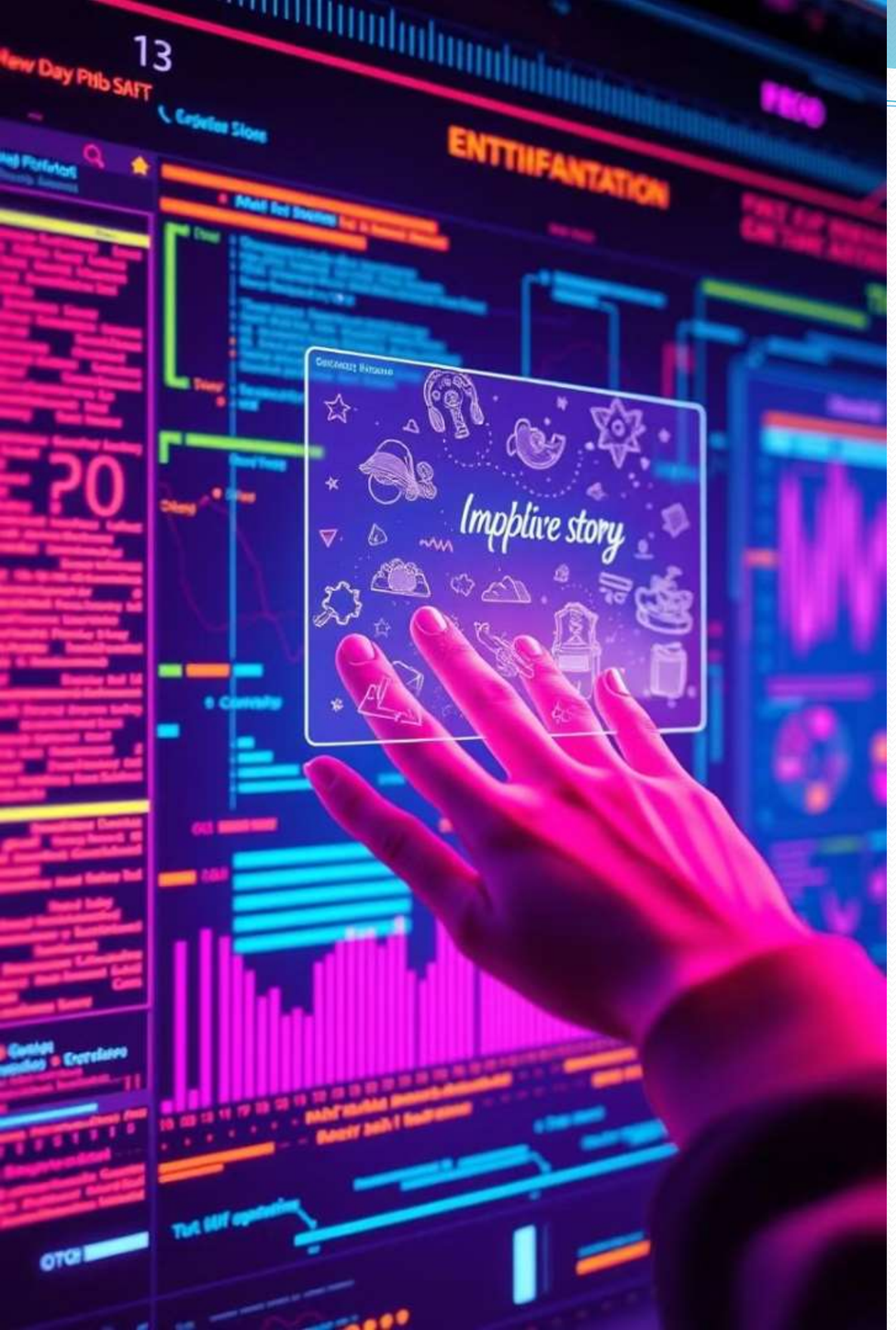


PANIMALAR ENGINEERING COLLEGE

Detectale : Story Generation with Image & Voice Recognition

Detectale is an innovative application that uses image and voice recognition to generate engaging and interactive stories. It combines cutting-edge AI with creative storytelling to bring your imagination to life.

By G.SHALINI

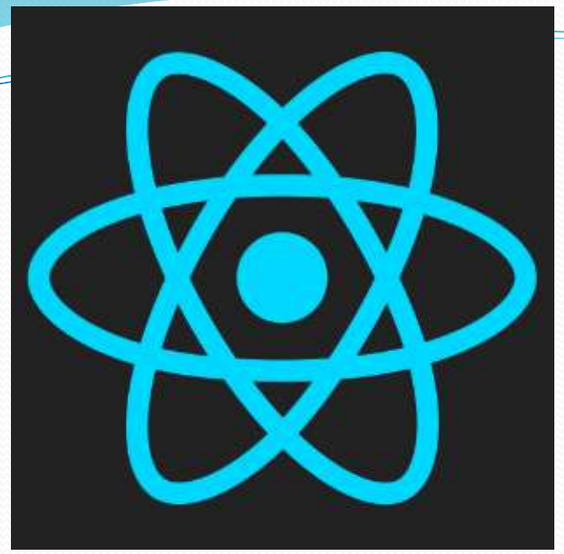
(Reg No : 211422104453)

