



AI Multimodal Communicator:

An AI streamlit Web Application for Conversations, Image Captioning, and PDF Interaction

Team Member: Komal Suthar

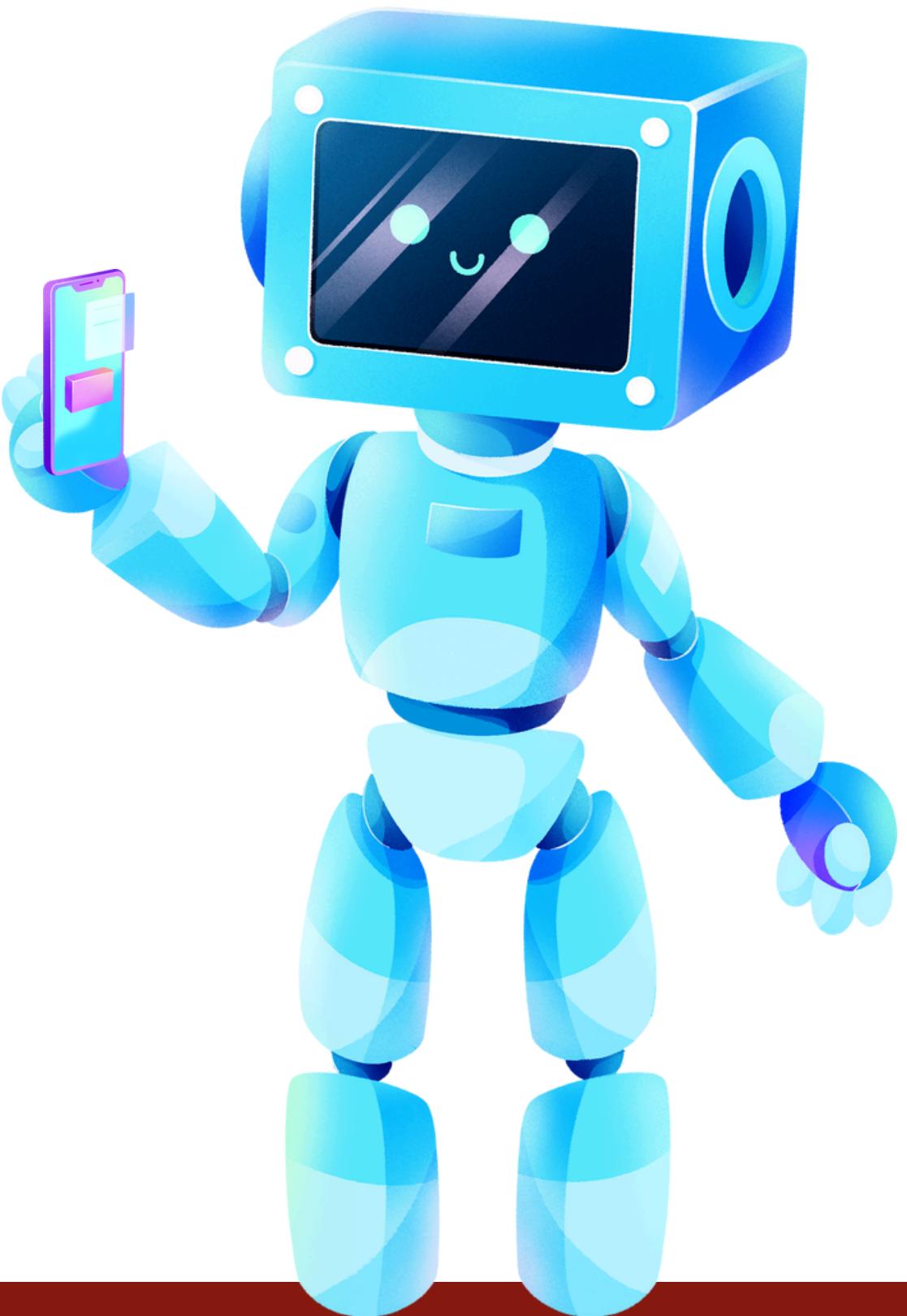
Guided By: Abdul Aziz md
Master Trainer
Edunet Foundation

OUTLINE

- Abstract
- Problem Statement
- Aims, Objective & Proposed System/Solution
- System Design/Architecture
- System Development Approach (Technology Used)
- Algorithm & Deployment
- Conclusion
- Future Scope
- References
- Video of the Project

