Project - High Level Design

on

Multimodal Education Creator

Course Name: Gen AI

Institution Name: Medicaps University – Datagami Skill Based Course

Student Name(s) & Enrolment Number(s):

Sr no	Student Name	Enrolment Number
01	AMAN VERMA	EN22CS301110
02	ANJALI JAISWAL	EN22CS301132
03	ANKIT PARMAR	EN22CS301135
04	ANSH LAAD	EN22CS301141
05	APOORVA GUPTA	EN22CS301190
06	ARYA BHAWSAR	EN22CS301209

Group Name: 12D2

Project Number: GAI-24

Industry Mentor Name:

University Mentor Name: Dr. Hemlata Patel

Academic Year: 2025-2026

Table of Contents

1. Introduction:-

The rapid advancement of Artificial Intelligence (AI) and Generative AI (GenAI) technologies has significantly transformed the education sector by enabling intelligent, interactive, and personalized learning solutions. Traditional education systems primarily rely on text-based learning materials, which may not be equally effective for all learners, especially visual learners. To overcome this limitation, this project proposes a **Multimodal GenAI Education System** that integrates both **text and image generation models** to deliver comprehensive educational content.

The proposed system allows users to enter an educational topic and automatically generates **detailed textual explanations along with relevant visual representations** such as diagrams, flashcards, and illustrations. By combining multiple learning modalities, the system enhances conceptual understanding, learner engagement, and knowledge retention. This document presents the **High-Level Design (HLD)** of the system, detailing its architecture, workflow, components, and design considerations.

1.1 Scope of the Document

The scope of this document is to describe the **High-Level Design (HLD)** of the Multimodal GenAI Education System. It provides an overview of the system architecture, functional workflow, and major system components.
Specifically, this document covers:

- Overall system architecture and design
- Application workflow and processing logic
- Component-level design
- Information and data flow
- API interactions
- Non-functional requirements including performance, scalability, and security

This document does not include detailed implementation-level specifications, source code, or deployment configurations.

1.2 Intended Audience

This document is intended for the following stakeholders:

- **Students:** To understand the system design, workflow, and project implementation approach.
- **Faculty Members and Academic Mentors:** To evaluate the architecture, learning objectives, and academic relevance.
- **Developers:** To gain insight into system components and integration mechanisms.
- **Project Evaluators:** For assessment, validation, and grading of the project work.

