

# MultiModel Education Creator Report

Course Name: GenAI

**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*

## Problem Statement & Objectives

### 1. Problem Statement

Traditional learning systems often rely heavily on either textual explanations or static visual materials, which may not be sufficient for effective understanding of complex concepts. Many learners struggle to grasp abstract or scientific topics when presented only in text form, while others require visual representations to fully comprehend the material. Additionally, existing educational tools lack the capability to dynamically generate personalized and context-aware content. There is a need for an intelligent system that can automatically generate both **textual explanations and visual representations** for educational topics in real time.

Hence, the problem is to design and develop a system that:

- Accepts educational topics from users
- Validates the relevance of the topic
- Generates meaningful textual explanations
- Creates corresponding visual content
- Delivers both outputs together for enhanced learning

### 2. Project Objectives

The primary objective of this project is to develop a **Multimodal Generative AI-based Education System** that improves learning through both text and visual content.

#### Specific Objectives:

- To build a system that accepts educational topics as input
- To implement a validation mechanism to filter non-educational topics
- To generate structured textual explanations using an LLM
- To generate visual content such as diagrams, flashcards, and illustrations
- To integrate text and image outputs into a single response
- To enhance conceptual understanding through multimodal learning
- To provide a user-friendly interface for seamless interaction

### 3. Scope of the Project

The scope of this project focuses on applying **Generative AI and Multimodal Learning** in the education domain.

#### Key Scope Areas:

##### Educational Domain Focus

The system is designed to process only educational topics such as science, mathematics, history, and technology.

##### Multimodal Content Generation

It generates both textual explanations and visual representations for better understanding.

### Automated Content Creation

The system automatically produces learning material without human intervention.

#### Use Cases

- E-learning platforms
- Smart tutoring systems
- Self-learning applications
- Educational content generation tools

#### Scalability

The system can be extended to include:

- More subjects
- Interactive quizzes
- Voice-based learning
- Personalized recommendations

#### Limitations

- Only supports educational topics
- Image quality depends on the generation model
- Requires internet and computational resources

## Proposed Solution

The proposed solution is a **Multimodal Generative AI-based Education System** designed to enhance learning by providing both textual explanations and visual representations of educational topics. The system integrates advanced AI models to process user input, validate its relevance, and generate meaningful educational content in multiple formats.

The workflow begins when the user enters an educational topic through the interface. The system first validates whether the topic is relevant to education. If valid, it is processed by a language model to generate structured textual explanations. This text is then used by an image generation model to create visual representations such as diagrams and flashcards. Finally, both outputs are combined and presented to the user, ensuring an interactive and comprehensive learning experience.

### 1. Key Features

- Topic validation to ensure educational relevance
- Automatic text content generation
- Image and visual content generation
- Multimodal output (text + images)
- User-friendly interface
- Structured learning content (concepts, examples, key points)
- Fast and automated content generation
- Scalable architecture

### 2. Overall Architecture / Workflow

The system follows a step-by-step workflow:

#### User Interface

The user inputs an educational topic (e.g., Photosynthesis, Newton's Laws).

**Topic Validation Module**

The system checks whether the entered topic is related to education.

- If valid → forward to the next stage
- If invalid → return an error message

**Text Generation Module (LLM)**

A language model processes the input and generates:

- Concept explanations
- Key points
- Examples
- Structured content

**Image Generation Module**

The generated text is passed to an image generation model, which produces:

- Diagrams
- Visual explanations
- Flashcards
- Illustrations

**Response Integration Module**

The system combines the generated text and images.

**Final Output**

The user receives a complete response containing both textual and visual content.

### 3. Tools & Technologies Used (If applicable)

#### Programming Languages

Python (for backend development and AI integration)

#### AI/ML Models

Text Generation Model (LLM) for content creation

Image Generation Model for visual outputs

#### Frameworks & Libraries

Streamlit for frontend interface

#### AI Platforms

Generative AI APIs for text generation

Image generation frameworks (e.g., diffusion-based models)

#### Other Tools

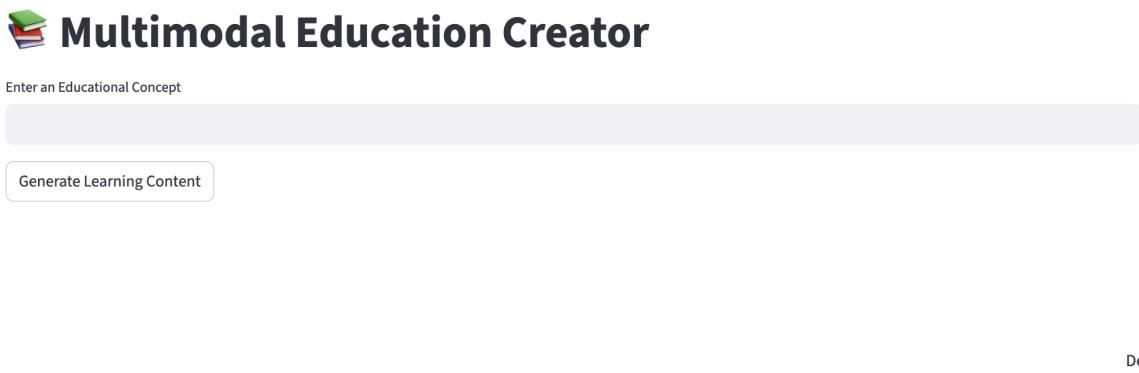
Git/GitHub for version control

Vs code for experimentation

## Results & Output

The developed **Multimodal GenAI Education System** was successfully implemented and tested with various educational topics. The system was able to generate both **textual explanations and visual representations**, providing a comprehensive learning experience. The results demonstrate the effectiveness of integrating multiple AI models for educational purposes.