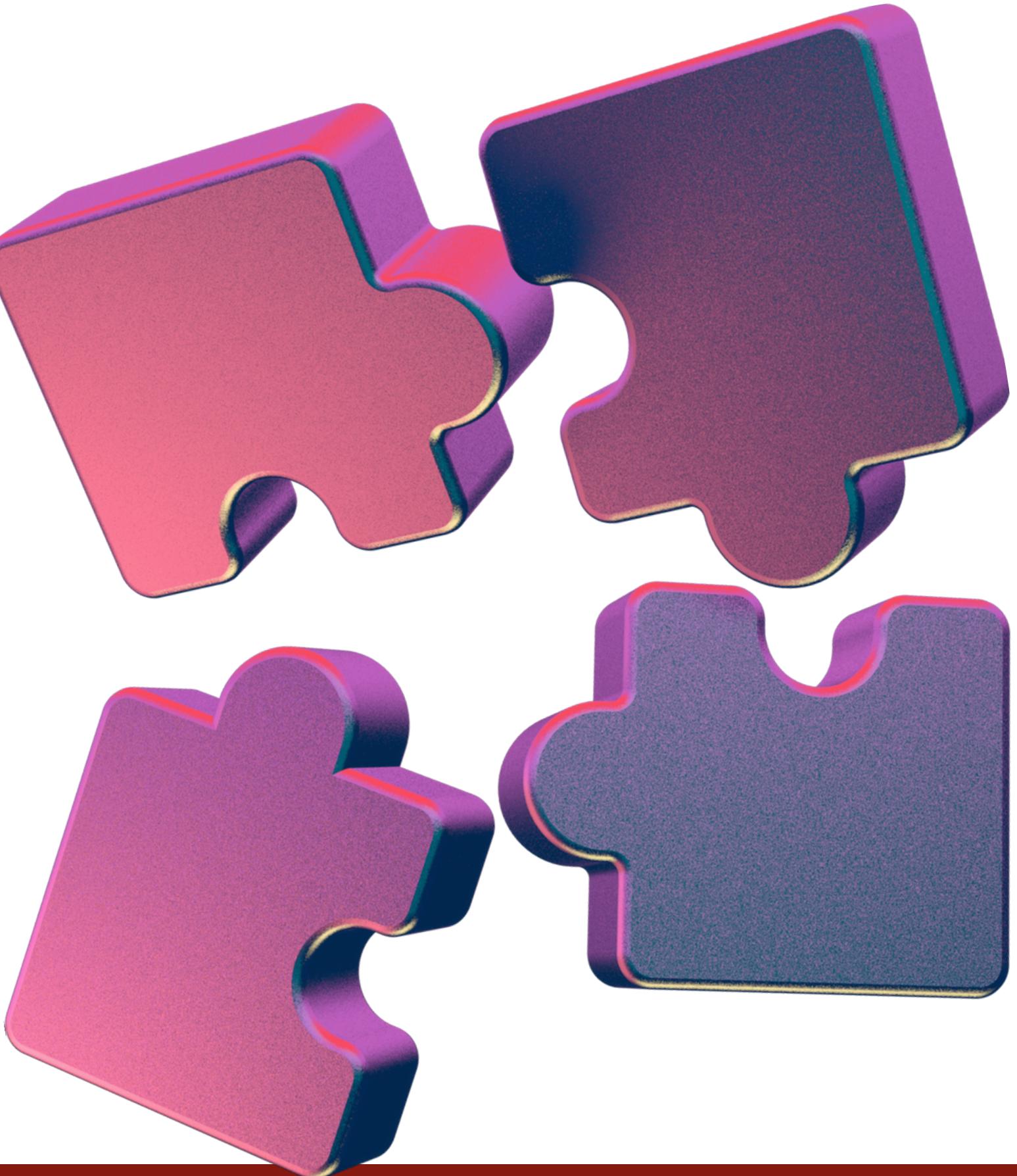
Abstract

- This project presents a generative AI system that integrates chatbot, image captioning, Ask me a Question, chat with PDF capabilities and About page.
- The system is built using Streamlit, Python, and various generative AI models.