Guide : Sophana Jenifer

Batch : C-27

Abstract

This innovative project combines the power of AI and image recognition to generate personalized stories from text descriptions or uploaded images. By reflecting elements from users' own surroundings, the app enhances engagement and educational value, especially for children, making reading more relatable and interactive. This approach not only fosters a love for reading but also supports learning through enriched narratives. In conclusion, this project exemplifies how blending AI with storytelling can significantly elevate traditional educational methods, offering a compelling tool for the next generation. The app's seamless integration of text and image-based storytelling creates a truly immersive and captivating experience for users of all ages.



SOFTWARE CONFIGURATION

* *ReactJS*: For building the user interface. Ensure you have Node.js installed to create and manage your React app.
* *React Router:* For managing navigation within your application.
* *Axios or Fetch API* : For making API calls to the Gemini API.
* *CSS Framework* : Tailwind CSS for styling the application.
* *Image Recognition Library* : Used TensorFlow-coco-ssd .
* *Web Speech API* : For capturing voice input directly in the browser.
* *Gemini API* : Utilize the given input and generate stories .
* *Code Editor* : Visual Studio Code or any other preferred IDE.
* *Version Control* : Git and GitHub for version control and collaboration.
* *Deployment* : Netlify

HARDWARE CONFIGURATION

* *A computer with at least :*

Processor: Intel i5 or equivalent.

* *Microphone* :

A quality microphone for accurate voice input.

* *Camera* (optional) :

If you plan to capture images directly through the application, a webcam or external camera would be helpful.

* *Internet Connection* :

To use this application, Internet connectivity is needed.

EXISTING SYSTEM

Story Generation from Images Using Deep Learning

The existing system described in the paper "Story Generation from Images Using Deep Learning" employs a deep learning approach where a pre-trained model generates descriptive captions for images. This system has two main components: an image-feature extractor and a story generator. The image-feature extractor identifies objects within the image and generates a list of object names. This list is then fed into a story generator, which has been trained on a dataset of short descriptive sentences to create coherent short stories. While the model successfully generates expressive narratives with a BLEU score of 0.59, it relies on the specific training data and may lack versatility in creating varied stories across different contexts.

PROPOSED SYSTEM

Interactive Storytelling with TensorFlow COCO-SSD

The proposed system offers a novel approach to storytelling by using TensorFlow's COCO-SSD for real-time object detection and AI-driven story generation. Unlike the existing system that depends on a pre-trained story generator, this system leverages advanced AI to create versatile and dynamic narratives. Users can generate stories by either inputting text descriptions or uploading images. TensorFlow COCO-SSD identifies and analyzes objects in the images, allowing the AI to generate rich and contextually relevant stories. This method enhances the versatility of story creation, providing more personalized and engaging content tailored to the user's input, and offering a more interactive and adaptive storytelling experience.

Backend Module

generateStory.js	The backend uses a google's public API called "Gemini AI api", and Google SDK for the generation of stories in different languages with the characters mentioned by the user.

The use of AI for the backend allows to generate dynamic stories every time.

Frontend Modules

1

WriteCharacter.js

This module gets the input from user as text format and use it to generate stories.

2

ImageRecognition.js

This module uses tensorflow-coco-ssd module to recognize objects from images and generate story with those objects.

