



University College Dublin

School of Computer Science
MSc in Computer Science NL

Lumina:

An Intelligent Guide for Cosmic Self-Reflection

Submitted by:

Saumya Vani Namballa

Student ID: **24217190**

Email: `saumya.namballa@ucdconnect.ie`

Module:

COMP47980: Generative AI And Language Models

Project Report - 2025

Contents

1	Introduction	3
1.1	Objective	3
1.2	Motivation	3
1.3	Implementation	3
2	Assistant Design: A Schematic View	4
2.1	Design Objective	4
2.2	Assistant Tools Used	4
2.3	Excluded Features	4
2.4	System Diagram Overview	4
2.5	Design Summary	4
3	Added Value: More than Mere ChatGPT or Basic LLM	5
3.1	Limitations of Basic ChatGPT Usage	5
3.2	How Lumina Adds Value	6
3.3	Beyond LLM-as-a-Service	6
3.4	Tool Justification	6
4	Outside Knowledge: Curated Data Sources	6
4.1	Data Sources and Purpose	6
4.2	Corpus Structure and Preprocessing	7
4.3	Expertise Enabled by Documents	7
4.4	Sample Knowledge Outputs	7
5	Worked Examples: Your Assistant in Action	7
5.1	Example 1: Personalized Personality Summary	7
5.2	Example 2: Follow-Up on Ascendant in Scorpio	8
5.3	Example 3: Philosophical Query – “What is Vedic Astrology?”	8
5.4	Design Insights	9
6	Summary and Conclusions	9
	Acknowledgements	10
	References	10

List of Figures

1	Schematic Overview of Lumina 's Dual-Channel Architecture	5
2	Lumina generating a personality summary based on user birth details	8
3	Lumina explaining the significance of an Ascendant in Scorpio using RAG-based PDF grounding	8
4	Lumina providing a canonical overview of Vedic astrology	9

1 Introduction

In a world more and more dominated by algorithms and automation, consumers are looking for tools that offer both personalized understanding and contextual awareness. This project presents **Lumina**—a hybrid assistant that blends retrieval-augmented generation (RAG) on user-uploaded documents with real-time Vedic astrology-derived personality briefs. Unlike one-size-fits-all chatbots, Lumina is designed to explain, interpret, and advise with domain-specific reasoning, rather than empathize or counsel.

Lumina provides two main abilities:

1. A document-grounded Q&A engine with LangChain and FAISS [4, 5], to provide accurate, context-dependent responses from user PDFs.
2. An astrology personality tool based on birth chart data fetched from FreeAstrologyAPI [6] and described with GPT-4o-mini [3].

1.1 Objective

The objective is to develop a dual-purpose assistant — one that provides answers based on uploaded documents and another one that offers deterministic astrological overviews from real-time data. It works through a pipeline of structured form integrating retrieval, semantic reasoning, and symbolic interpretation.

1.2 Motivation

In contrast to most general-question or transactional-task assistants, Lumina aids both introspective (astrology) and fact-based (document) questions. Vedic astrology is a thinking system, while RAG provides reliable, tailored responses from proximate sources. Together, these provide a particularly practical and reflective experience.

1.3 Implementation

Developed with **FastAPI** in a Google Colab environment, Lumina has two exposed endpoints:

- **/query** — Searches semantically relevant chunks from PDFs indexed in it and returns grounded responses through GPT-4o-mini [3].
- **/astro** — Makes API calls using birth information to FreeAstrologyAPI [6] and produces personality summaries.

Uploaded documents are stored and indexed through FAISS. Chart data is converted into structured prompting for summarization through LLM. The system employs two important techniques:

1. **RAG-based File Search** (LangChain + FAISS [4, 5])
2. **Function-Based API Integration** (FreeAstrologyAPI [6])

Interaction Logic: For directing the query in the correct direction, Lumina employs an easy keyword recognition strategy. If the message of the user includes any astrology-related keywords of the pre-listed ones:

```
astro_keywords = ["horoscope", "personality", "planets"]
```

the assistant activates the astrology pipeline and asks for birth information. Otherwise, it presumes a document-oriented question and sets the retrieval-augmented generation flow.

2 Assistant Design: A Schematic View

Lumina is designed as a document-grounded dual-modality assistant that synergizes document-based intelligence and real-time astrological analysis. It uses the OpenAI Assistants API paradigm incorporating Retrieval-Augmented Generation (RAG) and external API functionality calls into a modular, thread-aware system.