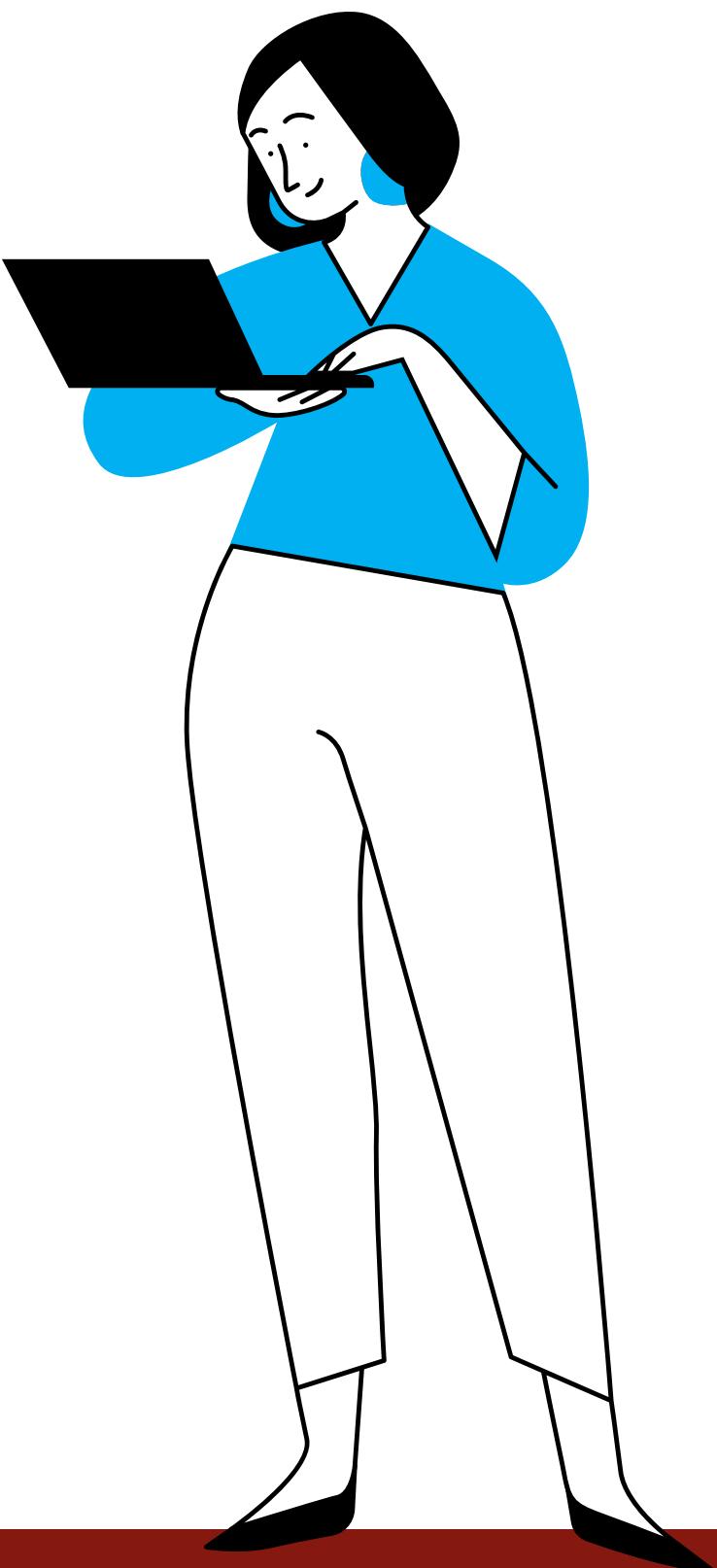
Problem Statement

- The need for a comprehensive generative AI system that can handle a variety of tasks, including natural language processing, image processing, and PDF processing.



Aim and Objective

- The aims of this project are to:
- Develop a generative AI system that can engage in natural conversations with users.
- Develop a generative AI system that can generate comprehensive captions for images.
- Develop a generative AI system that can chat with users based on text extracted from PDF documents.
- The proposed system is a web application that integrates these three capabilities into a single user-friendly interface.