3

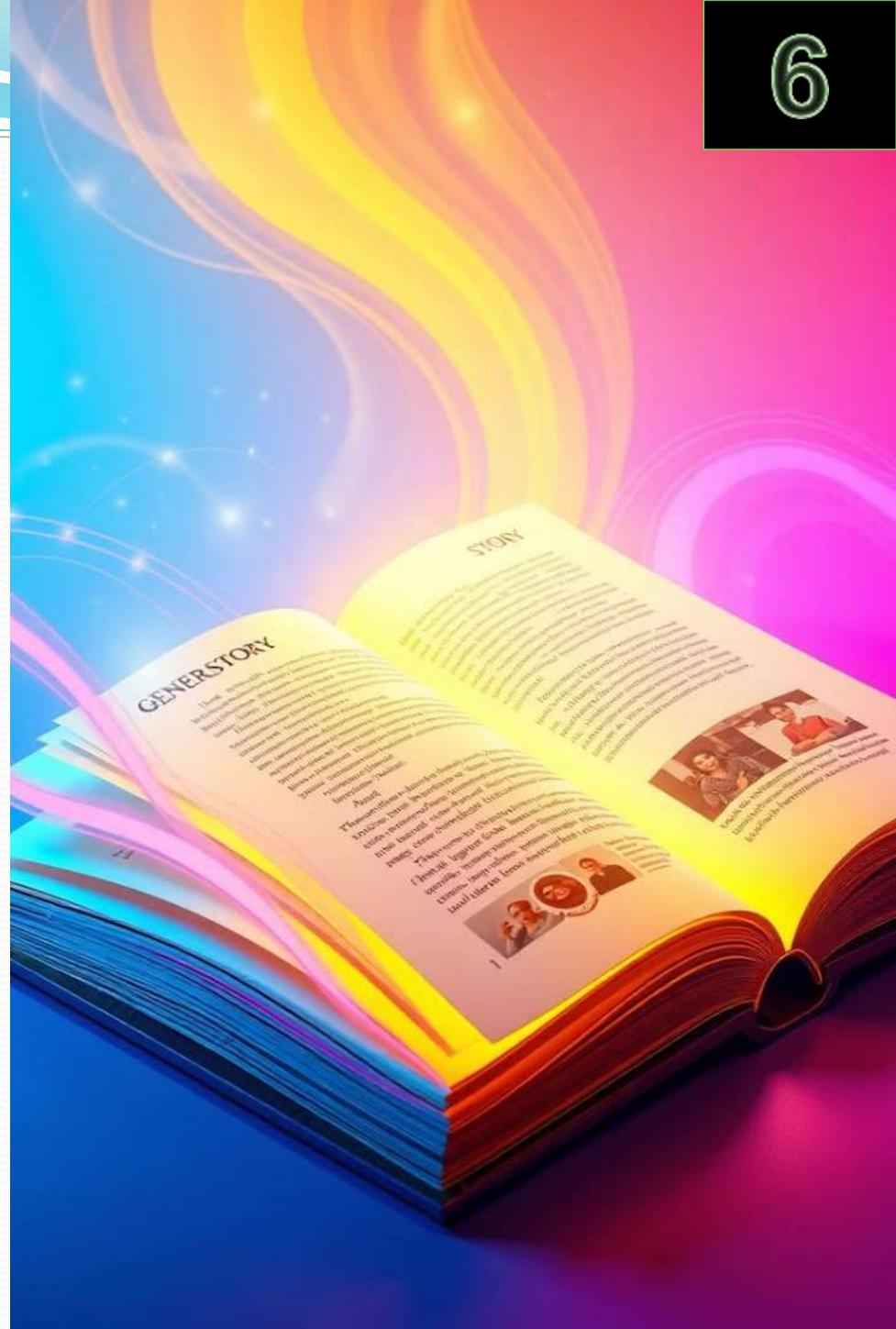
SpeakCharacter.js

This module uses react-speech library to recognize voice and takes the input characters of the story and displays on the screen .

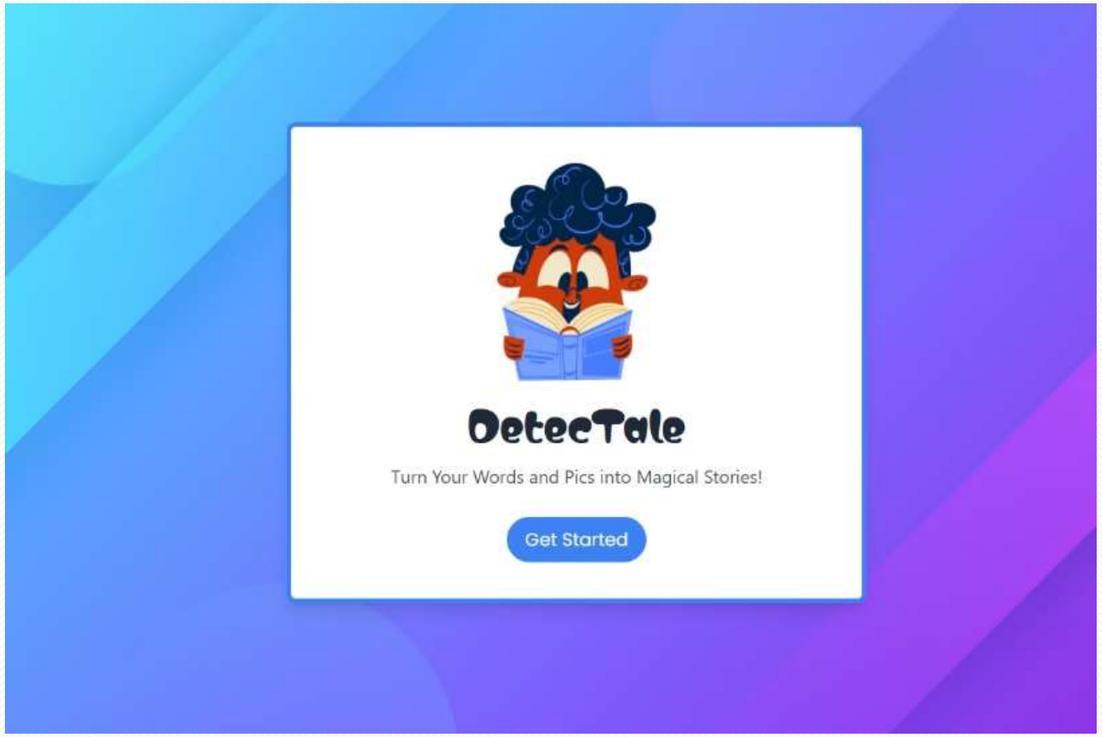
4

StoryPage.js

This page displays the story generated by the backend , with multilingual facility and dark mode functionality .

