



UniVerse: ChatBot for Universities in Lebanon

Prepared by:

RIM CHEHADE – KAREN TIMANI

Supervised by:

DR. HUSSEIN HAZIMEH

Table of content

1. Motivation
2. Problem Statement
3. Meet "UniVerse"
4. Why Agentic RAG
5. Technologies Used
6. Data Preparation Pipeline
7. Data Flow
8. Prompt engineering
9. Evaluation Results
10. Challenges and Limitations
11. Future Work
12. Conclusion

Motivation

Every year, students and parents across Lebanon search for answers to urgent academic questions, but what they find is:

- ✗ Comparing universities is frustrating
- ✗ Scattered and inconsistent information
- ✗ No unified platform to explore options
- ✗ Time-consuming and stressful search processes

There must be a smarter way to get answers.





Problem Statement

How can a chatbot be designed and developed to assist students in comparing and exploring universities in Lebanon more effectively?



Meet “UniVerse”

Your AI Guide to Lebanese Universities

UniVerse is an intelligent AI chatbot designed to make academic exploration effortless. Ask it anything, from tuition fees to program offerings, and get smart, instant, and accurate answers.

Why UniVerse?

- Trained on official Lebanese university websites.
- Understands your question, finds relevant info, and responds like a real assistant.
- Powered by an agentic Retrieval-Augmented Generation (RAG)



One question. One chat. The whole university universe.

Why Agentic RAG

AGENTIC RAG COMBINES TWO POWERFUL IDEAS:



1. RETRIEVAL-AUGMENTED GENERATION (RAG)

Finds real info → Generates smart answers



2. AGENTIC REASONING

Plans, thinks in steps, uses tools like web search, and memory

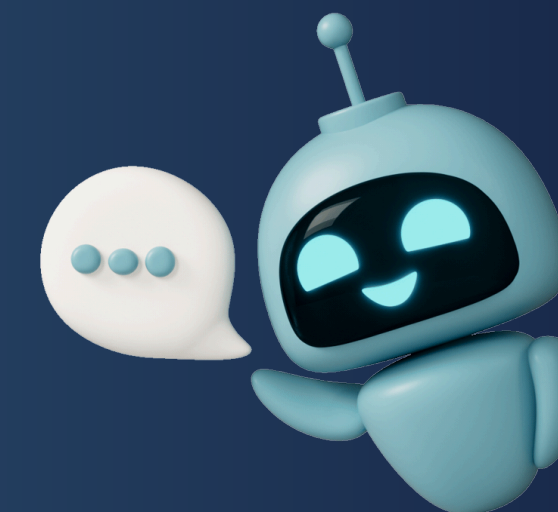
WHY WE CHOSE IT FOR UNIVERSE?

- ✓ More accurate
- ✓ Avoids hallucinations
- ✓ Handles complex questions
- ✓ Smarter than basic Q&A





