Responses

By gt3.ai

Find the latest
research about AI

I'll provide a research
paper link; Please ...

I will upload a PDF
paper; Use critical
skills to read it

Type "LS" to list my
built-in critical
reading skills

Retrieve related poems to this concept from different languages:
"Love in the first sight"





ChatGT3



Explore GT3s



GT3's Retrieval System

Available Information



English Literature.pdf
PDF



Italian Literature.html
HTML



Persian Literature.pdf
PDF



Indian Literature.docx
Microsoft Word



French Literature.txt
Plain Text



German Literature.pdf
PDF



Russian Literature.pdf
PDF



Dutch Literature.docx
Microsoft Word



Upgrade plan

How does the attached PDF describe methods to measure similarity? What approaches are highlighted?



Italian Literature.html
HTML



English Literature.pdf
PDF

O mi senti' svegliar dentro a lo core
Un spirito amoroso che dormia:
E poi vidi venir da lungi Amore
Allegro sì, che appena il conoscia.

Your love was a gift, from our first meeting
the way you looked at me, made me smile.
Love songs played on my heart, repeating
with you my love, I'd go that extra mile!



French Literature.txt
Plain Text



Indian Literature.docx
Microsoft Word

Un éclair... puis la nuit! — Fugitive beauté
Dont le regard m'a fait soudainement
renaître,
Ne te verrai-je plus que dans l'éternité?

हुआ दिल तुम्हारा, पहली नजर से
ना खुद पे है काबू, पहली नजर से

Message ChatGT3





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...

Similarity and performance ...

Next steps ...

Conclusion ...



Upgrade plan

What can I help with?



ADSP - P9 - RAG MARCO.pdf
PDF

What problem does this project aim to solve based on the attached PDF? How is it defined? |



Create Image



Code



Summarize



Get advice

More





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...

Similarity and performance ...

Next steps ...



Upgrade plan



ADSP - P9 - RAG MARCO.pdf
PDF

What problem does this project aim to solve based on the attached PDF? How is it defined?



Let's dive into optimizing information quality in a RAG system, with a spotlight on retrieval and response generation. When it comes down to it, the quality of information is everything.

Based on insights from the PDF, here's a breakdown of the main objectives:

- **Objective 1:** Refine the retrieval system.
- **Objective 2:** Improve the contextual quality of information fed into the response generation stage.

And here's a potential **research question** to dig deeper:

“How optimizing the retrieval stage in a RAG system ensures high-quality and contextually accurate responses?”

Also, for this project, the **value proposition** could be:

“Enhanced Retrieval, Delivering Smarter Responses”

Aligned with the following user persona:

“Michael Foster, the AI software manager”

This approach keeps the focus sharp - enhancing accurate retrieval and delivering relevant.



Message ChatGT3





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...

Similarity and performance ...

Next steps ...



Upgrade plan


What can I help with?




ADSP - P9 - RAG MARCO.pdf
PDF


What key terms are essential to understanding this topic, as explained in the attached PDF? How are they defined? |



 Create Image

 Code

 Summarize

 Get advice

More





ChatGT3



ChatGT3



Explore GT3s



GT3's Retrieval System

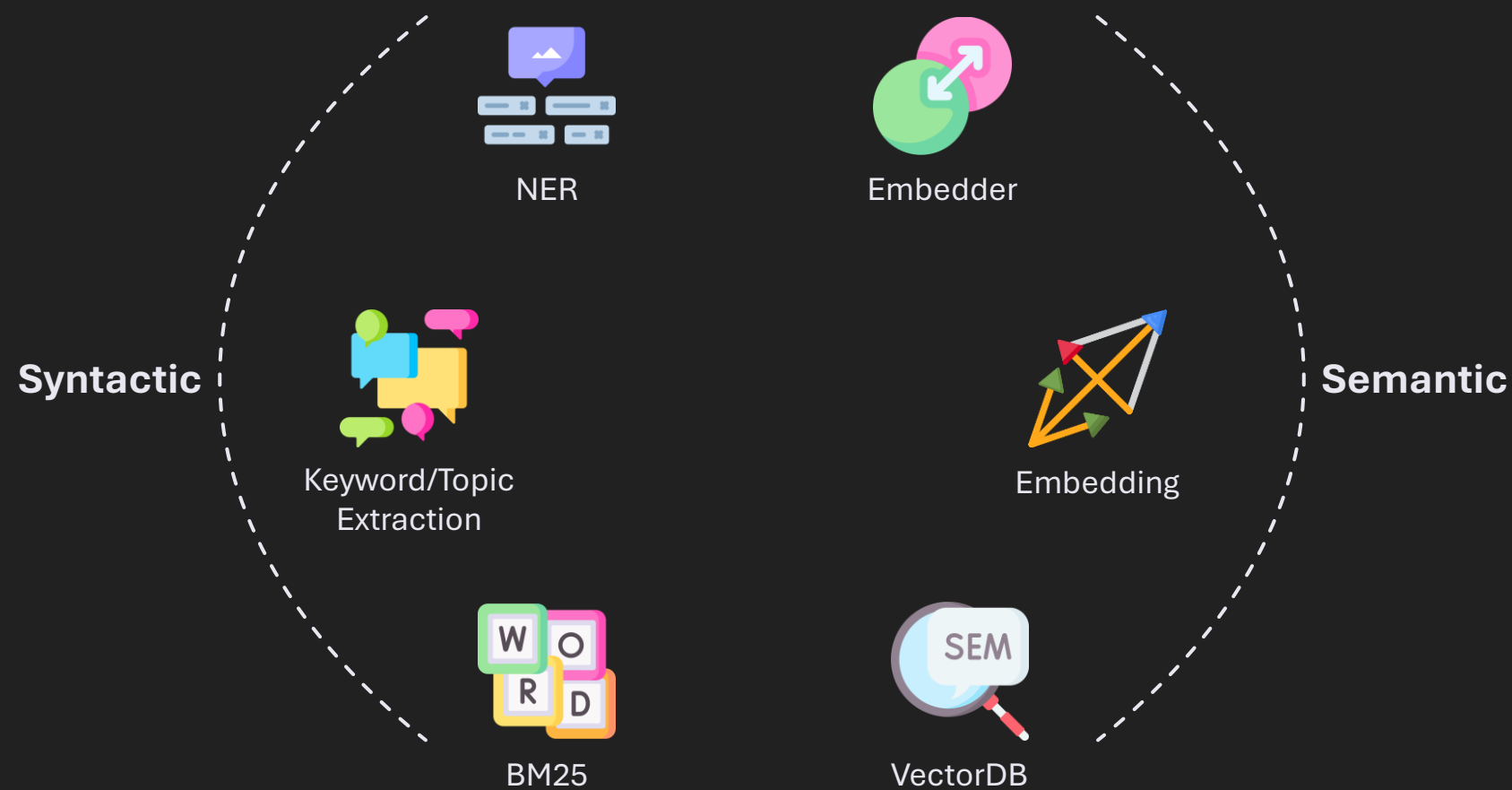


ADSP - P9 - RAG MARCO.pdf
PDF

What key terms are essential to understanding this topic, as explained in the attached PDF? How are they defined?



The terminology of this project is summarized in the following illustration:



Message ChatGT3



Upgrade plan

ChatGT3 can make mistakes. Check important info.





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...

Similarity and performance ...



Upgrade plan

What can I help with?



ADSP - P9 - RAG MARCO.pdf
PDF

How does the attached PDF describe methods to measure similarity? What approaches are highlighted? |



Create Image



Code



Summarize



Get advice

More





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...