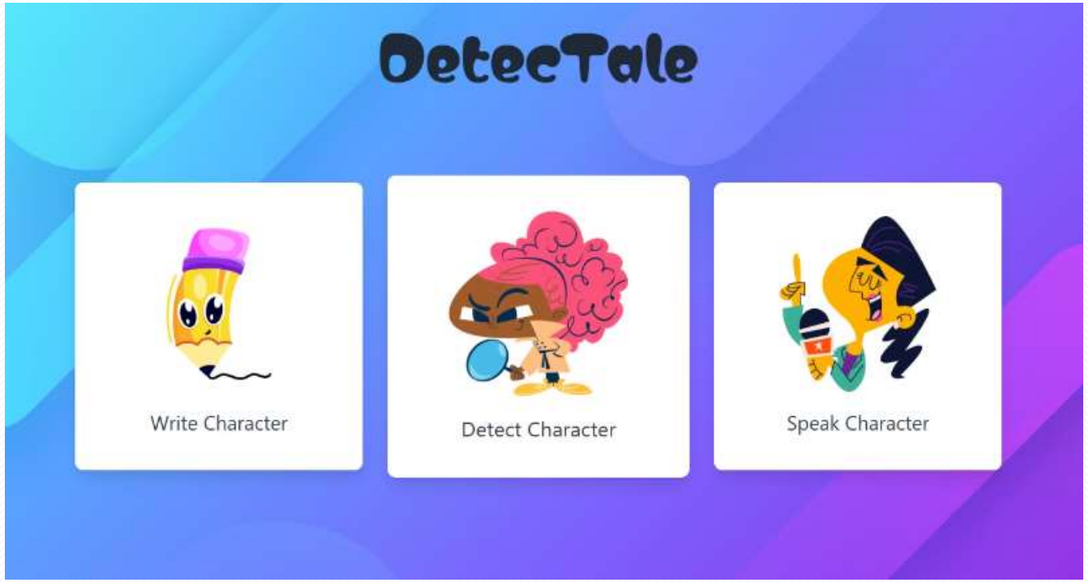


Frontend Screens



Landing Page

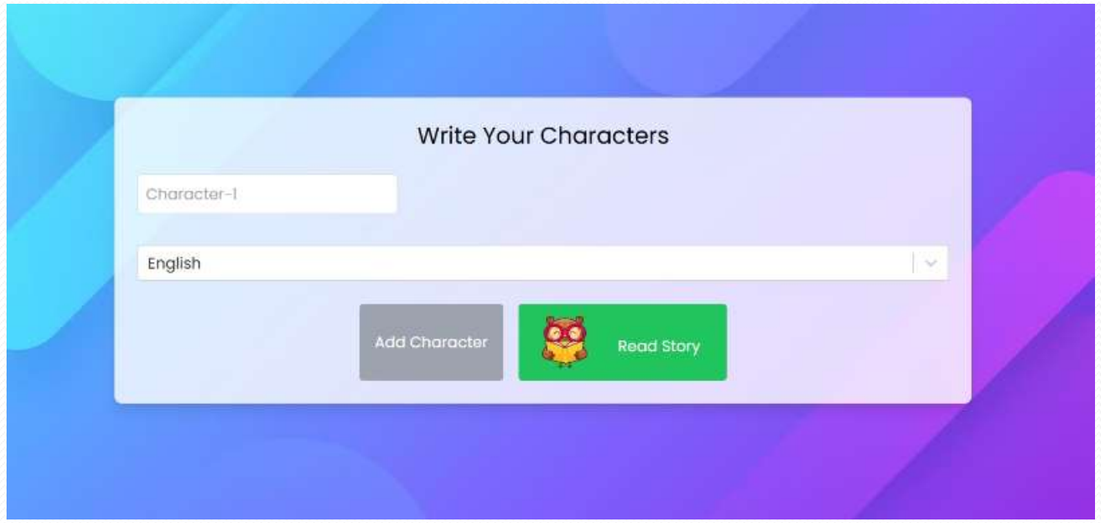