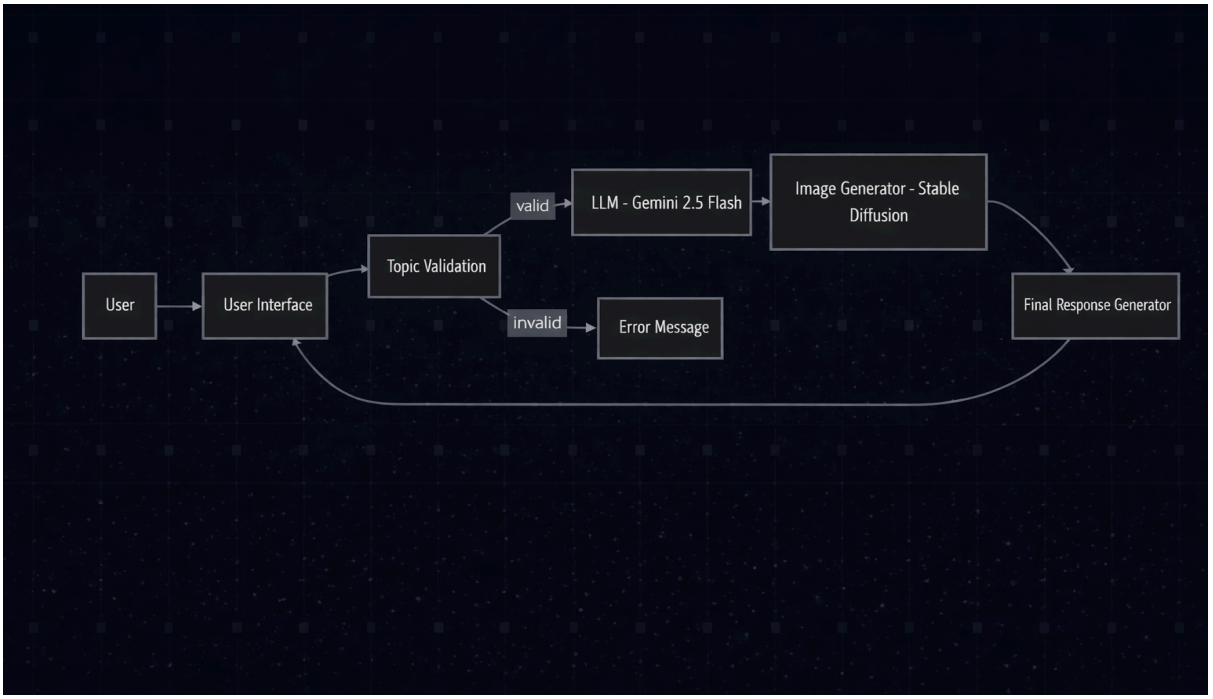
1.3 System Overview

The Multimodal GenAI Education System is designed to generate **both textual explanations and visual educational content** from a single user input. The system integrates **Large Language Models (LLMs)** and **Image Generation Models** to provide an enhanced learning experience.

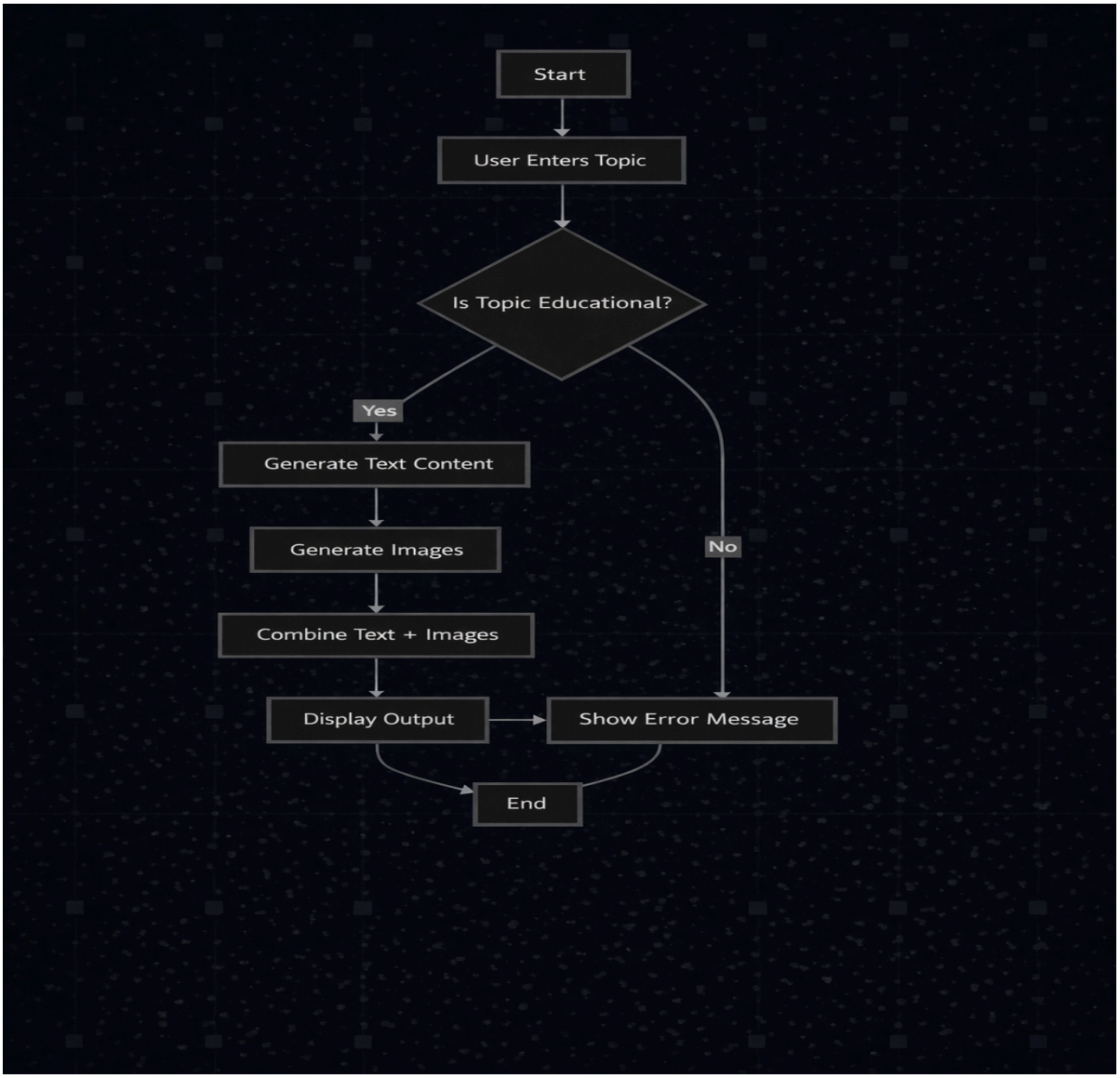
2. System Workflow:

The user enters an **educational topic** through the User Interface (UI).
The system validates whether the topic is **education-related.**
If valid, the topic is forwarded to the **LLM (Gemini 2.5 Flash)**, which generates structured textual content including explanations, key points, and examples.
The generated textual output is sent to the **Image Generation Model (Stable Diffusion)** to produce relevant diagrams, illustrations, and flashcards.
The system integrates both **text and image outputs** and presents them to the user as a **final multimodal educational response**.

2.1Application Design



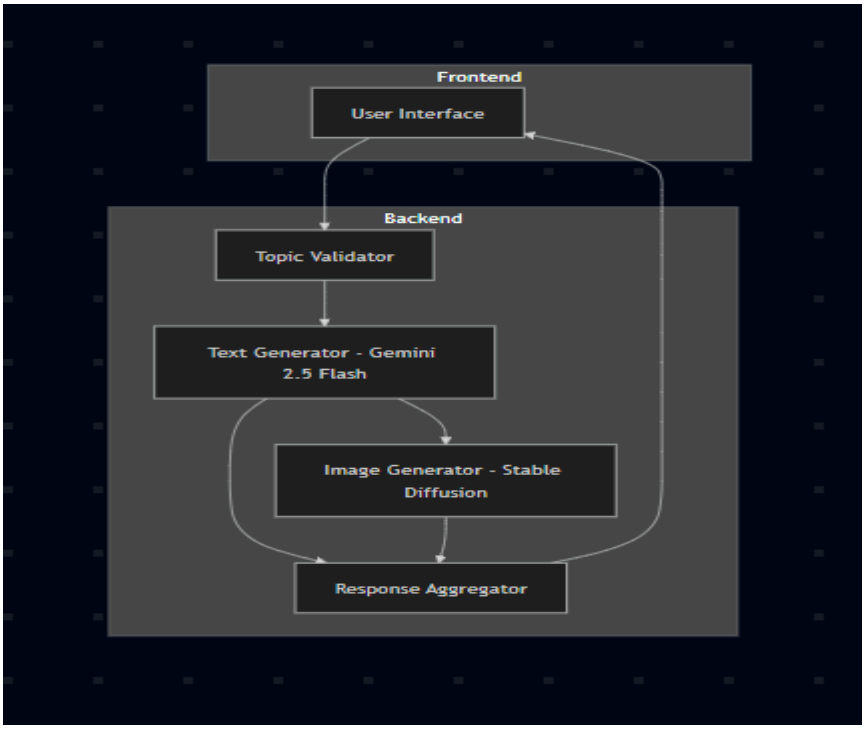
2.2 Process Flow



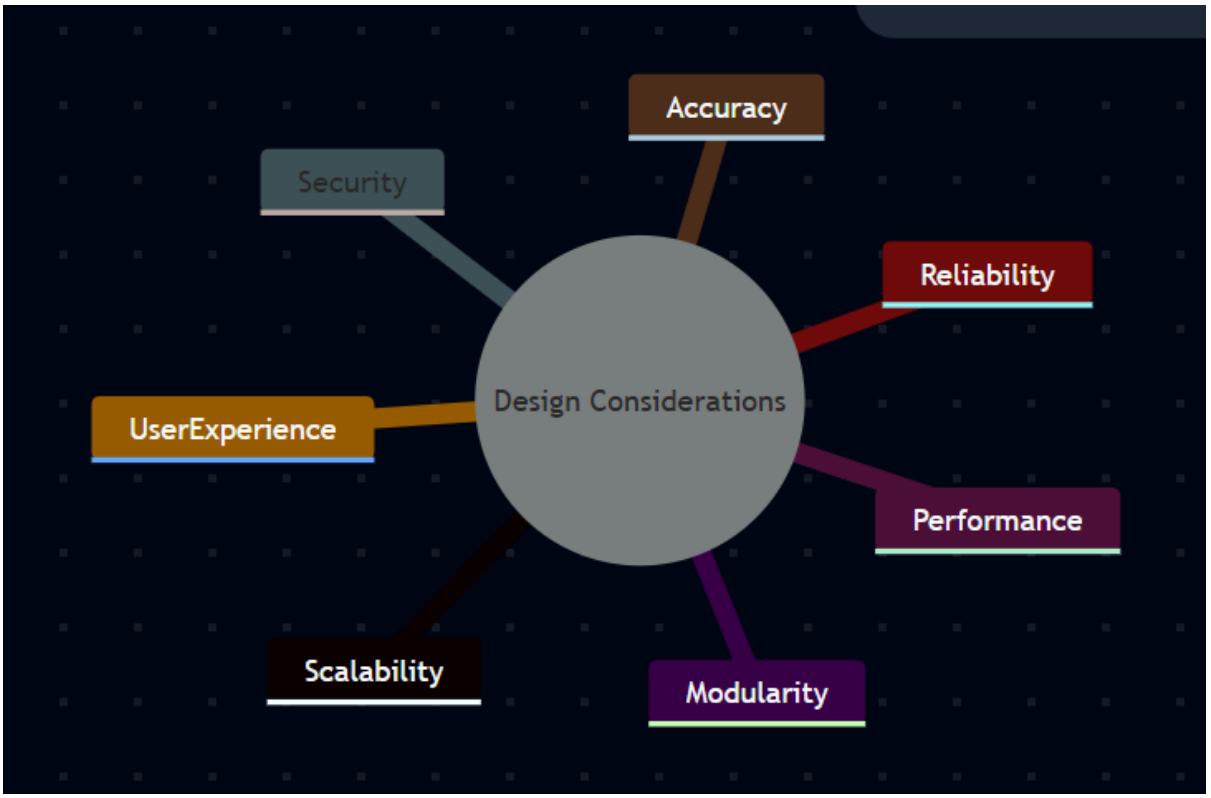
2.3 Information flow

User Input
↓
Prompt Construction
↓
LLM API Call
↓
Text Response
↓
Diffusion Model Call
↓
Image Output
↓
UI Rendering

2.4 Components Design



2.5 Key Design Considerations



2.6 API Catalogue.

1 Text Generation API

Provider: Google Gemini