Upgrade plan

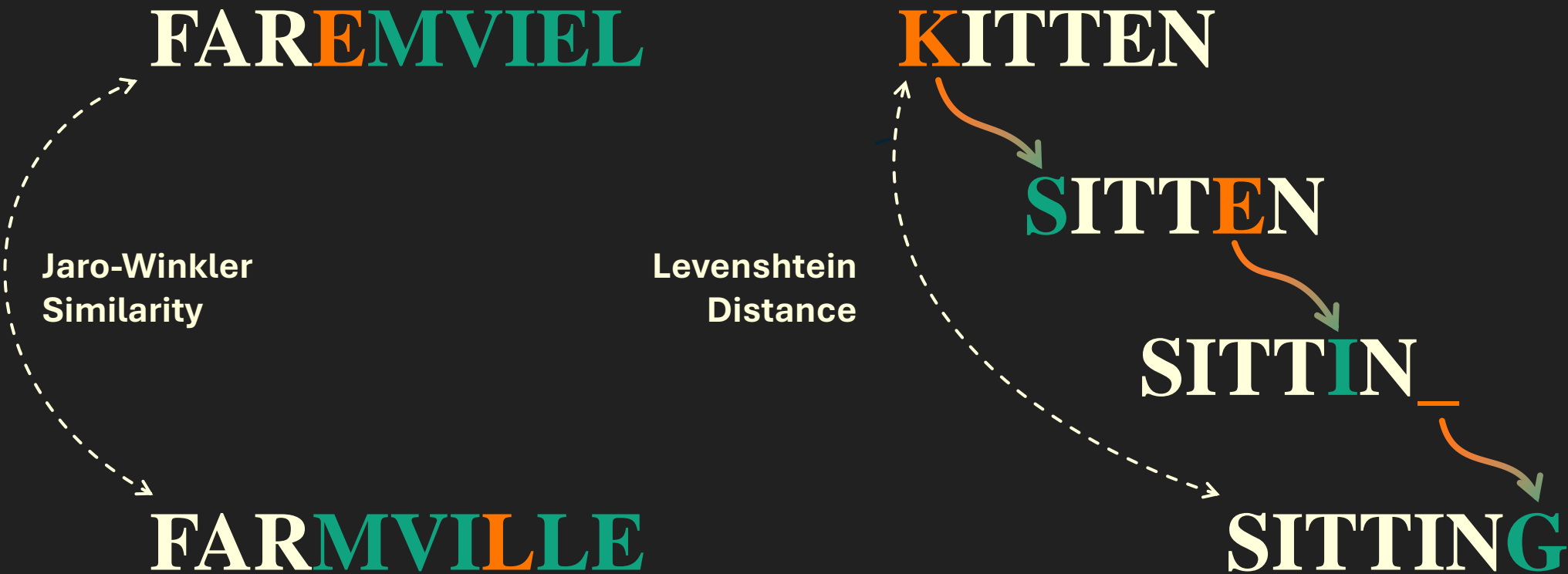


ADSP - P9 - RAG MARCO.pdf
PDF

How does the attached PDF describe methods to measure similarity? What approaches are highlighted?



Firstly, to measure the **Syntactic Similarity**, we can use metrics such as **Jaro-Winkler Similarity** and **Levenshtein Distance**.



Message ChatGT3





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...



Upgrade plan

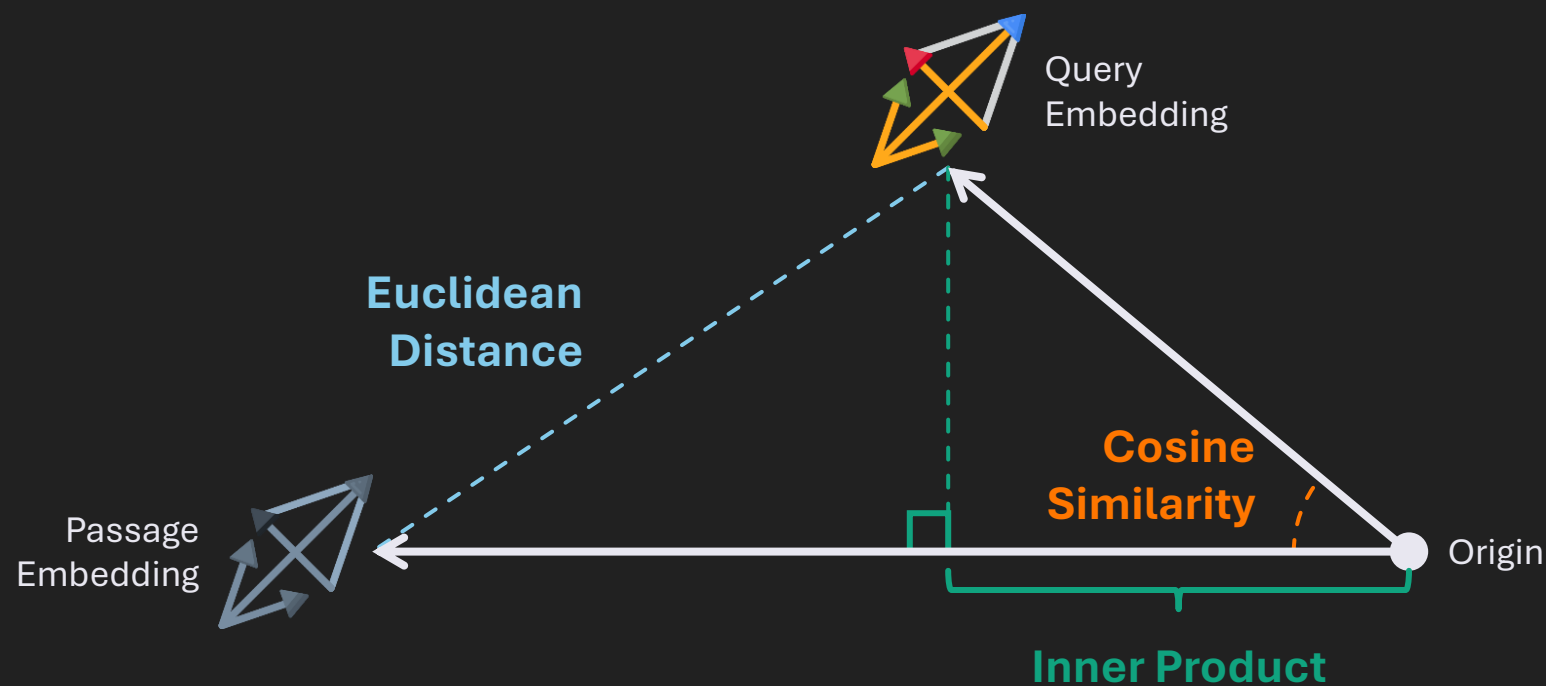


ADSP - P9 - RAG MARCO.pdf
PDF

How does the attached PDF describe methods to measure similarity? What approaches are highlighted?



Secondly, to measure the **Semantic Similarity**, we rely on metrics like **L2 distance, Inner product, and Cosine Similarity**, which determine the closeness between vector representations.



Message ChatGT3





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Embedders and VectorDBs ...



Upgrade plan


What can I help with?




ADSP - P9 - RAG MARCO.pdf
PDF


What performance metrics are discussed in the attached PDF? How do they indicate success? |



 Create Image

 Code

 Summarize

 Get advice

More





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...