1



Home Page

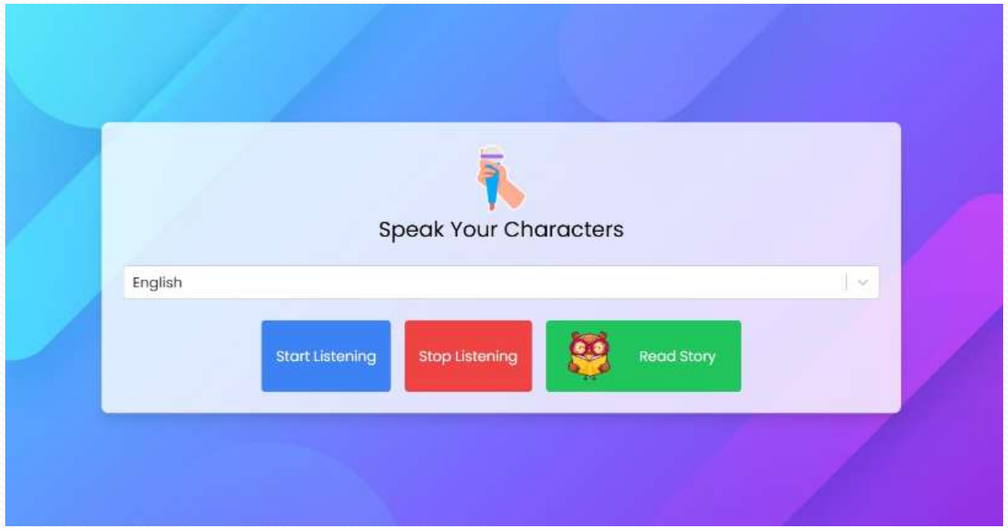
2

3

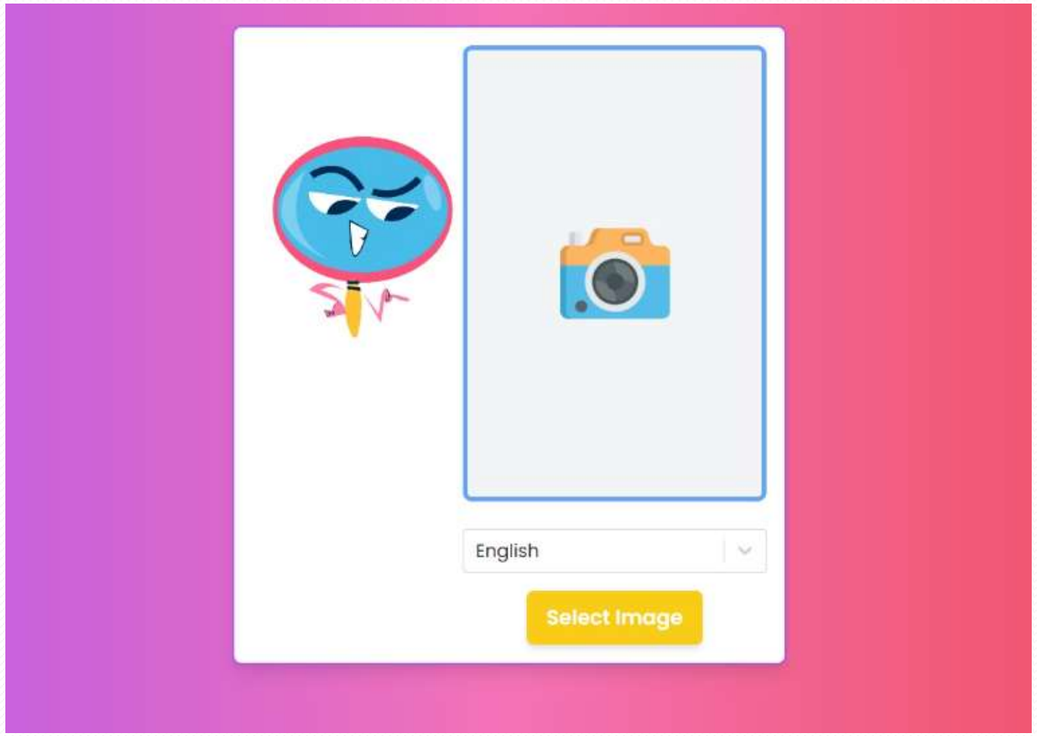