Model: gemini-2.5-flash

Input:

- Topic
- Prompt Template

Output:

- Structured explanation text

2 Image Generation API

Provider: Stable Diffusion

Input:

- Topic description

Output:

- Generated educational image

3. Interfaces

This section describes the interfaces between different system components including the **User Interface, LLM services, and Image Generation services**.

Types of Interfaces:

- User Interface (Web-based Frontend)
- LLM Integration Interface
- Image Generation Interface

External Interfaces

- Google Gemini API
- Stable Diffusion API

Internal Interfaces

- Function-based communication
- Module imports

Multimodal Education Creator

Enter an Educational Concept

Generate Learning Content

localhost:8501/#multimodal-education-creator

Deploy

Multimodal Education Creator

Enter an Educational Concept

Science

Generate Learning Content

Explanation

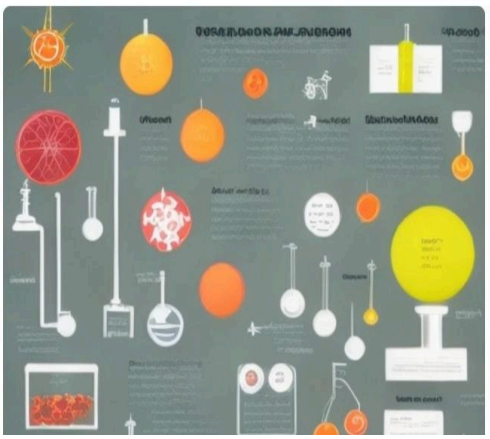
Okay, let's explain "Science" for a student!