Upgrade plan

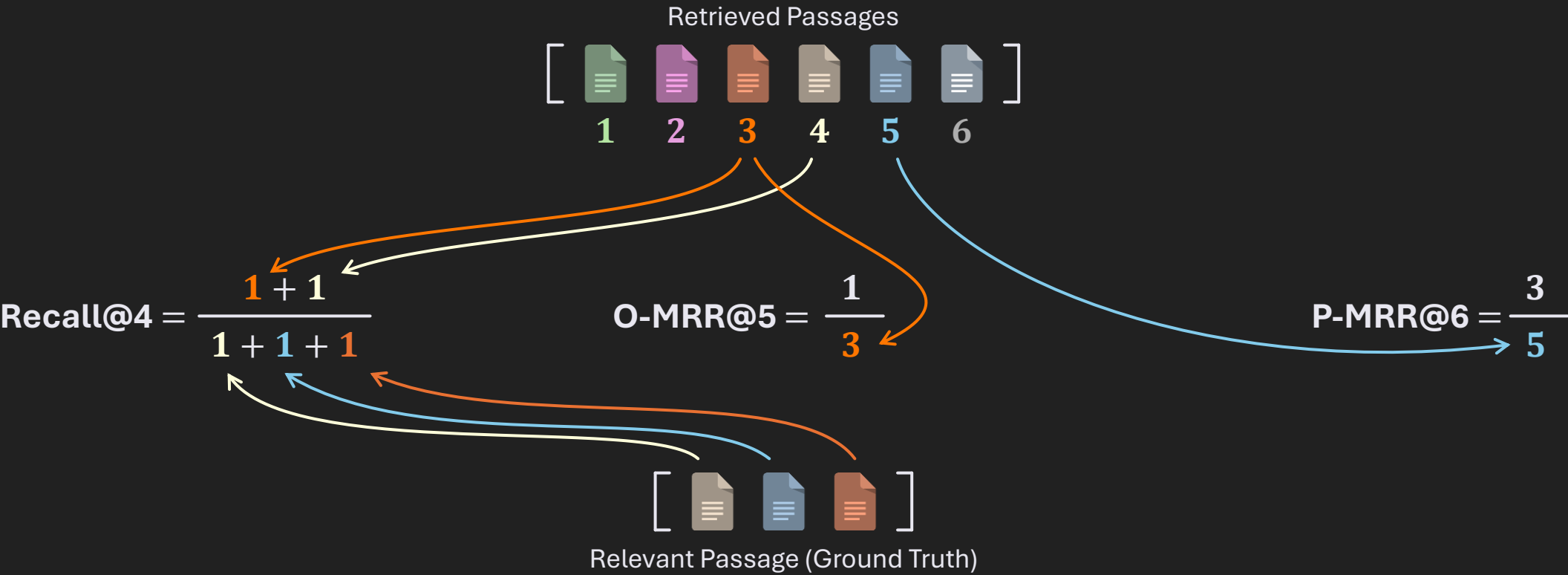


ADSP - P9 - RAG MARCO.pdf
PDF

What performance metrics are discussed in the attached PDF? How do they indicate success?



For evaluating system performance, we use metrics such as **Recall@k** and **MRR@K**, ensuring relevance and ranking quality in retrieved results. Below, you'll find a comparison of these similarity measures and performance metrics.



Message ChatGT3





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

December

Project objectives ...

Research question ...

Main task ...

Datasets ...



Upgrade plan

What can I help with?



ADSP - P9 - RAG MARCO.pdf
PDF

What core ideas of the methodology are outlined in the attached PDF? Why were they chosen? |



Create Image



Code



Summarize



Get advice

More





ChatGT3 ▾



ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

December

Project objectives ...

Research question ...

Main task ...



Upgrade plan

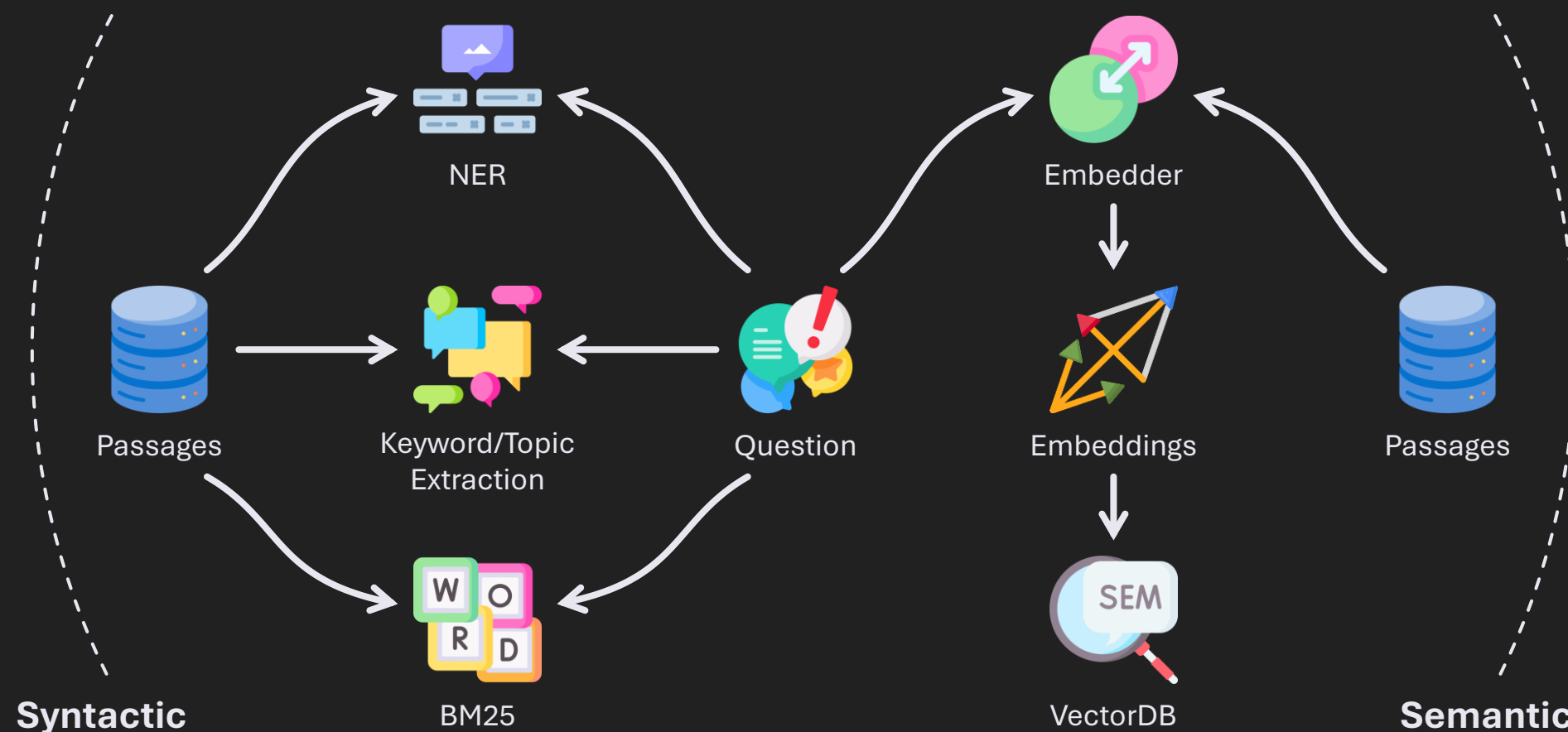


ADSP - P9 - RAG MARCO.pdf
PDF

What core ideas of the methodology are outlined in the attached PDF? Why were they chosen?



You can find the proposed ideas in the figure below:



Message ChatGT3



ChatGT3 can make mistakes. Check important info.





ChatGT3



ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

December

Project objectives ...

Research question ...

Main task ...