WriteCharacter Page

4

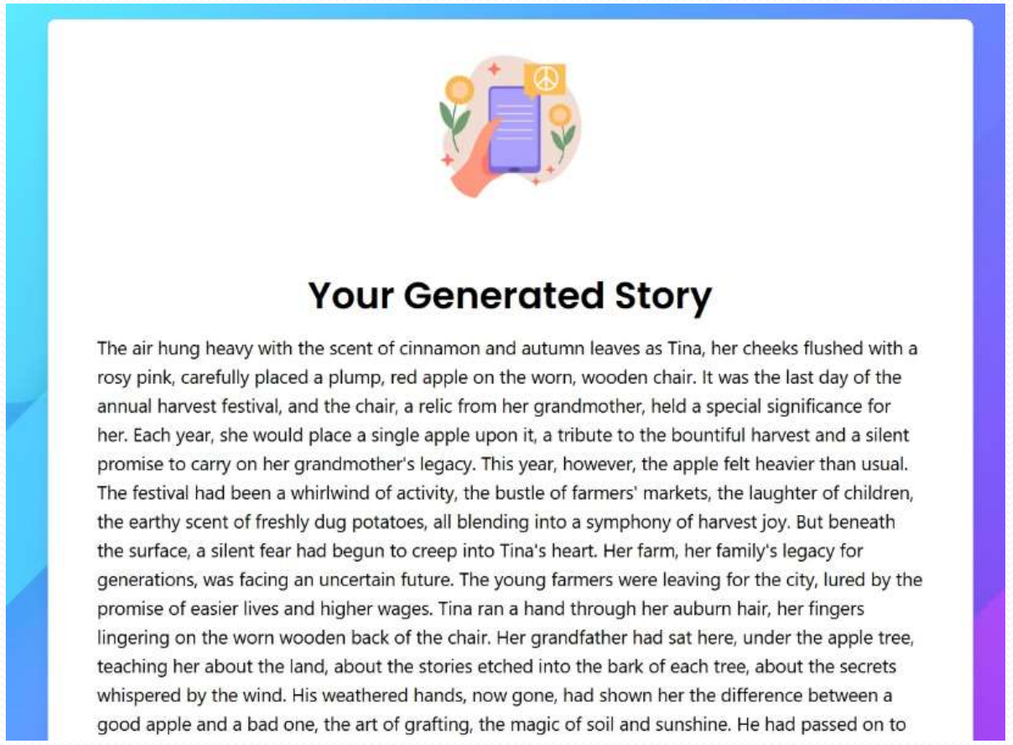


SpeakCharacter Page



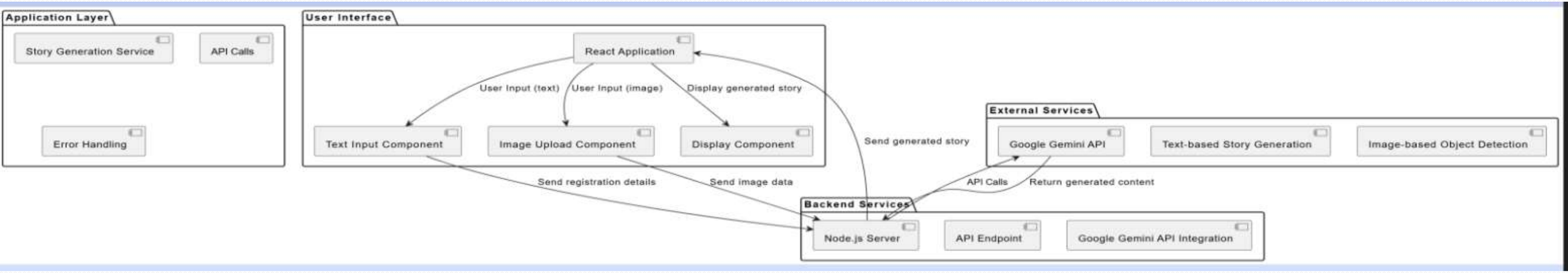