2.1 Design Objective

Lumina is designed to serve two core purposes:

1. Deliver accurate, context-grounded answers from user-uploaded PDFs using RAG.
2. Generate structured personality summaries from real-time birth chart data via astrology APIs.

2.2 Assistant Tools Used

Two primary tools shape Lumina’s architecture:

- **File Search (LangChain + FAISS [4, 5]):** PDFs are split into semantic chunks using `RecursiveCharacterTextSplitter`, embedded with `OpenAIEmbeddings`, and indexed using FAISS [5]. User queries retrieve relevant chunks passed into a `ConversationalRetrievalChain` to generate grounded answers via GPT [4].
- **Function Calling (FreeAstrologyAPI [6]):** Birth details are submitted to FreeAstrologyAPI [6], which returns Rāśi and Navamsa data. This is parsed into structured prompts for GPT-based personality summarization [3].

2.3 Excluded Features

Certain Assistants API capabilities were intentionally omitted:

- **Code Interpreter:** Not required due to absence of mathematical or programmatic reasoning.
- **Persistent Thread Memory:** Lumina is designed statelessly, maintaining context only through a temporary chat history parameter.
- **Native Function Schema:** API calls are implemented via direct REST integration for transparency and flexibility.

2.4 System Diagram Overview

Figure 1 visualizes the dual-pipeline structure:

Key components include:

- **User Input:** PDFs and birth details.
- **RAG Pipeline:** Chunking → Embedding → FAISS Index → Retrieval → GPT Response.
- **Astrology Pipeline:** API Call → Chart Data → Prompt Template → GPT Summary.
- **Output:** Structured JSON returned via REST API.

2.5 Design Summary

Lumina’s design segregates concerns and aggregates delivery. The two-pipeline architecture provides modularity, accuracy, and personalization — making it distinct from generic assistants on the basis of domain knowledge and tool coordination.