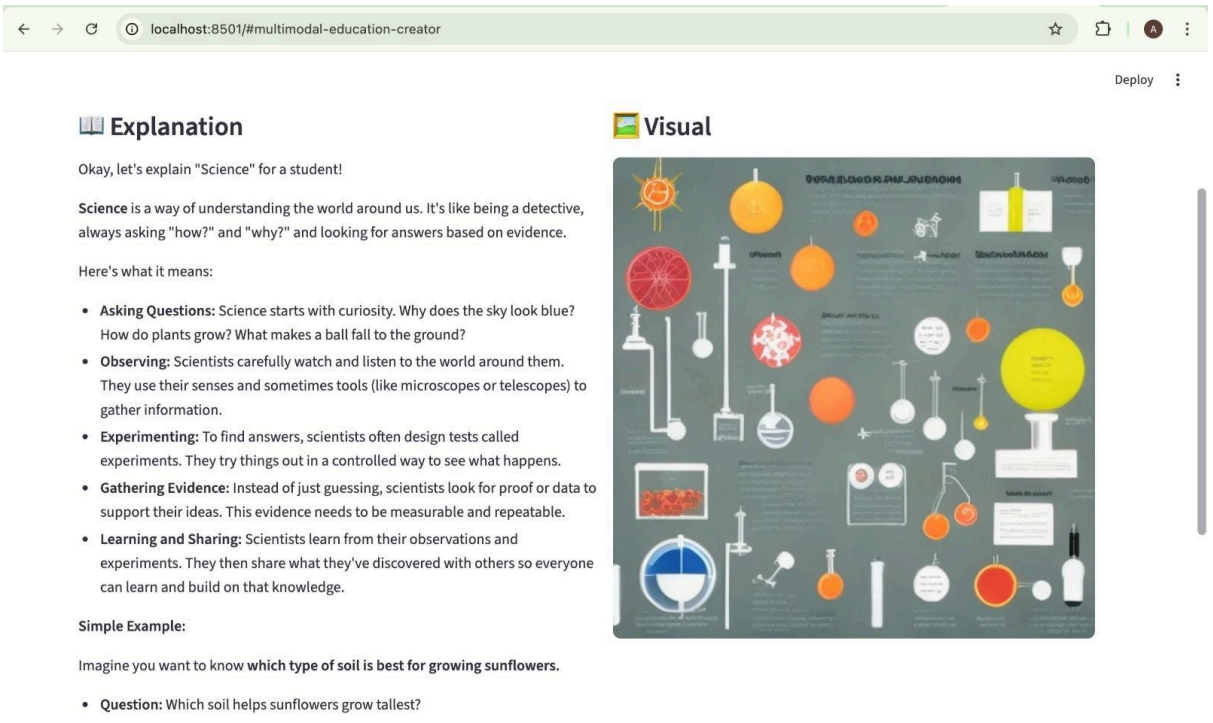
Science is a way of understanding the world around us. It's like being a detective, always asking "how?" and "why?" and looking for answers based on evidence.