ImageUploading Page

5

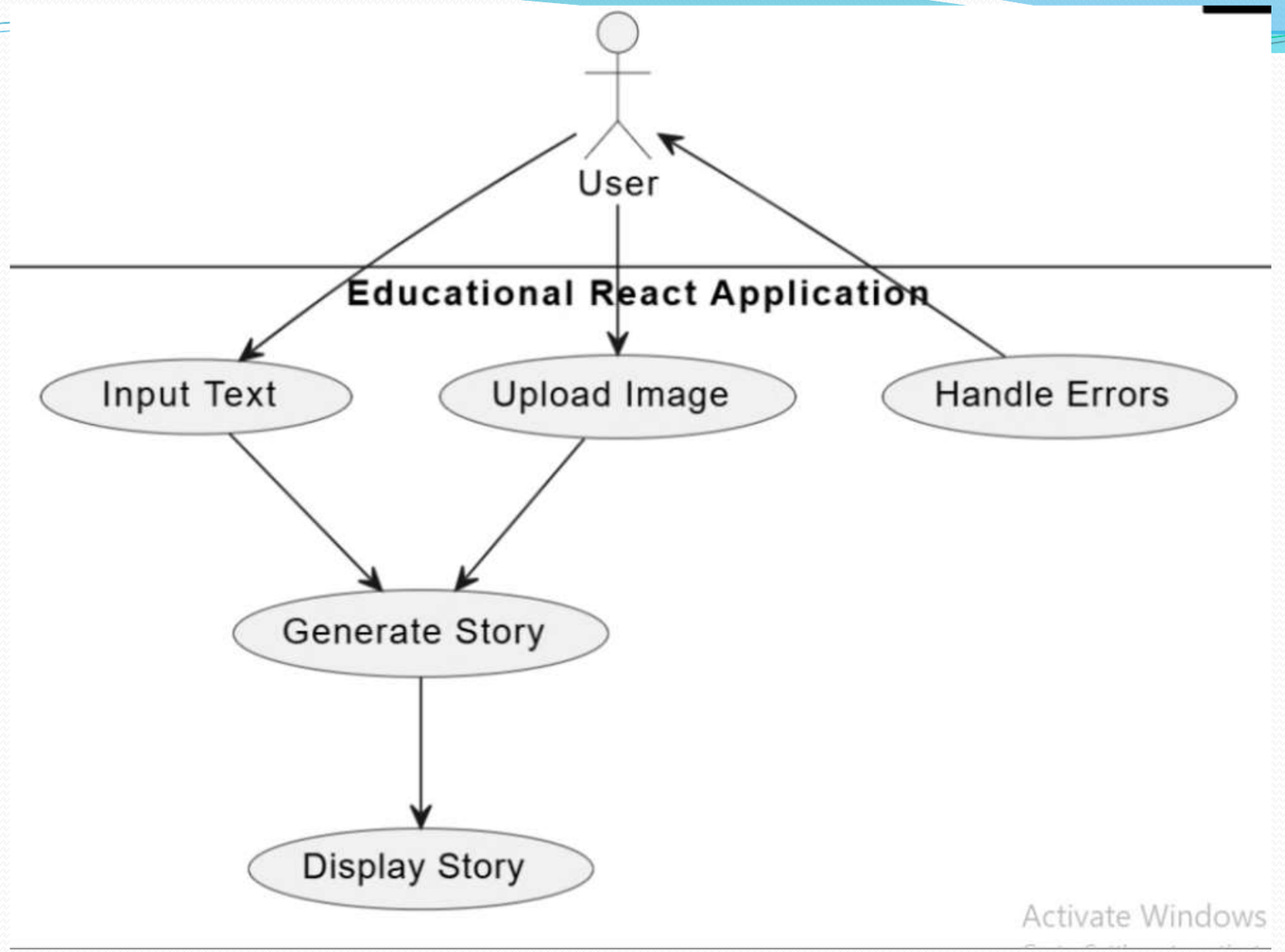


Story Page