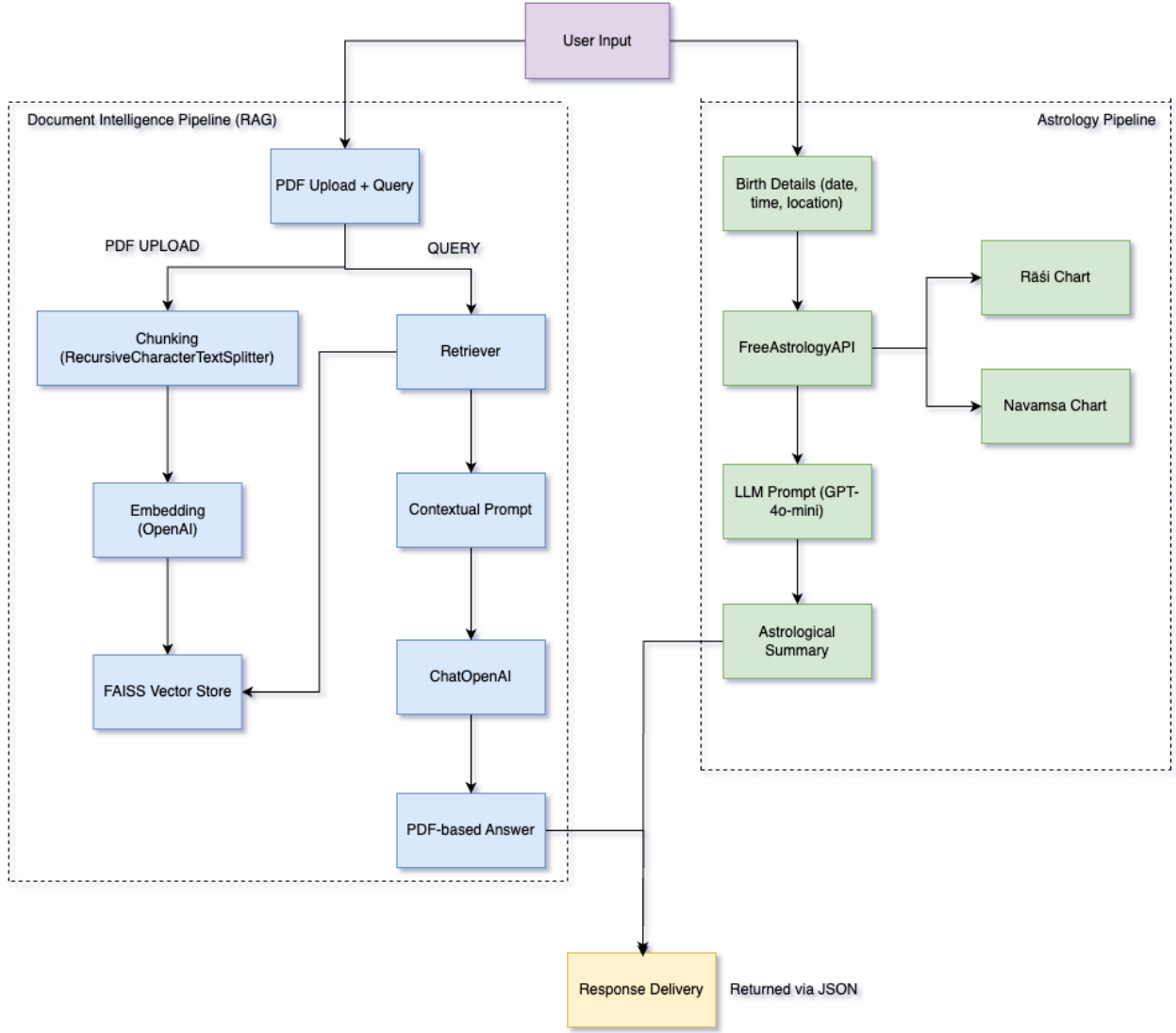


Figure 1: Schematic Overview of **Lumina**’s Dual-Channel Architecture

3 Added Value: More than Mere ChatGPT or Basic LLM

Lumina goes beyond an ordinary ChatGPT interaction by integrating document-grounded smarts, API-integrated execution at runtime, and domain-knowledge-based reasoning. It is designed as a task-specialized assistant that remedies the shortcomings of generic LLMs via tool orchestration and external tools, along with modular control over context.

3.1 Limitations of Basic ChatGPT Usage

Standard ChatGPT lacks:

- **Contextual Access to Local Files:** It cannot ingest or retrieve multi-document data without manual input.
- **Real-Time Integration:** No support for API-based planetary chart generation or personalized birth data interpretation.
- **Hallucination Guardrails:** General-purpose prompting increases the risk of speculative or misleading outputs.

3.2 How Lumina Adds Value

Lumina overcomes these gaps with:

- **RAG-Powered File Search:** PDFs are chunked, embedded, and indexed via FAISS [5], enabling grounded answers based on user-specific documents [4].
- **Astrology API Integration:** Real-time chart data is fetched from FreeAstrologyAPI [6] and interpreted via structured GPT prompts [3].
- **Controlled Prompting:** Domain-aware templates ensure interpretability, relevance, and tone consistency across both document and astrology pipelines.

3.3 Beyond LLM-as-a-Service

Lumina enhances the LLM with:

- Scoped context management per task
- Real-world external data through APIs
- Task-specific role conditioning (e.g., astrology interpreter)
- Isolated, modular pipelines for flexibility and reuse

3.4 Tool Justification

- **LangChain + FAISS** [4, 5]: For accurate, document-grounded retrieval and Q&A.
- **FreeAstrologyAPI** [6]: For dynamic chart generation and personality analysis.
- **OpenAI API** [3]: For structured generation based on retrieved or computed context.

The assistant does not employ tools such as the Code Interpreter or persistent memory, as they are unnecessary for its core use cases. Computation-heavy tasks are not present, and user sessions are intentionally stateless to preserve simplicity and portability. Instead of OpenAI’s native function schema, REST-style local implementations are used for transparency and better compatibility with Colab-based deployments.

Collectively, such design allows Lumina to be a grounded, dynamic, and extensible personal assistant, significantly beyond the one-size-fits-all scope of a standard ChatGPT session.

4 Outside Knowledge: Curated Data Sources

Lumina leverages carefully curated Vedic astrology texts to facilitate true and document-grounded answers through its RAG pipeline. These sources provide depth, authority, and fidelity to ancient interpretations.

4.1 Data Sources and Purpose

Two foundational texts were selected and uploaded via Google Drive:

1. *Vedic Astrology Textbook* [2]
2. *The Art and Science of Vedic Astrology: Intermediate Principles* [1]

Both documents offer structured insights into planetary roles, house meanings, karmic principles, and chart interpretation techniques—forming the knowledge base for Lumina’s astrology-related queries.

4.2 Corpus Structure and Preprocessing

Combined, the PDFs span over 1000 pages and are preprocessed as follows:

- Loaded via `PyPDFLoader`
- Split using `RecursiveCharacterTextSplitter` (1000-character chunks, 150-character overlap)
- Embedded using OpenAI’s `text-embedding` model
- Indexed with FAISS for fast semantic retrieval

This ensures semantic continuity while optimizing retrieval precision.