Upgrade plan

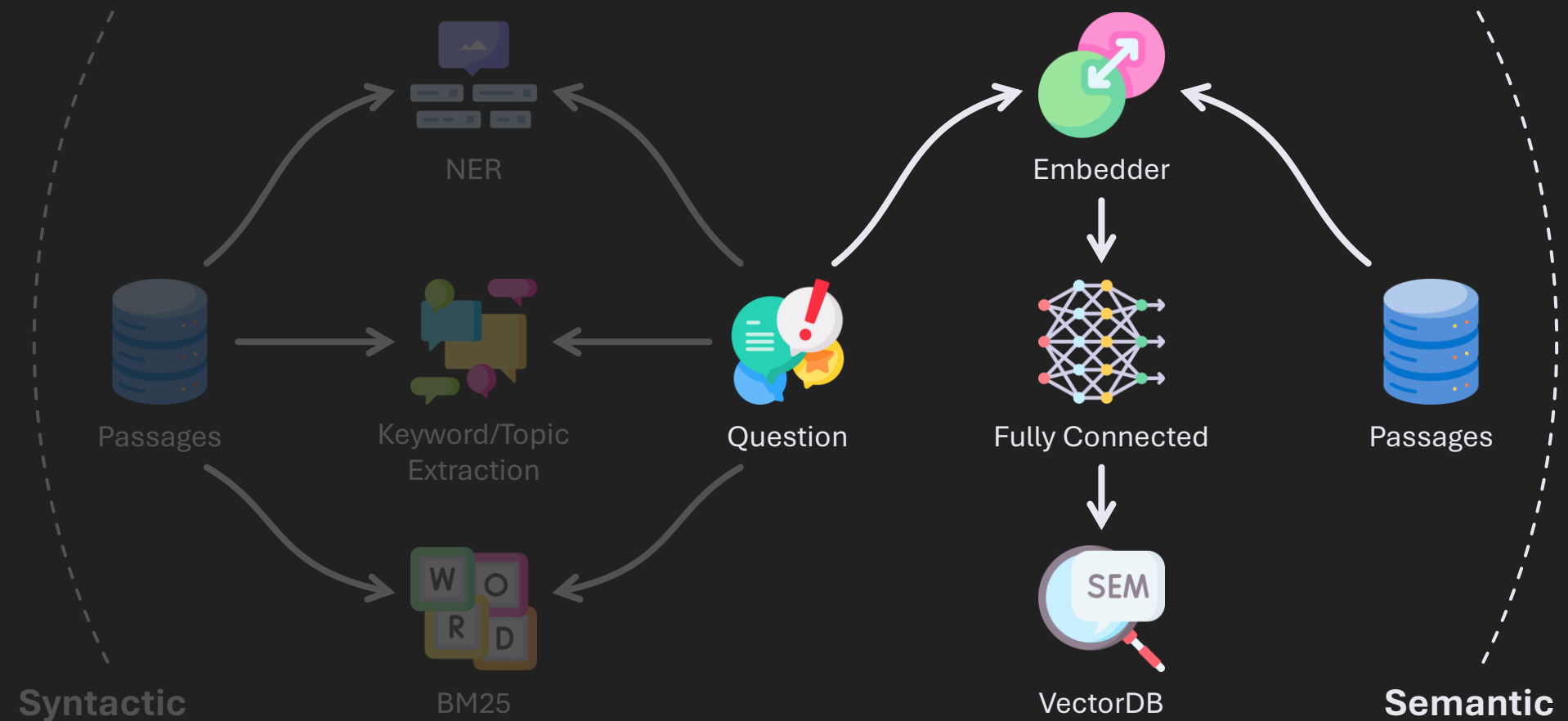


ADSP - P9 - RAG MARCO.pdf
PDF

What core ideas of the methodology are outlined in the attached PDF? Why were they chosen?



Now, we focus on the fully connected idea:



Message ChatGT3



ChatGT3 can make mistakes. Check important info.





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

December

Project objectives ...

Research question ...

Main task ...



Upgrade plan



ADSP - P9 - RAG MARCO.pdf
PDF

What core ideas of the methodology are outlined in the attached PDF? Why were they chosen?



The purpose was to calculate an **Overall Similarity Score** as a combination of syntactic and semantic similarity scores:

$$S = w_1 S_{NER} + w_2 S_{KEYWORD} + w_3 S_{TOPIC} + w_4 S_{EMB}$$

where, for the weights, we have:

$$w_1 + w_2 + w_3 + w_4 = 1$$

Then, **Instead of Manually Tuning** the weights, w_1 to w_4 , we proposed to use **Attention Mechanism** to learn them:

- **Keys (K):** The retrieved passages.
- **Queries (Q):** The input queries.
- **Values (V):** The similarity scores computed for each passage (S_{NER} , $S_{KEYWORD}$, S_{TOPIC} , and S_{EMB}).

Note that, in this transformer architecture, we compute the attention scores, which correspond to w_1 to w_4 according the following formula:

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = \text{softmax} \left(\frac{Q \times K^T}{\sqrt{d_k}} \right) V$$



Message ChatGT3





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

December

Project objectives ...

Research question ...

Main task ...



Upgrade plan


What can I help with?




ADSP - P9 - RAG MARCO.pdf
PDF


What tools or frameworks are mentioned in the attached PDF to support the methodology?
How are they applied? |



 Create Image

 Code

 Summarize

 Get advice

More





ChatGT3 ▾



ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

Methodology: Tools ...

December

Project objectives ...

Research question ...



Upgrade plan

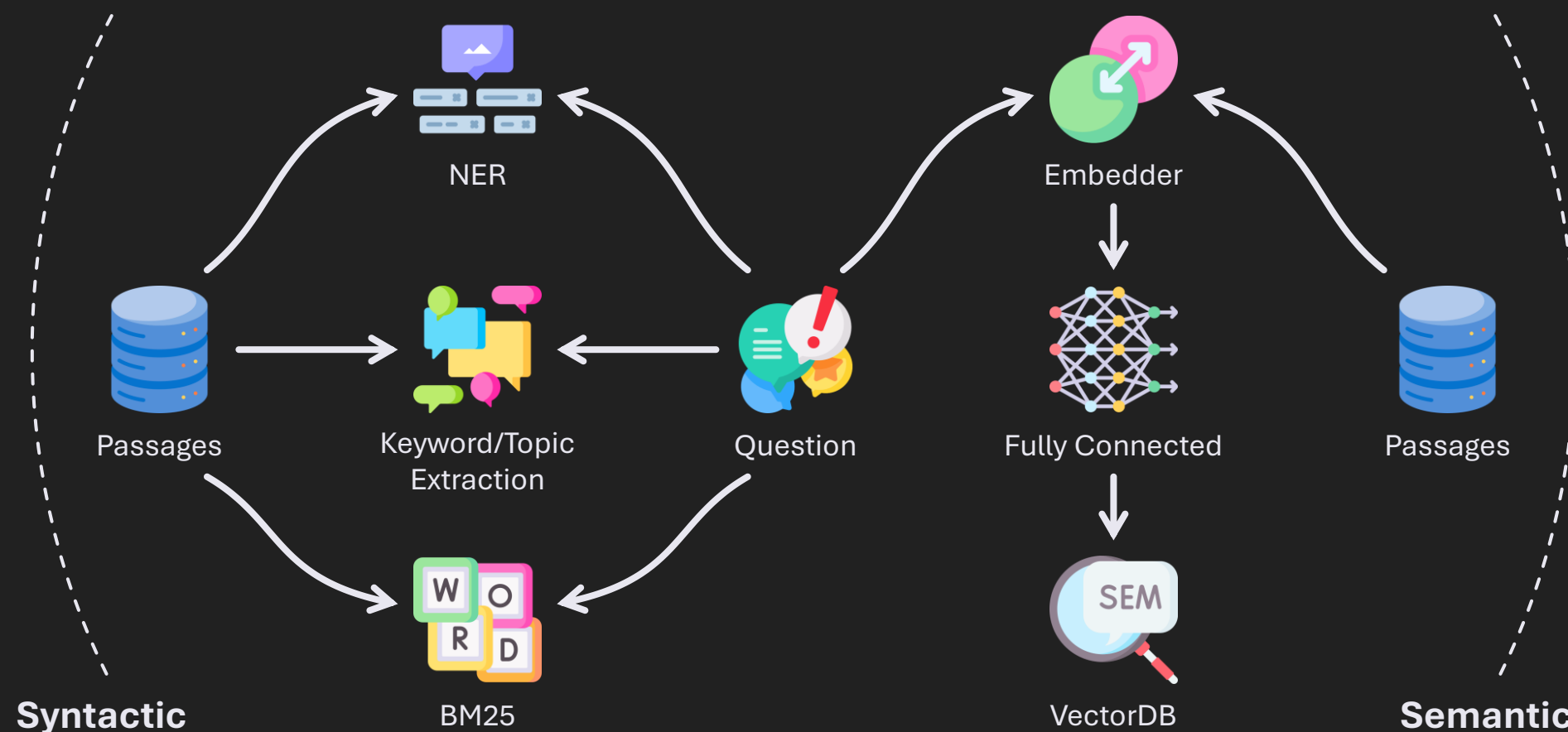


ADSP - P9 - RAG MARCO.pdf
PDF

What tools or frameworks are mentioned in the attached PDF to support the methodology? How are they applied?



First, let's have a reminder on the proposed ideas:



Message ChatGT3



ChatGT3 can make mistakes. Check important info.





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

- Introduction: RAG & poetry ...
- Project Definition ...
- Terminology ...
- Metrics: Similarity ...
- Metrics: Performance ...
- Methodology: Ideas ...
- Methodology: Tools ...

December

- Project objectives ...
- Research question ...