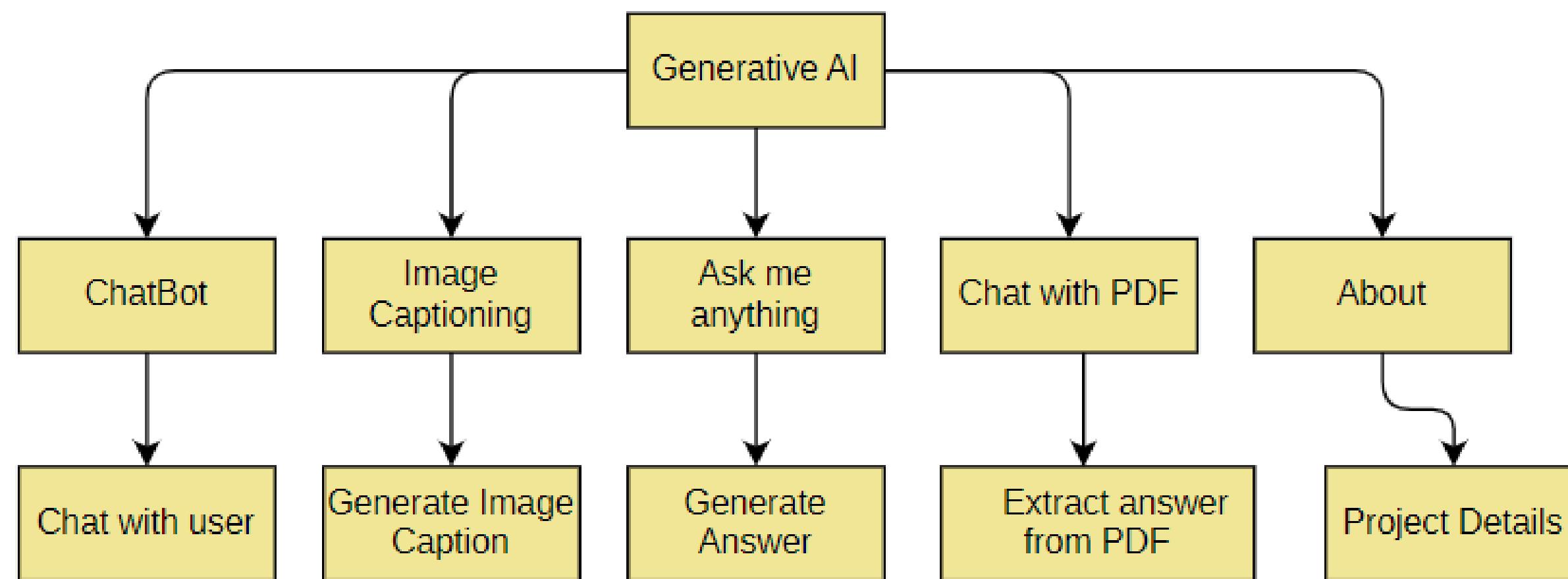


Proposed Solution

- The system will be divided into the following components:
- **Frontend:** The frontend will be responsible for handling the user interface and collecting the user's input. It will be implemented using Streamlit, a Python library that makes it easy to build and deploy web applications.
- **Backend:** The backend will be responsible for processing the user's input and generating the output. It will be implemented using Python and various generative AI models.
- **Database:** The database will store the data that is used by the system. This data will include the training data for the generative AI models, the user's input, and the system's output.



System Architecture



System Development Approach(Technology Used)

- The following technologies are used to develop the system:
- **Python**: Python is a programming language that is well-suited for machine learning and data science.
- **Streamlit**: Streamlit is a Python library that makes it easy to build and deploy web applications.
- **PyPDF2**: PyPDF2 is a Python library that allows users to read and write PDF files.
- **Langchain**: Langchain is a Python library that provides a set of tools for building and deploying conversational AI systems.
- **FAISS**: FAISS is a Python library that provides a set of tools for building and searching vector databases.



Algorithm & Deployment

- The following algorithms are used in the system:
- **Chatbot:** The system uses a generative AI model, such as Gemini-Pro, to generate text.
- **Image captioning:** The system uses a generative AI model, such as Gemini-Pro, to generate captions for images.
- **Chat with PDF:** The system uses a conversational AI chain, PyPDF2, Langchain and FAISS to provides a set of tools for building, deploying conversation AI with the user using FAISS searching vector databases.
- **Deployment:** The system is deployed on a server using Microsoft Azure. Users can access the system by visiting the web server's URL in their browser.