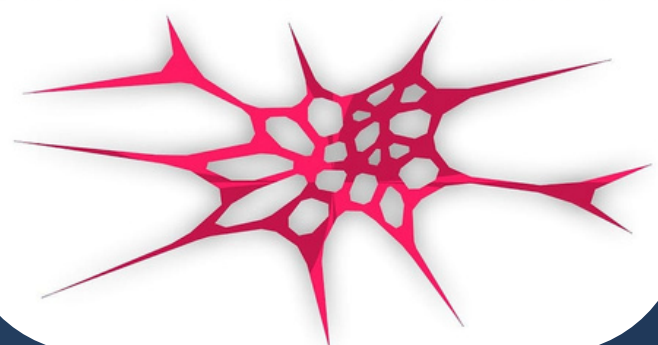
Technologies Used



Llama 2 13b Chat



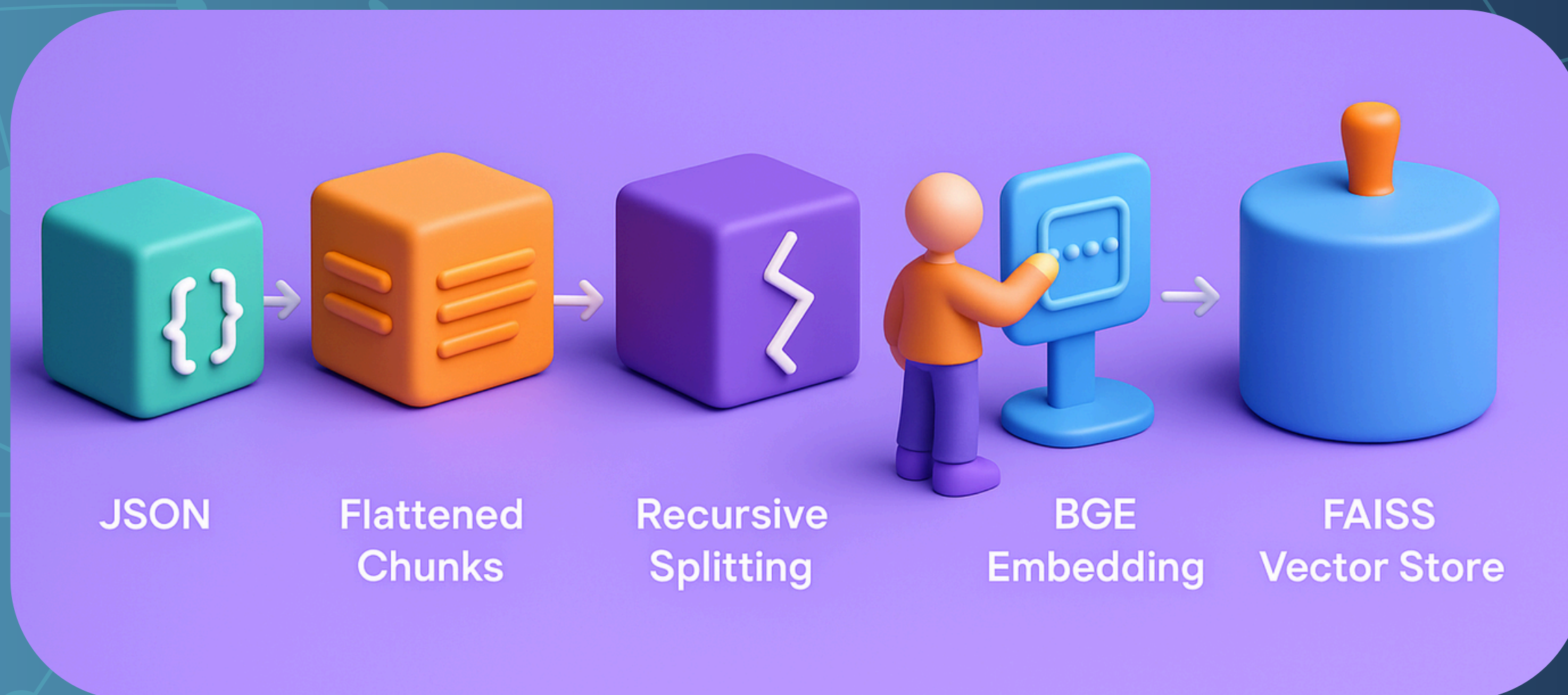
FAISS
Scalable Search With Facebook AI



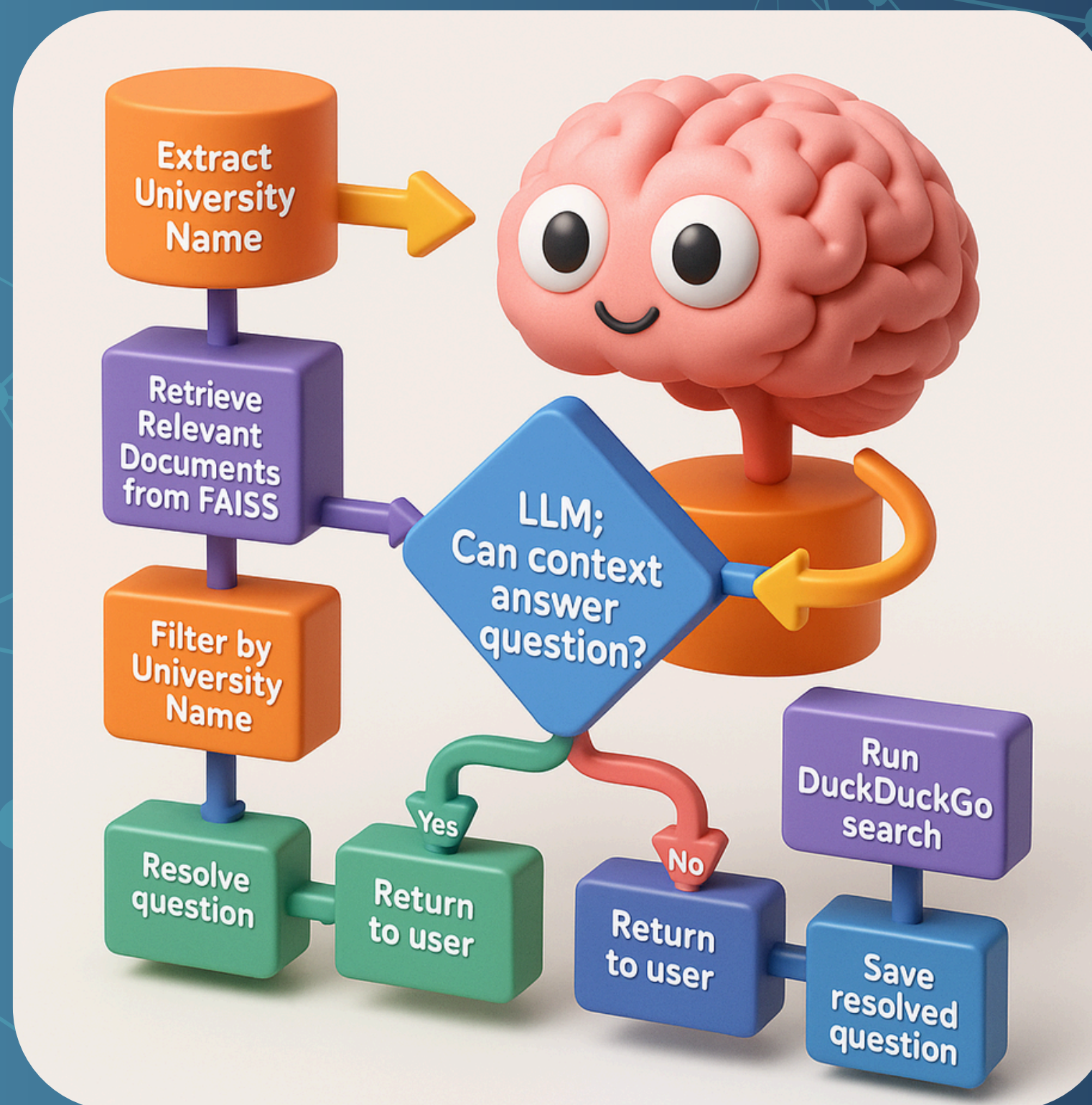

Flask

ngrok

Data Preparation Pipeline



Data Flow





Prompt Engineering

```
decision_system_prompt = """
You are an AI that only answers with a single number.

Respond:
1 - if the context CAN answer the question
0 - if the context CANNOT answer the question

Only reply with one of these digits. Do not explain or say anything else.

Context:
{context}
"""

user_prompt = """
Question: {question}

Answer: """
```

```
def resolve_query(current_question, previous_question):
    prompt = f"""You are helping resolve follow-up questions in a conversation. Given the previous user question and the current one, rewrite the current question so that it is complete and fully understandable on its own.
Previous question: "{previous_question}"
Current follow-up question: "{current_question}"

Rewritten, standalone question:
"""

    return llm.invoke(prompt).strip()
```

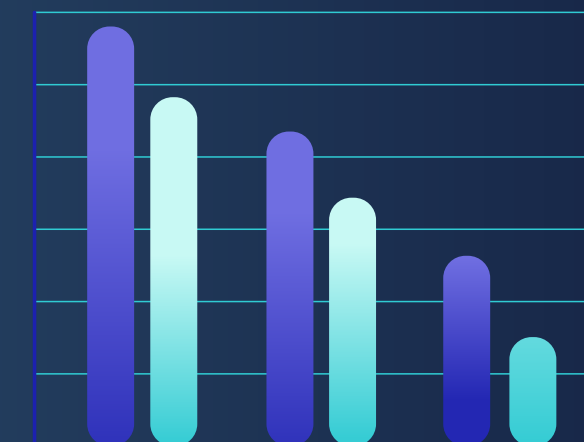
```
system_prompt = """You are an expert for answering questions. Answer the question according only to the given context.
If question cannot be answered using the context, simply say I don't know. Do not make stuff up. Your answer MUST be informative, concise, and action driven. Your response must be in Markdown.

Context: {context}
"""

user_prompt = """
Question: {question}

Answer: """
```

Evaluation Results



1. ROUGE-1 F1: 0.4734

Captures ~47% of key content words, good content alignment.



2. ROUGE-2 F1: 0.3112

31% bigram match, shows some fluency and phrasing similarity.



3. ROUGE-L F1: 0.3676