Upgrade plan

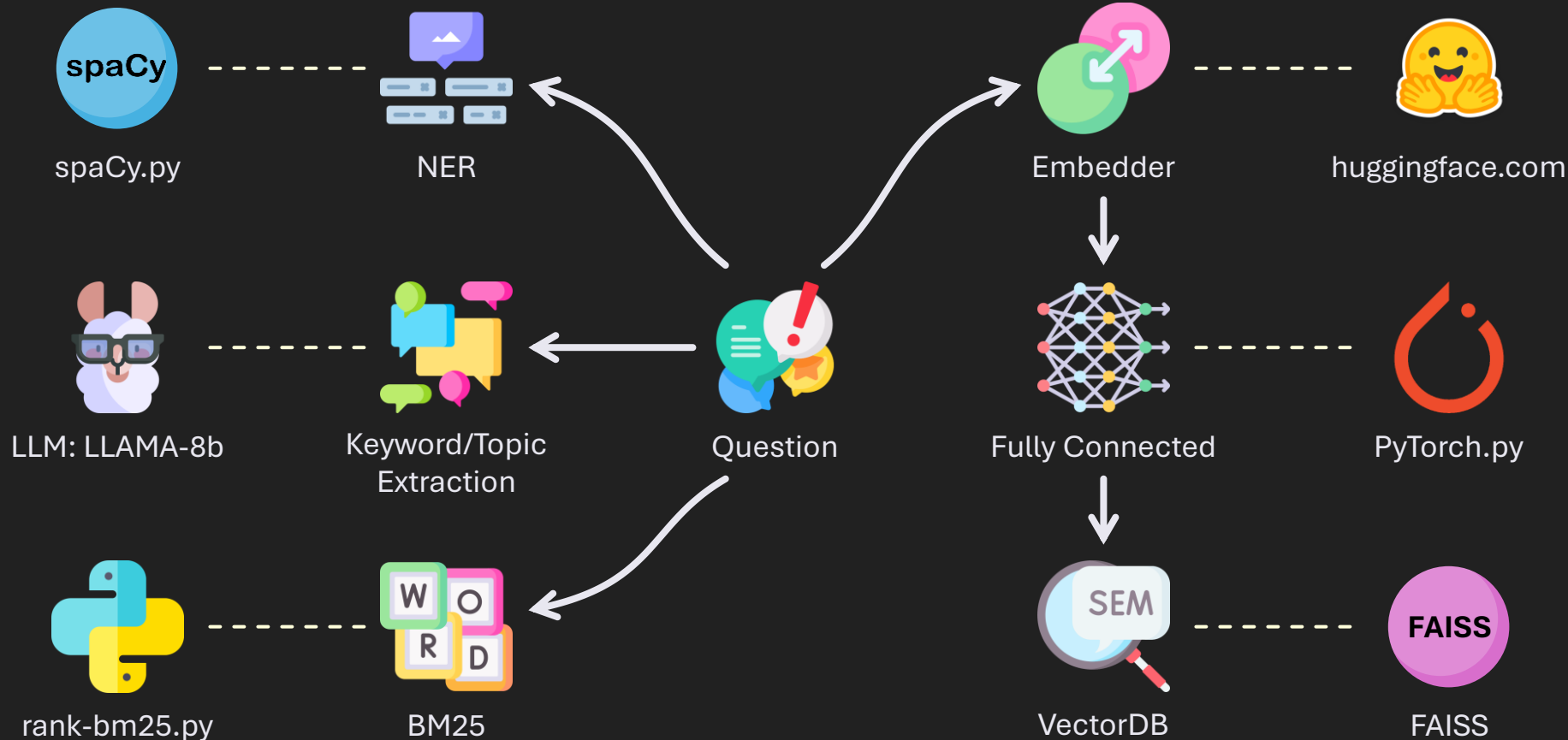


ADSP - P9 - RAG MARCO.pdf
PDF

What tools or frameworks are mentioned in the attached PDF to support the methodology? How are they applied?



Now, here's an elaboration on the tools incorporated for each idea:



Message ChatGT3





ChatGT3 ▾



ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

Methodology: Tools ...

December

Project objectives ...

Research question ...



Upgrade plan

What can I help with?



ADSP - P9 - RAG MARCO.pdf
PDF

What experimental findings are highlighted in the attached PDF? How do they relate to the objectives? |



Create Image



Code



Summarize



Get advice

More

ChatGT3 can make mistakes. Check important info.





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

Methodology: Tools ...

Experiments: Findings ...

December

Project objectives ...



Upgrade plan

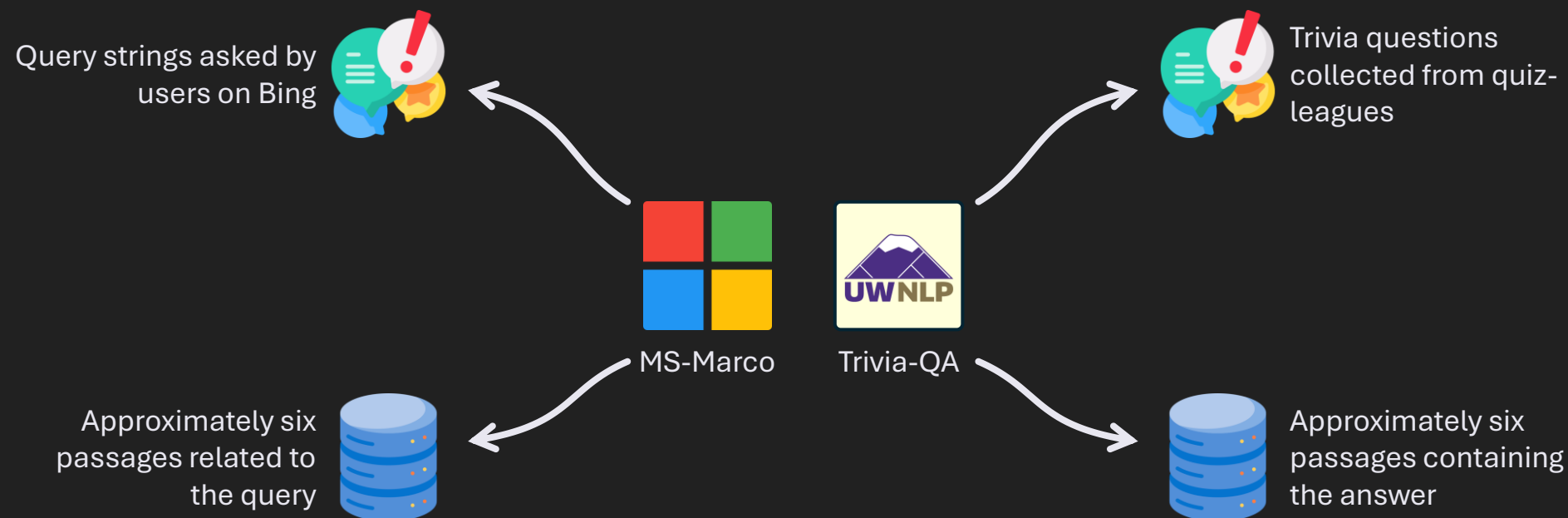


ADSP - P9 - RAG MARCO.pdf
PDF

What experimental findings are highlighted in the attached PDF? How do they relate to the objectives?



Let's have a recap on the datasets used. According to the provided PDF file, the datasets include MS-Marco and Trivia-QA. MS-Marco focuses on machine reading comprehension and passage retrieval with real-world queries, while Trivia-QA offers diverse question-answer pairs with supporting evidence. These were chosen to evaluate both retrieval and reasoning capabilities across varied scenarios.