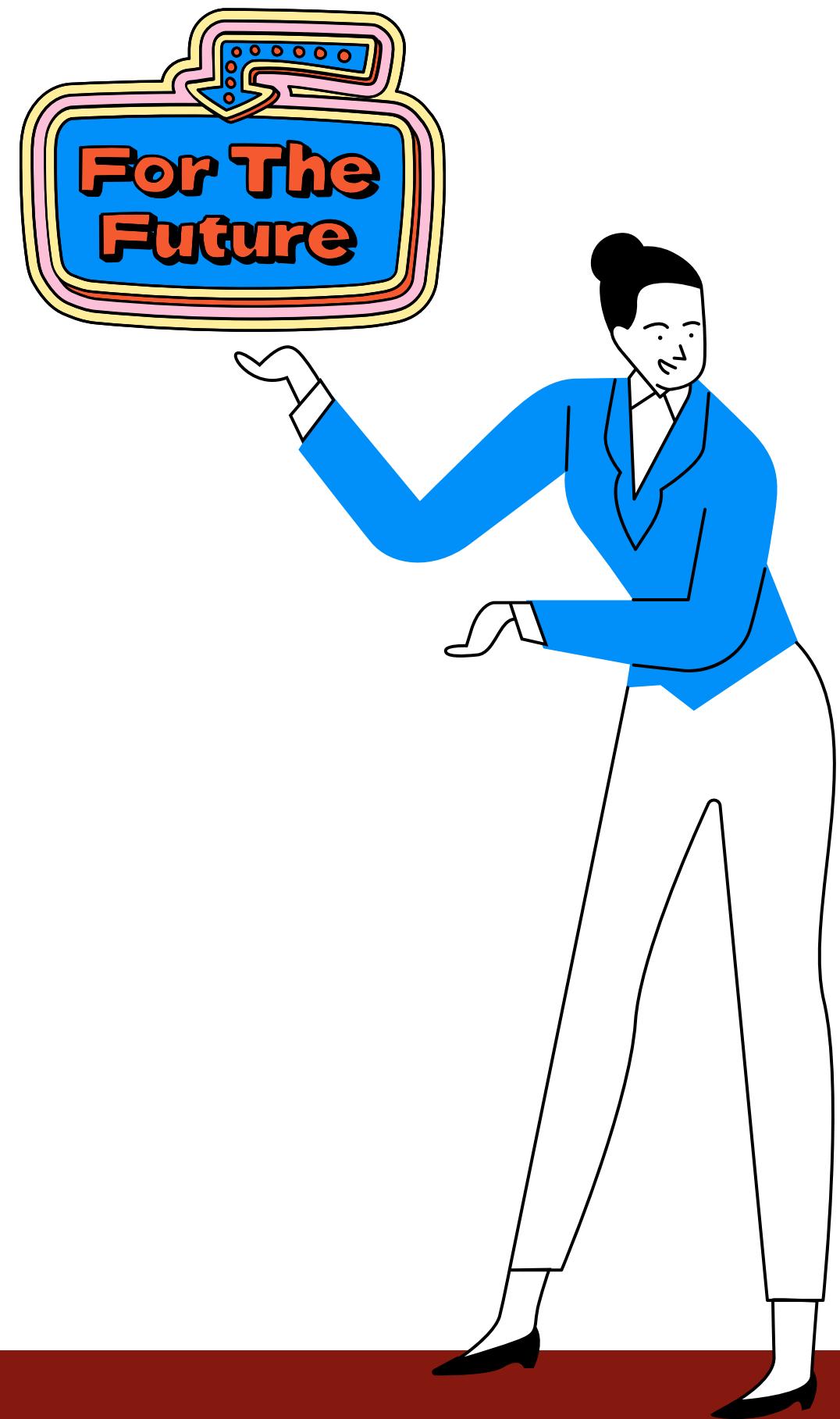
Conclusion

- This project demonstrates the power of generative AI to create a comprehensive system that can handle a variety of tasks. The system is easy to use and can be used by a wide range of users, including students, researchers, and professionals.