### 1. Screenshots / Outputs



**Multimodal Education Creator**

Enter an Educational Concept

Generate Learning Content

Deploy

#### Explanation

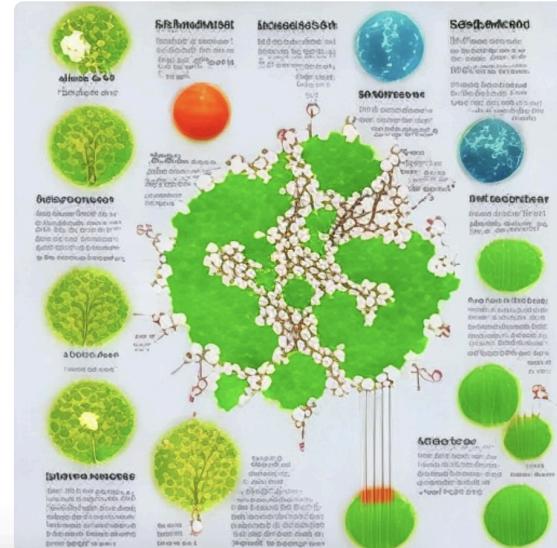
Okay, since "Photosynthesis" is a science topic commonly taught in education, I can help you with that!

Here's an explanation of Photosynthesis for a student:

Photosynthesis is the amazing process that plants use to make their own food! Think of plants as little chefs, and photosynthesis is their cooking recipe.

- **What it is:** Photosynthesis is how green plants (and some other organisms) turn light energy into chemical energy, which they store as sugar (their food).
- **Where it happens:** This process mainly happens in the leaves of plants, inside tiny parts called chloroplasts, which contain a green pigment called chlorophyll. That's why leaves are green!
- **What plants need (The Ingredients):**
  - **Sunlight:** They capture energy from the sun.
  - **Water ( $H_2O$ ):** They absorb water from the soil through their roots.

#### Visual



- **Carbon Dioxide (CO<sub>2</sub>):** They take in this gas from the air through tiny pores in their leaves called stomata.
- **What plants make (The Products):**
  - **Glucose (Sugar):** This is the plant's food, giving it energy to grow, flower, and produce fruits.
  - **Oxygen (O<sub>2</sub>):** This is released back into the air, and it's what we humans and many other animals breathe to live!

**Simple Example:**

Imagine you have a small tomato plant in your garden. The sun shines on its green leaves. The plant's roots suck up water from the soil. Invisible carbon dioxide from the air enters the leaves. Inside the leaves, the plant uses the sunlight to "cook" the water and carbon dioxide, turning them into sweet glucose (sugar) to help the tomato plant grow big and produce yummy tomatoes. As a bonus, it also releases oxygen into the air for you to breathe while you're admiring your plant!

## 2 Reports / Dashboards / Models

### Multimodal Education Creator

Enter an Educational Concept

boiouou

Generate Learning Content

#### Explanation

Sorry, this topic is not supported. Please ask about education-related topics.

#### Visual

No visual generated for this topic.

## 3. Key Outcomes

The project achieved the following key outcomes:

- Successfully developed a **multimodal AI system** for education
- Generated **accurate and structured textual explanations**
- Produced **relevant visual content** for better understanding
- Improved **concept clarity through combined text and visuals**
- Ensured **validation of educational topics** before processing
- Demonstrated the **integration of multiple AI models** in one system
- Enhanced **user engagement and learning experience**

## Conclusion

The **Multimodal GenAI Education System** successfully demonstrates how artificial intelligence can enhance the learning experience by combining both textual and visual content generation. The system takes an educational topic as input, validates its relevance, and generates structured explanations using a language model. It further enhances understanding by producing corresponding visual representations such as diagrams and flashcards.

By integrating text generation and image generation models, the project showcases the power of **multimodal AI** in simplifying complex concepts and making learning more interactive and engaging. The system ensures that users receive comprehensive learning material in a single response, improving both clarity and retention.

Through this project, key learnings include:

- Understanding of **Generative AI models** and their applications
- Integration of multiple AI models in a single workflow
- Designing scalable system architecture
- Handling real-world problems using AI solutions
- Improving user experience through intelligent automation

Overall, the project highlights the potential of AI-driven solutions in transforming modern education systems and making learning more accessible and effective.

## Future Scope & Enhancements

The current system provides a strong foundation for multimodal learning; however, several enhancements can further improve its capabilities:

### 1. Advanced Personalization

- Customize content based on user learning level (beginner, intermediate, advanced)
- Provide adaptive learning paths

### 2. Interactive Learning Features

- Add quizzes, MCQs, and assessments
- Provide instant feedback and performance tracking

### 3. Voice-Based Interaction

- Enable speech-to-text and text-to-speech features
- Support voice-assisted learning

### 4. Improved Visual Generation

- Enhance image quality and accuracy
- Generate more detailed and subject-specific diagrams

### 5. Multilingual Support

- Provide explanations in multiple languages
- Increase accessibility for diverse users

## 6. Real-Time Collaboration

Allow teachers and students to interact within the platform  
Enable content sharing and discussion

## 7. Integration with E-learning Platforms

Connect with Learning Management Systems (LMS)  
Provide seamless educational experiences

## 8. Offline Support

Enable limited functionality without internet access

## 9. Analytics & Insights

Track user progress and learning behavior  
Provide performance analytics and reports