4.3 Expertise Enabled by Documents

The indexed corpus allows Lumina to:

- Provide precise interpretations of planetary placements
- Answer configuration-specific questions using classical logic [1, 2]
- Minimize hallucinations by relying strictly on curated sources

4.4 Sample Knowledge Outputs

“When Mars is placed in the seventh house, it often indicates a fiery or passionate nature in partnerships...”

“The Moon in the first house reflects heightened emotional awareness... It increases intuition but may also bring emotional sensitivity.”

These samples illustrate how Lumina’s responses remain faithful to canonical Vedic interpretations through its embedded document intelligence.

5 Worked Examples: Your Assistant in Action

This section illustrates Lumina’s dual functionality with live interaction flows—demonstrating astrology-based personality generation and document-grounded question answering.

5.1 Example 1: Personalized Personality Summary

User Input:

personality check

System Flow:

- Prompts user for name, date, time, and location of birth
- Resolves location using `geopy` (Nominatim)
- Fetches planetary charts via `FreeAstrologyAPI` [6]
- Passes chart data to GPT-4o-mini [3] for summarization

Lumina’s Output:


```

Welcome to Lumina – your intelligent assistant for Vedic Astrology & personality summaries.

I can help you with:
1. Answering questions related to Vedic astrology concepts, planetary roles, charts, and principles.
2. Summarizing your personality and behavioral traits using your birth details based on Vedic astrology.

Just type your question to begin. Type 'exit' to end the session.

You: i want to do personality check
I'll need the birth details to get started.
Name: Saumya
Date of Birth (YYYY-MM-DD): 1999-10-18
Time of Birth (HH:MM, 24hr format): 08:25
City or Location Name (e.g. Dublin): Srikakulam
Resolved Location: Srikakulam -> Latitude: 18.2949307, Longitude: 83.8938844
Timezone Offset (e.g. 0.0 for Dublin): 5.5

Summarising Saumya's personality...

Lumina: Hey Saumya! You have a captivating blend of depth and warmth in your personality. With your Ascendant in Scorpio, you possess an intense and magneti

```

Figure 2: Lumina generating a personality summary based on user birth details

Hey Saumya! You have a captivating blend of depth and warmth in your personality. With your Ascendant in Scorpio, you possess an intense and magnetic presence that draws people in. ... You're likely to be ambitious yet nurturing, always striving for growth while caring deeply for those around you. Embrace your unique blend of passion and diplomacy; it's truly a gift!

Comment: Unlike generic chatbots, Lumina delivers a structured interpretation derived from real-time astrological data using domain-aware prompting and API integration.

5.2 Example 2: Follow-Up on Ascendant in Scorpio

User Input:

What is the meaning of Ascendant in Scorpio?

Lumina's Output:

Having an Ascendant in Scorpio means the person presents with intensity and depth. Scorpio's passionate and transformative ... This placement indicates a person unafraid of change, often drawn to personal growth and transformation.

```

You: {What is the meaning of Ascendant in Scorpio?

Here's what I found.

Lumina: Having an Ascendant in Scorpio means that the person is likely to present themselves with intensity and depth. Scorpio is known for its passionate a
In terms of health and vitality, a Scorpio Ascendant might indicate a robust constitution, but they may also be prone to emotional ups and downs, which can
Overall, a Scorpio Ascendant suggests a person who is not afraid to delve into the depths of life, embracing transformation and change, and often seeking de

```

Figure 3: Lumina explaining the significance of an Ascendant in Scorpio using RAG-based PDF grounding

Explanation: The answer is retrieved via RAG from indexed Vedic astrology PDFs [1, 2]. Lumina delivers interpretive depth grounded in classical texts—beyond what a generic LLM like ChatGPT can infer without sources.

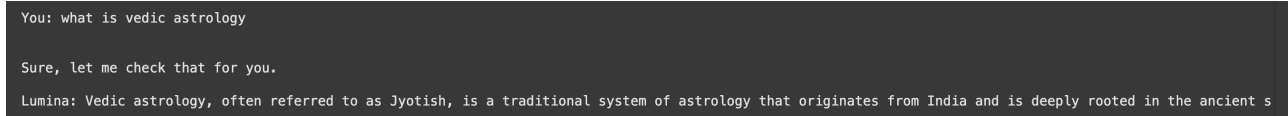
5.3 Example 3: Philosophical Query – “What is Vedic Astrology?”