Message ChatGT3





- ChatGT3
- Explore GT3s
- GT3's Retrieval System

Today

- Introduction: RAG & poetry
- Project Definition
- Terminology
- Metrics: Similarity
- Metrics: Performance
- Methodology: Ideas
- Methodology: Tools
- Experiments: Findings

December

- Project objectives



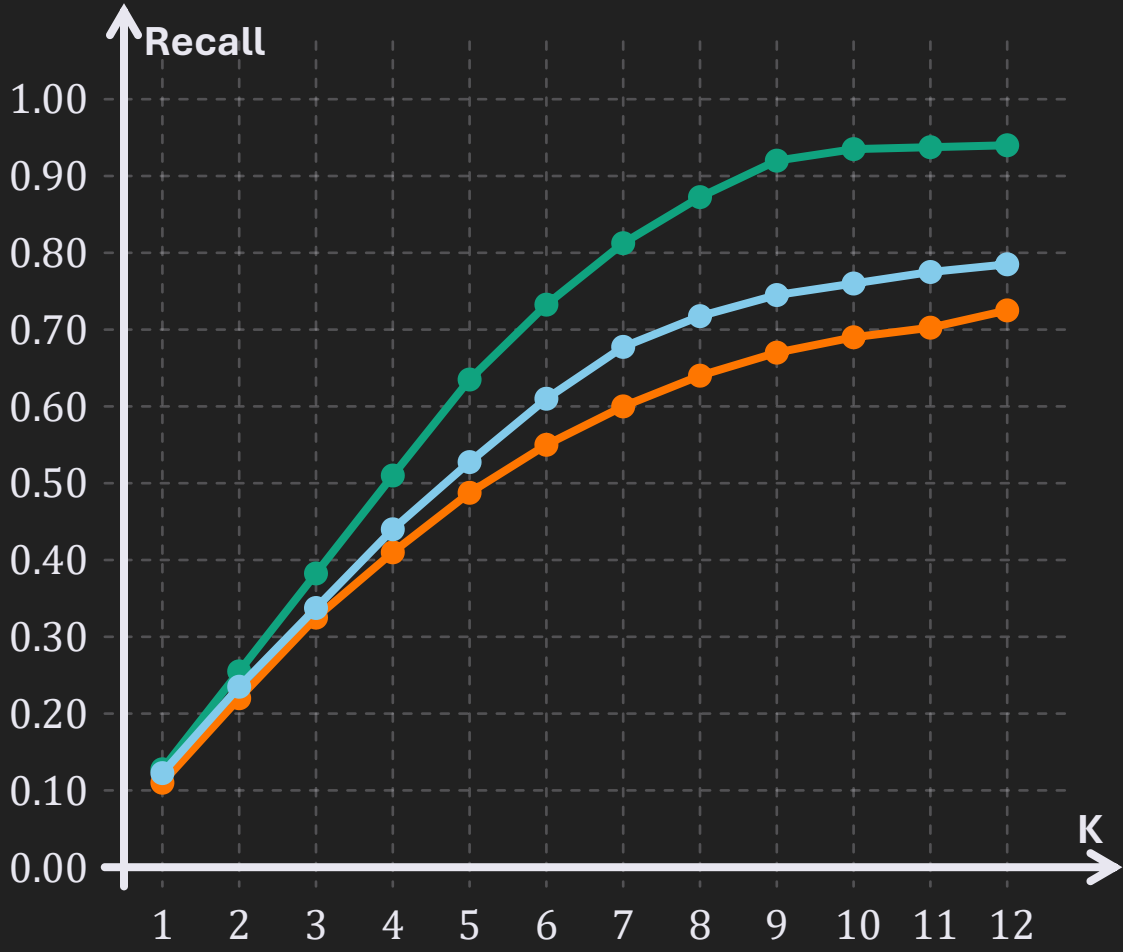
ADSP - P9 - RAG MARCO.pdf
PDF

What experimental findings are highlighted in the attached PDF? How do they relate to the objectives?



Now, let's summarize the experimental results using O-MRR, P-MRR, and recall. Together, these metrics provide a comprehensive assessment of the system's effectiveness.

Search Type	O-MRR	P-MRR
Syntactic with NER and Keyword/Topic Extraction	0.91	0.37
Syntactic with BM25	0.96	0.45
Semantic with Embedder and VectorDB	0.98	0.77



Message ChatGT3





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

Methodology: Tools ...

Experiments: Findings ...

December

Project objectives ...



Upgrade plan


What can I help with?




ADSP - P9 - RAG MARCO.pdf
PDF


What decisions were made based on experimental results, as per the attached PDF? What led to these decisions? |



 Create Image

 Code

 Summarize

 Get advice

More





ChatGT3 ▾



ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

Methodology: Tools ...

Experiments: Findings ...

Experiments: Decision ...

December



Upgrade plan

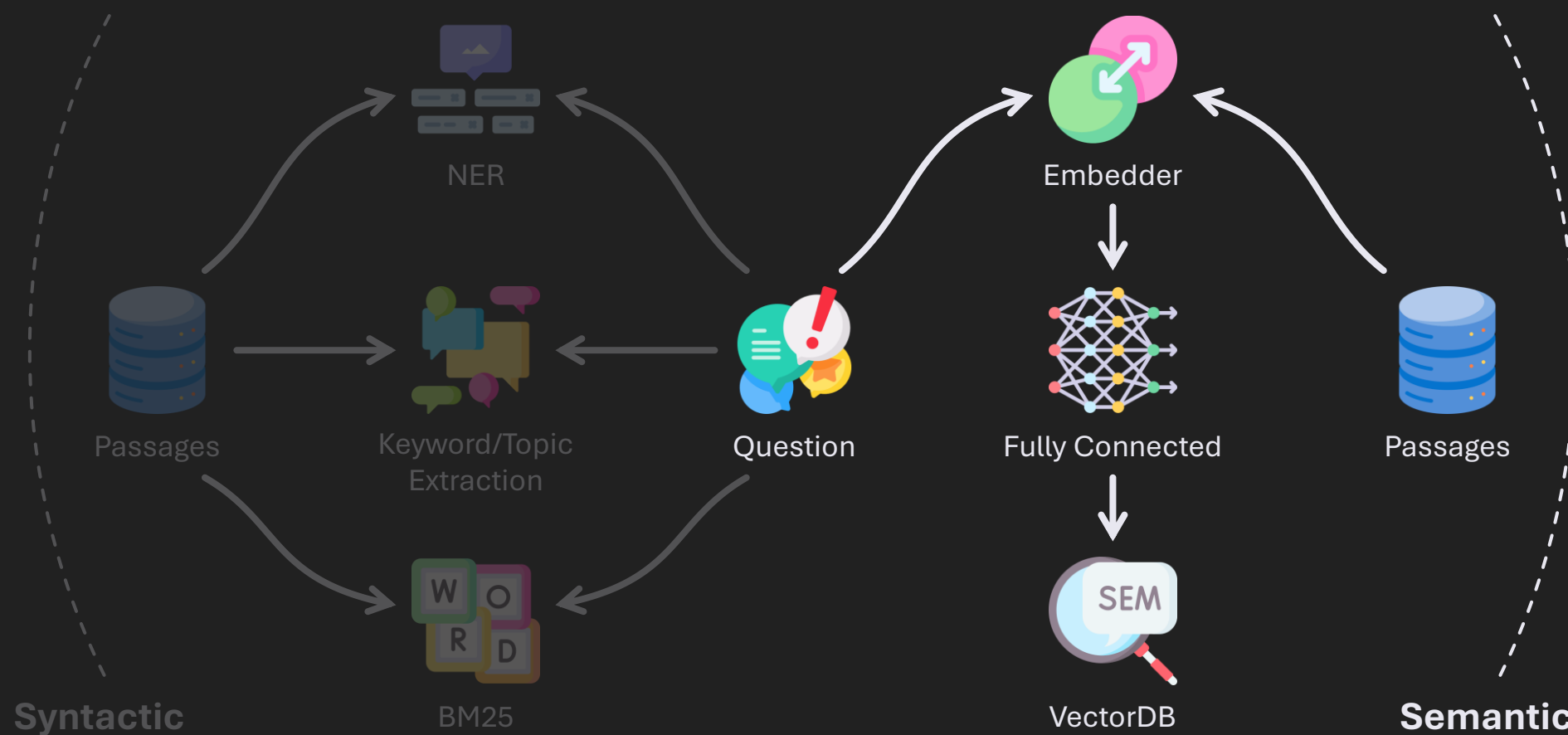


ADSP - P9 - RAG MARCO.pdf
PDF

What decisions were made based on experimental results, as per the attached PDF? What led to these decisions?



Let's start with the proposed ideas:



Message ChatGT3



ChatGT3 can make mistakes. Check important info.





ChatGT3 ▾



ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

Methodology: Tools ...

Experiments: Findings ...

Experiments: Decision ...