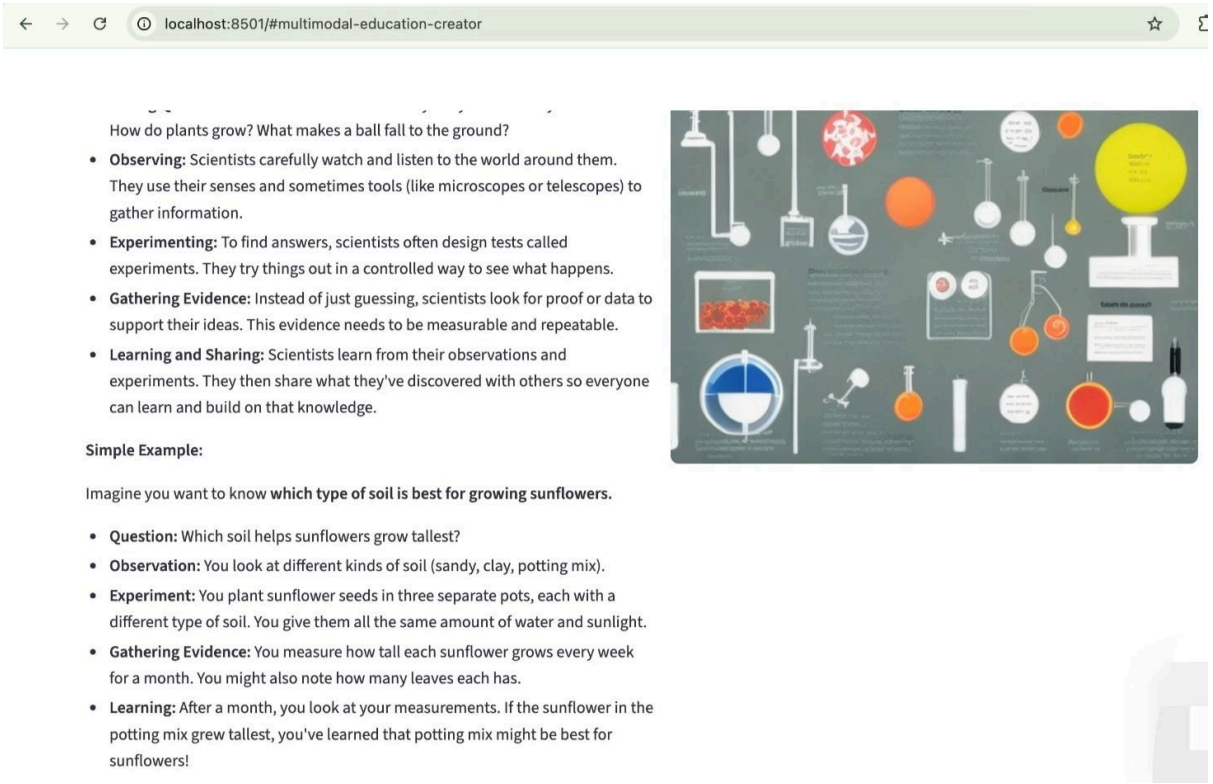
Here's what it means:

- **Asking Questions:** Science starts with curiosity. Why does the sky look blue? How do plants grow? What makes a ball fall to the ground?
- **Observing:** Scientists carefully watch and listen to the world around them. They use their senses and sometimes tools (like microscopes or telescopes) to gather information.
- **Experimenting:** To find answers, scientists often design tests called experiments. They try things out in a controlled way to see what happens.
- **Gathering Evidence:** Instead of just guessing, scientists look for proof or data to support their ideas. This evidence needs to be measurable and repeatable.

Visual




localhost:8501/#multimodal-education-creator
Deploy
Explanation
Okay, let's explain "Science" for a student!
Science is a way of understanding the world around us. It's like being a detective, always asking "how?" and "why?" and looking for answers based on evidence.
Here's what it means:
Asking Questions: Science starts with curiosity. Why does the sky look blue? How do plants grow? What makes a ball fall to the ground?
Observing: Scientists carefully watch and listen to the world around them. They use their senses and sometimes tools (like microscopes or telescopes) to gather information.
Experimenting: To find answers, scientists often design tests called experiments. They try things out in a controlled way to see what happens.
Gathering Evidence: Instead of just guessing, scientists look for proof or data to support their ideas. This evidence needs to be measurable and repeatable.
Learning and Sharing: Scientists learn from their observations and experiments. They then share what they've discovered with others so everyone can learn and build on that knowledge.
Simple Example:
Imagine you want to know which type of soil is best for growing sunflowers.
Question: Which soil helps sunflowers grow tallest?
Visual