Partial sentence structure preserved — helps maintain flow.



4. BLEU: 0.2433

Low word-for-word overlap — typical for paraphrased responses.



5. COSINE SIMILARITY: 0.7757

Strong semantic closeness to reference answers — meaning is well preserved.

Challenges & Limitations

1.

LIMITED DATA SCOPE

Only 3 universities (LU, LIU, LAU) included in the knowledge base.

2.

MODEL CONSTRAINTS

Lack of GPU power limited experimenting with LLMs affecting scalability.

3.

MULTI-QUESTION INPUTS LIMITATION

It may miss parts of the response when users ask multiple questions at once.

4.

NO REAL-TIME DATA

No live integration for updates like application deadlines.



Future Work

EXPAND UNIVERSITY COVERAGE

Add more Lebanese and regional universities.

ADD A QUIZ FEATURE

introduce a short quiz to help students match with majors.

IMPROVE FALLBACK SEARCH

Use headless browsers for complex pages.



UPGRADE DEPLOYMENT

Host on stable, cloud-based infrastructure.

ENHANCE LONG TERM MEMORY

Smarter context retention in conversations.

SUPPORT ARABIC LANGUAGE

Add multilingual and Arabic speech support.

Conclusion

In this project, we developed UniVerse – an intelligent, agentic RAG-based chatbot that helps students navigate Lebanese university information quickly and accurately.

By combining semantic search, agentic reasoning, and retrieved content grounding, UniVerse offers a more reliable and scalable alternative to traditional Q&A systems.

While still a prototype, it lays the foundation for a powerful educational assistant, one that can evolve to support more universities, Arabic language, multi-turn reasoning, and live data updates in the future.



Thank You

FOR YOUR ATTENTION

Any Question?

[Github Repository Link](#)

[Demo](#) 