December



Upgrade plan

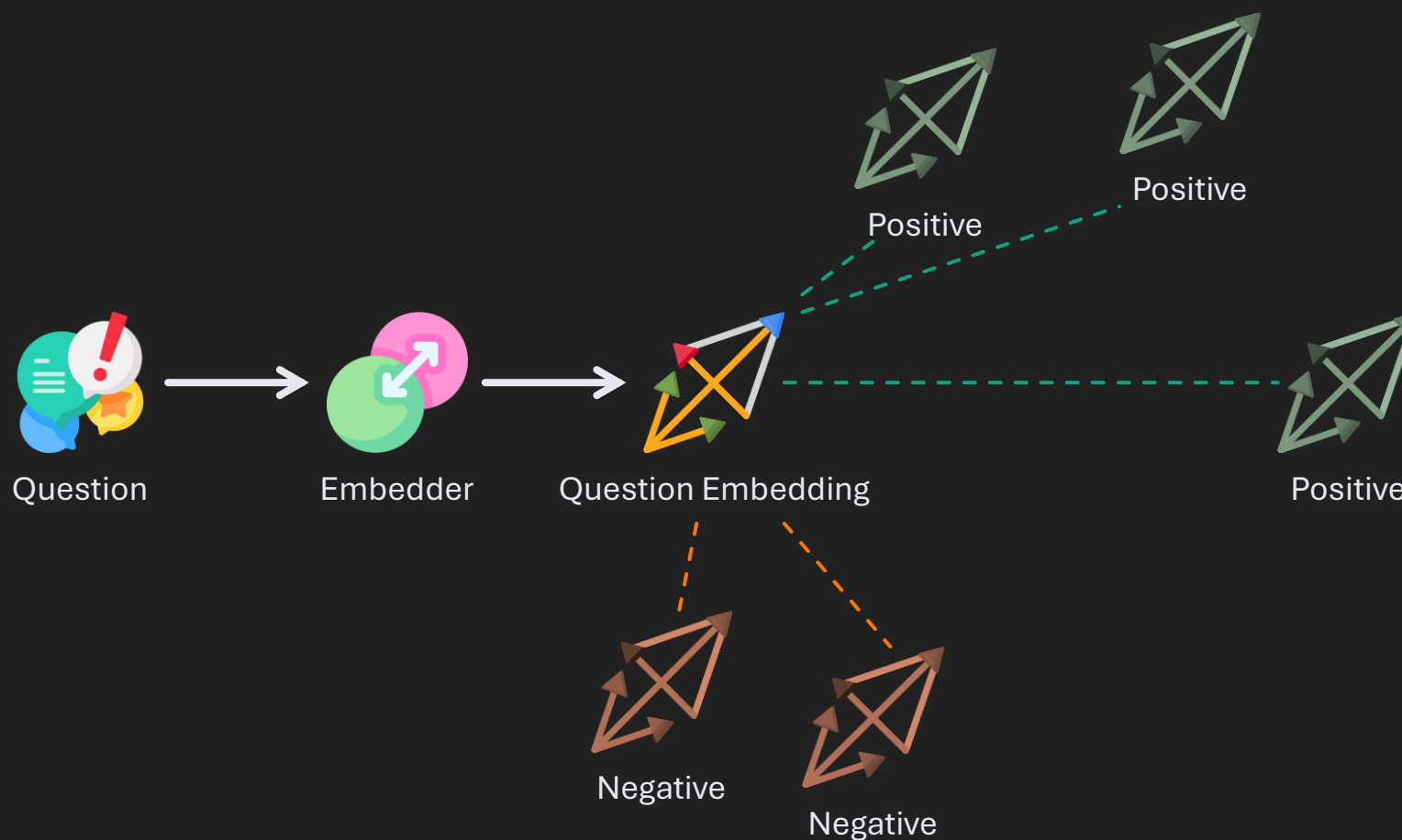


ADSP - P9 - RAG MARCO.pdf
PDF

What decisions were made based on experimental results, as per the attached PDF? What led to these decisions?



Here's the baseline:



Message ChatGT3



ChatGT3 can make mistakes. Check important info.





ChatGT3 ▾



ChatGT3



Explore GT3s



GT3's Retrieval System

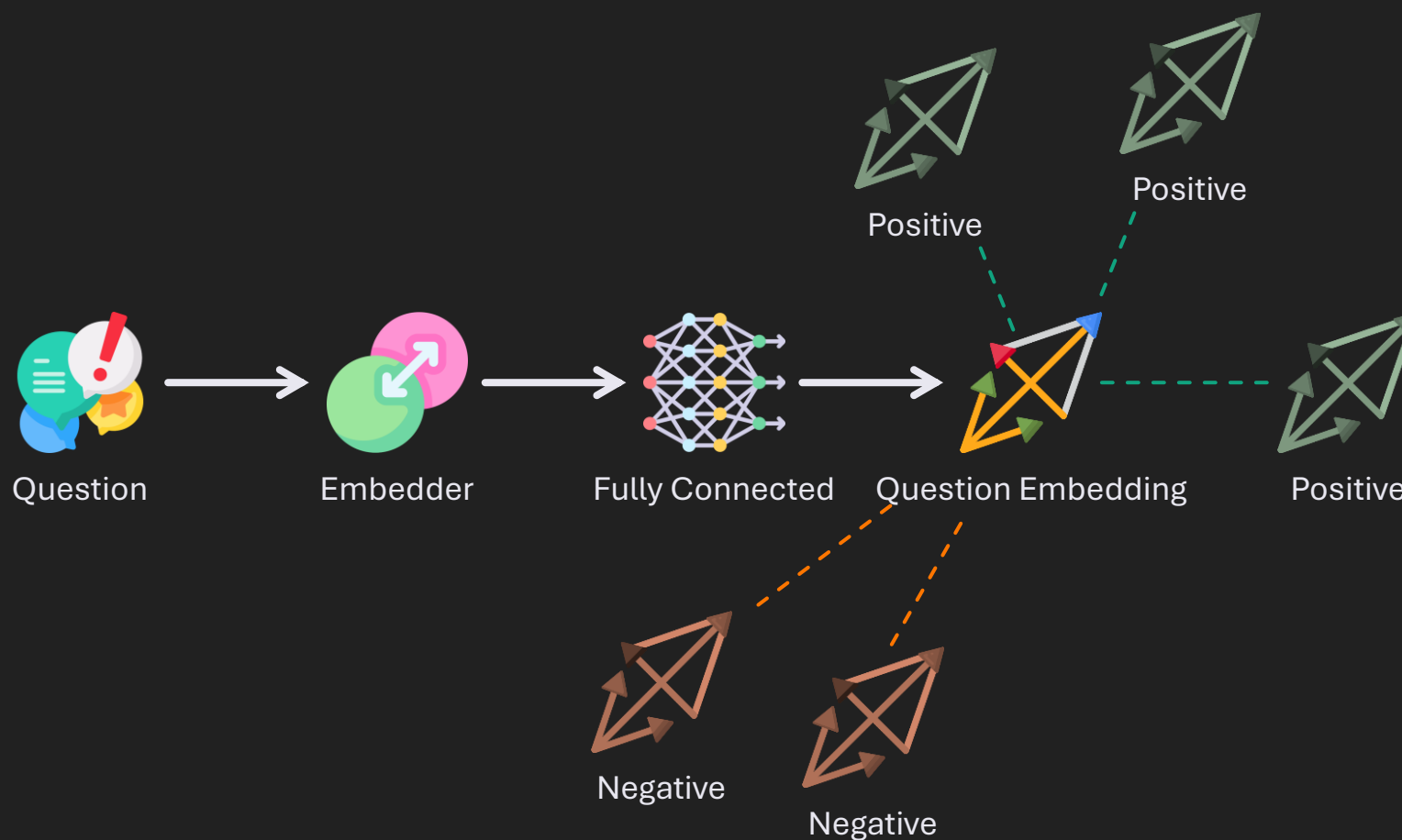


ADSP - P9 - RAG MARCO.pdf
PDF

What decisions were made based on experimental results, as per the attached PDF? What led to these decisions?



And this is how the main idea is to work in practice:



Message ChatGT3



December



Upgrade plan

ChatGT3 can make mistakes. Check important info.





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

Methodology: Tools ...

Experiments: Findings ...

Experiments: Decision ...