User Input:

What is Vedic astrology?

Lumina's Output:

Vedic astrology, often referred to as Jyotish, is a traditional system of astrology that originates from India and is deeply rooted in the ancient scriptures known as the Vedas. The term "Vedic" signifies its connection to these sacred texts, which are considered to encompass knowledge across various subjects. Vedic astrology is viewed as a "Vedaanga," or a limb of the Vedas, and is often described as the "eye of the Vedas." This practice focuses on the positions and movements of celestial bodies, or grahas (planets), to gain insights into human life and the universe. It's a fascinating blend of spirituality and science, aiming to illuminate the path of individuals through the wisdom of the stars.



The screenshot shows a chat interface with a dark background. The user's input is "You: what is vedic astrology". The model's response is "Sure, let me check that for you." followed by "Lumina: Vedic astrology, often referred to as Jyotish, is a traditional system of astrology that originates from India and is deeply rooted in the ancient s".

Figure 4: Lumina providing a canonical overview of Vedic astrology

Explanation: Lumina’s response is sourced from classical Vedic texts [1, 2] and showcases how it blends document retrieval with structured summarization. Unlike ChatGPT, Lumina delivers grounded and semantically rich insights on traditional knowledge systems like astrology.

5.4 Design Insights

These examples highlight Lumina’s differentiated value:

- **Personalization:** API-powered charts deliver hyper-relevant summaries.
- **Contextual Integrity:** Document responses are grounded in authoritative texts, not fabricated.
- **Dual Utility:** The assistant seamlessly merges mental wellness exploration with precise information retrieval.

6 Summary and Conclusions

Lumina is a two-way assistant that connects document smarts and introspection. It integrates Retrieval-Augmented Generation (RAG) from vetted Vedic astrology texts with live chart analysis through API [6], providing personalized, context-sensitive advice rooted in tradition and user input.

From queries such as *"Mars in the seventh house"* to complete personality profiles, Lumina shows how LLMs can be extended through templated prompts and external information to provide meaningful guidance.

Future Directions

A number of additions can extend Lumina’s functionality:

- **Compatibility Analysis:** Comparing two users’ birth charts to analyze relationship dynamics through synastry.
- **Reflective Support:** Providing journaling prompts and emotional insights through psychological models and customized prompts.
- **Time-Based Forecasting:** Incorporating dasha and transit systems for monthly/yearly forecasts.

- **Multimodal UX:** Including chart visualizations and voice interfaces for more immersive interaction.
- **Adaptive Memory:** Utilizing thread memory in order to tailor responses in accordance with repeated themes or issues.

Essentially, Lumina is a new generation of AI assistants that combine analytical form with symbolic depth. As AI continues to progress, such applications might have a real impact on the way users think, develop, and connect inwardly and outwardly.

Acknowledgements

This project represents independent and original work completed solely by **Saumya Vani Namballa (Student ID: 24217190)**. Every sentence and idea presented in this report was conceived, written, and structured by the author alone. No part of this document has been copied from another individual's work. Wherever external references have been used, they have been clearly cited in the References section, and any quotations are properly enclosed within quotation marks and attributed to their original sources.

The author would like to express sincere gratitude to the teaching staff of the module **COMP47980: Generative AI and Language Models** at University College Dublin. The module has provided not only the technical foundation required to undertake this project, but also the intellectual inspiration to explore language models in innovative and meaningful ways. The guidance, discussions, and timely feedback offered throughout the semester played a pivotal role in shaping the direction and execution of this assistant system.

Finally, appreciation is extended to the broader open-source community and API providers whose tools and documentation made the implementation of this assistant possible.

References

- [1] A. Ahuja and K. Ahuja, *The Art and Science of Vedic Astrology: Intermediate Principles*. Bloomington, IN: Balboa Press, 2013.
- [2] S. K. Rao, *Vedic Astrology Textbook*. Hyderabad, India: Saptarishis Astrology, 2021.
- [3] OpenAI, "OpenAI API Documentation," 2023. [Online]. Available: <https://platform.openai.com/docs>
- [4] LangChain, "LangChain Documentation," 2023. [Online]. Available: <https://docs.langchain.com>
- [5] Facebook AI Research, "FAISS: Facebook AI Similarity Search," 2023. [Online]. Available: <https://github.com/facebookresearch/faiss>
- [6] FreeAstrologyAPI, "Free Astrology API: Real-Time Birth Chart and Planetary Calculations," 2024. [Online]. Available: <https://freeastrologyapi.com>