Future Scope

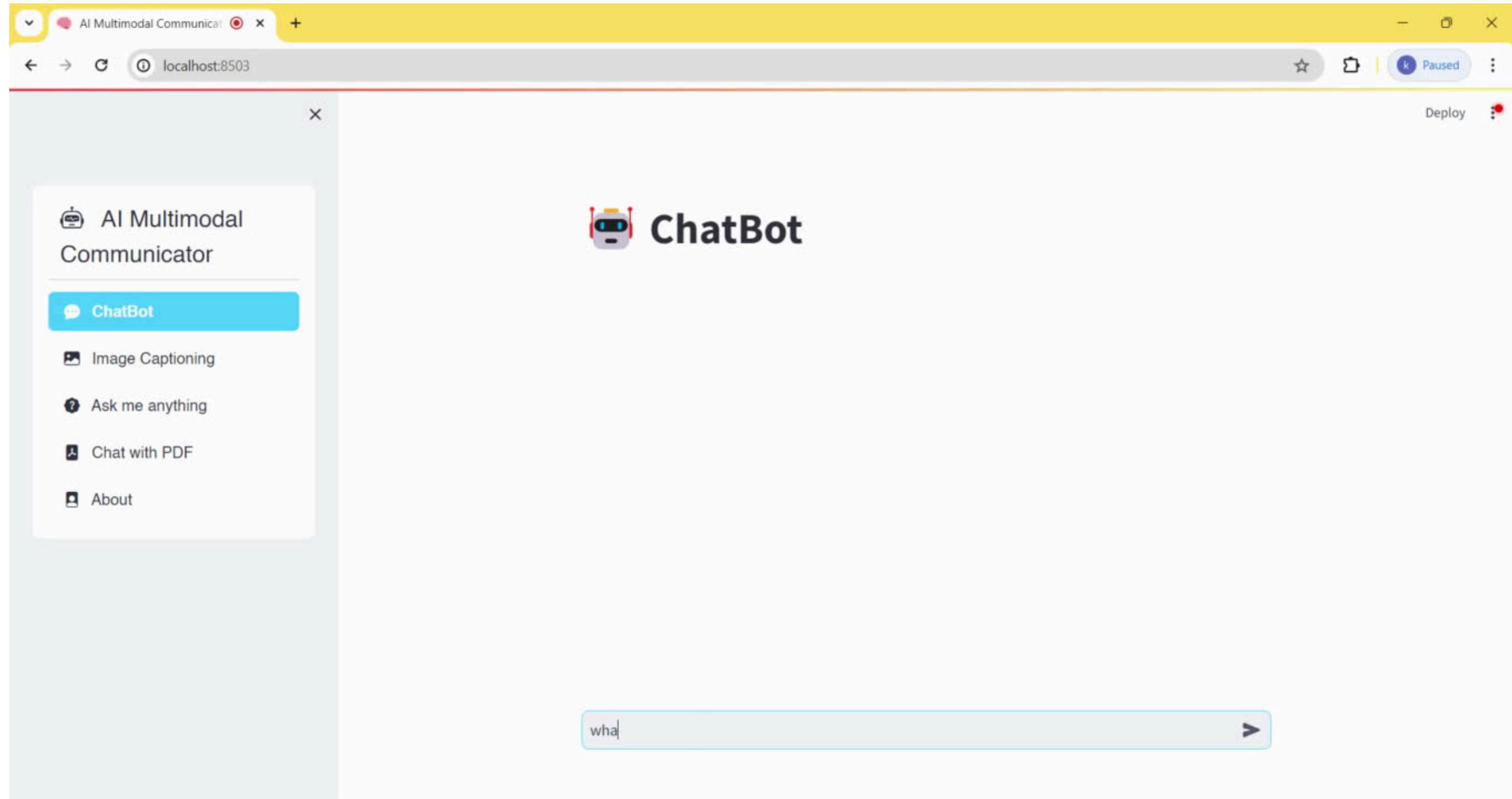
- Integrating additional generative AI models.
- Adding new categories features to perform various other task.
- Image and Video Generation.
- Explore advanced NLP techniques like named entity recognition, coreference resolution, and question classification.
- Implement safeguards and guidelines to ensure responsible and ethical use of the application.
- Developing a mobile application version of the system.



Reference

- https://youtu.be/sf5MrM0AlU?si=kBBFB_H84MxuT4VK
- <https://github.com/siddhardhan23/gemini-pro-streamlit-chatbot>

Video



THANK YOU