localhost:8501/#multimodal-education-creator
How do plants grow? What makes a ball fall to the ground?
Observing: Scientists carefully watch and listen to the world around them. They use their senses and sometimes tools (like microscopes or telescopes) to gather information.
Experimenting: To find answers, scientists often design tests called experiments. They try things out in a controlled way to see what happens.
Gathering Evidence: Instead of just guessing, scientists look for proof or data to support their ideas. This evidence needs to be measurable and repeatable.
Learning and Sharing: Scientists learn from their observations and experiments. They then share what they've discovered with others so everyone can learn and build on that knowledge.
Simple Example:
Imagine you want to know which type of soil is best for growing sunflowers.
Question: Which soil helps sunflowers grow tallest?
Observation: You look at different kinds of soil (sandy, clay, potting mix).
Experiment: You plant sunflower seeds in three separate pots, each with a different type of soil. You give them all the same amount of water and sunlight.
Gathering Evidence: You measure how tall each sunflower grows every week for a month. You might also note how many leaves each has.
Learning: After a month, you look at your measurements. If the sunflower in the potting mix grew tallest, you've learned that potting mix might be best for sunflowers!

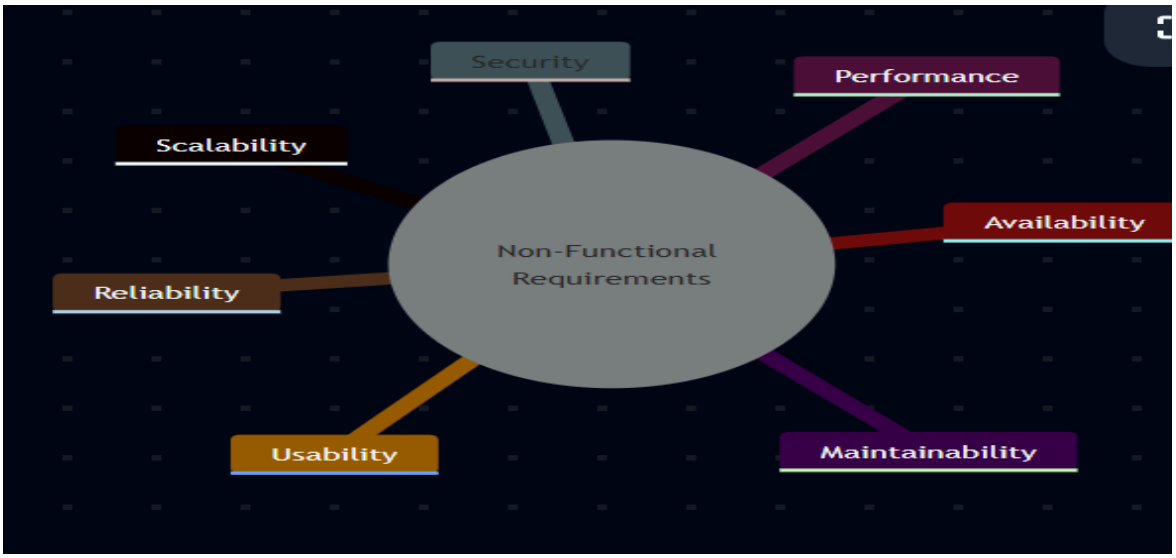
4 Non-Functional Requirements

Non-functional requirements define the **quality attributes** of the system.

Key NFRs:

Scalability
Performance
Security
Reliability
Availability
Maintainability
Usability

NFR Mindmap



4.1 Security Aspects

- API keys stored in .env file
- No hardcoding of credentials
- HTTPS-based API communication
- No persistent user data storage

4.2 Performance Aspects

Performance ensures fast response time and smooth user experience.

- Lightweight UI framework
- Prompt optimization
- Model caching
- Efficient API calls
- Minimal runtime memory usage

5. References

Gemini 2.5 Flash – Official Documentation
Stable Diffusion – Model Architecture & API Guide
Research Papers on Multimodal Learning Systems
Software Engineering – System Design and Architecture Textbook.
IEEE Journals on AI in Education