December



Upgrade plan

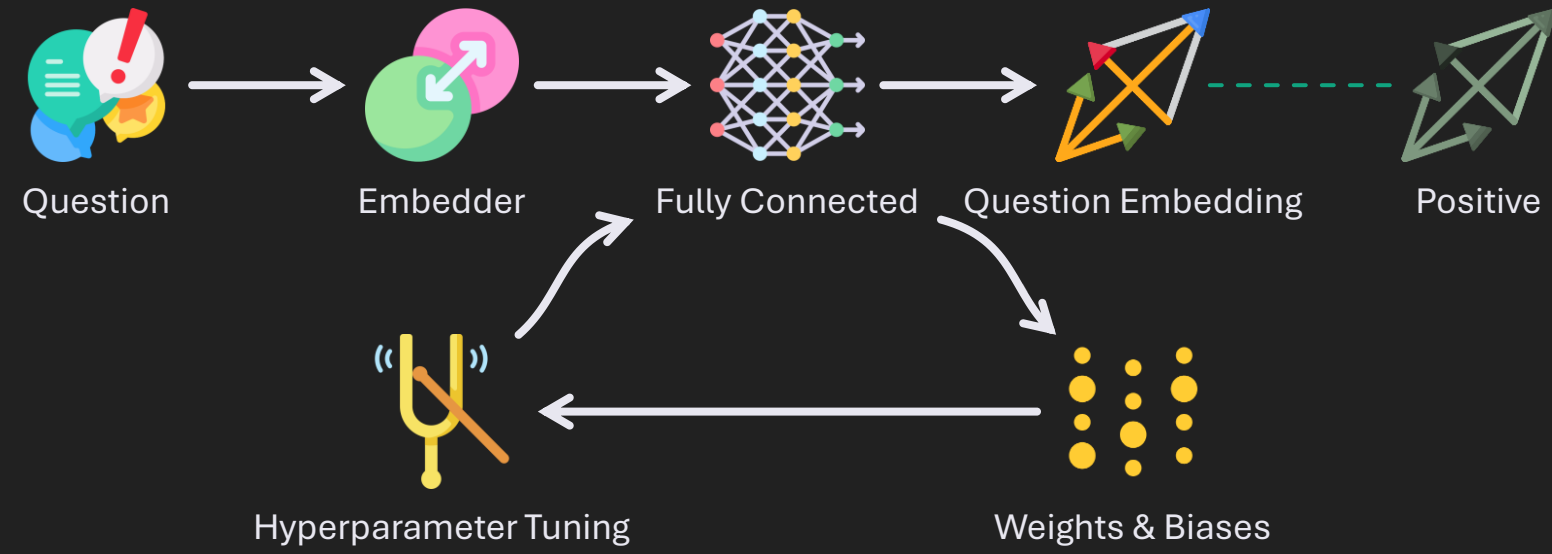


ADSP - P9 - RAG MARCO.pdf
PDF

What decisions were made based on experimental results, as per the attached PDF? What led to these decisions?



Based on the findings, it is recommended to use the **Weights & Biases Service (wandb.ai)** for **Hyperparameter Tuning** of the fully connected layer. This platform provides advanced tracking and visualization tools, enabling systematic experimentation with various configurations such as learning rates, activation functions, and layer dimensions. By integrating this service, you can efficiently compare model performances, identify optimal hyperparameters, and streamline the tuning process.



Message ChatGT3





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

Introduction: RAG & poetry ...

Project Definition ...

Terminology ...

Metrics: Similarity ...

Metrics: Performance ...

Methodology: Ideas ...

Methodology: Tools ...

Experiments: Findings ...

Experiments: Decision ...

December



Upgrade plan


What can I help with?




ADSP - P9 - RAG MARCO.pdf
PDF


What key takeaways does the attached PDF provide about this project? What impact does it highlight? |



 Create Image

 Code

 Summarize

 Get advice

More





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

- Introduction: RAG & poetry ...
- Project Definition ...
- Terminology ...
- Metrics: Similarity ...
- Metrics: Performance ...
- Methodology: Ideas ...
- Methodology: Tools ...
- Experiments: Findings ...
- Experiments: Decision ...
- Conclusion ...



Upgrade plan



ADSP - P9 - RAG MARCO.pdf
PDF

What key takeaways does the attached PDF provide about this project? What impact does it highlight?



To conclude,

- **Introduction:** The project began with a small application—retrieving Persian poems—to showcase its potential and set the stage for further exploration.
- **Terminology:** Key terms were defined to establish a shared understanding of the domain and create a foundation for clear communication.
- **Metrics:** For similarity, different metrics were defined to evaluate content relevance. Also, regarding performance, it was measured using optimistic MRR, pessimistic MRR, and recall to assess system effectiveness.
- **Methodology:** The approach combined innovative ideas with advanced tools, creating a structured framework for implementation.
- **Experiments and Results:** Experiments on datasets like MS-Marco and Trivia-QA provided critical insights into system performance, influencing decisions on design and optimization.
- **Next Steps:** Future improvements focus on refining the fully connected layer, using the weights and biases service for efficient hyperparameter tuning to enhance performance further.





ChatGT3



Explore GT3s



GT3's Retrieval System

Today

- Introduction: RAG & poetry ...
- Project Definition ...
- Terminology ...
- Metrics: Similarity ...
- Metrics: Performance ...
- Methodology: Ideas ...
- Methodology: Tools ...
- Experiments: Findings ...
- Experiments: Decision ...
- Conclusion ...



Upgrade plan



ADSP - P9 - RAG MARCO.pdf
PDF

What key takeaways does the attached PDF provide about this project? What impact does it highlight?



To conclude,

- **Introduction:** The project began with a small application—retrieving Persian poems—to showcase its potential and set the stage for further exploration.
- **Terminology:** Key terms were defined to establish a shared understanding of the domain and create a foundation for clear communication.
- **Metrics:** For similarity, different metrics were defined to evaluate content relevance. Also, regarding performance, it was measured using optimistic MRR, pessimistic MRR, and recall to assess system effectiveness.
- **Methodology:** The approach combined innovative ideas with advanced tools, creating a structured framework for implementation.
- **Experiments and Results:** Experiments on datasets like MS-Marco and Trivia-QA provided critical insights into system performance, influencing decisions on design and optimization.
- **Next Steps:** Future improvements focus on refining the fully connected layer, using the weights and biases service for efficient hyperparameter tuning to enhance performance further.



Can you help us with the exam?



- ChatGT3
- Explore GT3s
- GT3's Retrieval System

Today

- Introduction: RAG & poetry
- Project Definition
- Terminology
- Metrics: Similarity
- Metrics: Performance
- Methodology: Ideas
- Methodology: Tools
- Experiments: Findings
- Experiments: Decision
- Conclusion

ADSP - P9 - RAG MARCO.pdf

PDF

What key takeaways does the attached PDF provide about this project? What impact does it highlight?

- To conclude,
- Introduction:**

The project began with a small application—retrieving Persian poems—to showcase its potential and set the stage for further exploration.

Terminology:

Key terms were defined to establish a shared understanding of the domain and create a foundation for clear communication.

Metrics:

For similarity, different metrics were defined to evaluate content relevance. Also, regarding performance, it was measured using optimistic MRR, pessimistic MRR, and recall to assess system effectiveness.

Methodology:

The approach combined innovative ideas with advanced tools, creating a structured framework for implementation.

Experiments and Results:

Experiments on datasets like MS-Marco and Trivia-QA provided critical insights into system performance, influencing design decisions and optimization.

Next Steps:

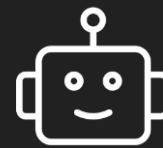
Future work includes refining the retrieval process, exploring advanced models, and integrating user feedback for efficient and accurate results.
- You've hit the Free plan limit for GT3.

You need GT3-4o to continue this chat because there's an attachment.

Your limit resets after January 31st.

Get Plus

and biases service
- Can you help us with the exam?



Oops!

Thank you for your attention, but our systems are busy at the moment.
Please take a break and ask your questions and good luck with your exam!



Homayoun Afshari



Arash Daneshvar



Hossein Khodadadi

You've hit the Free plan limit for GT3.

You need GT3-4o to continue this chat because there's an attachment.
Your limit resets after January 31st.

Get Plus