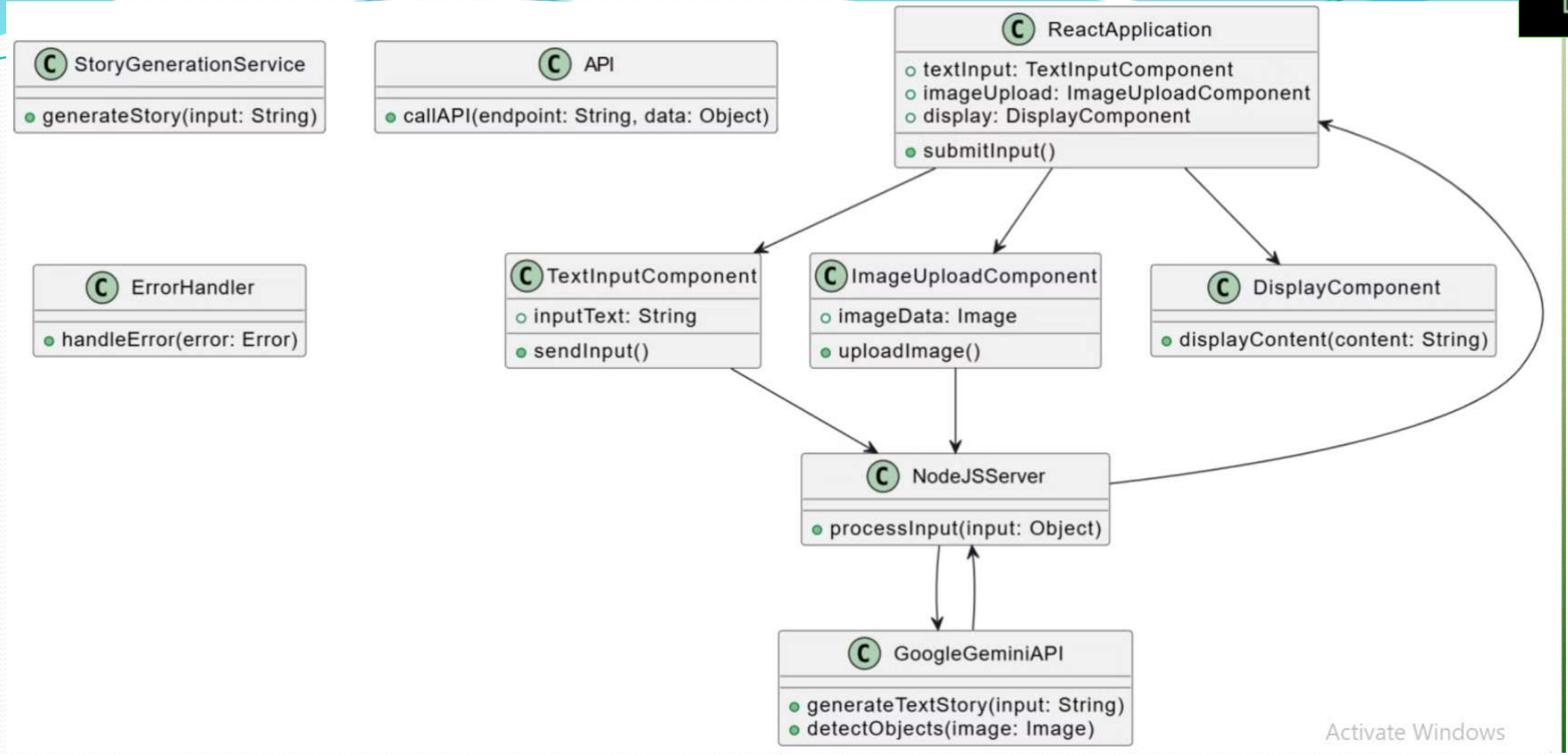
6



Architecture Diagram

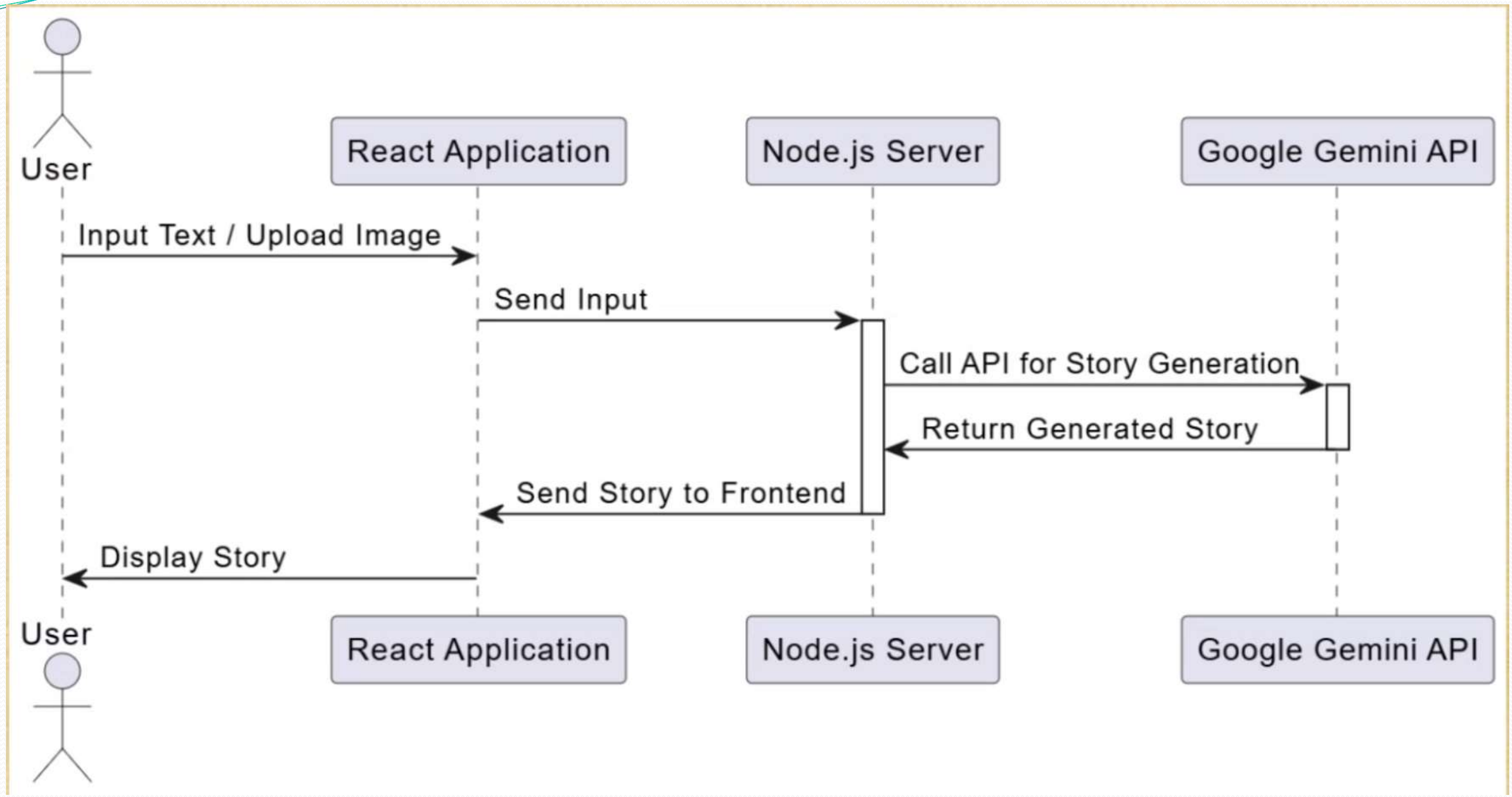


Use Case Diagram

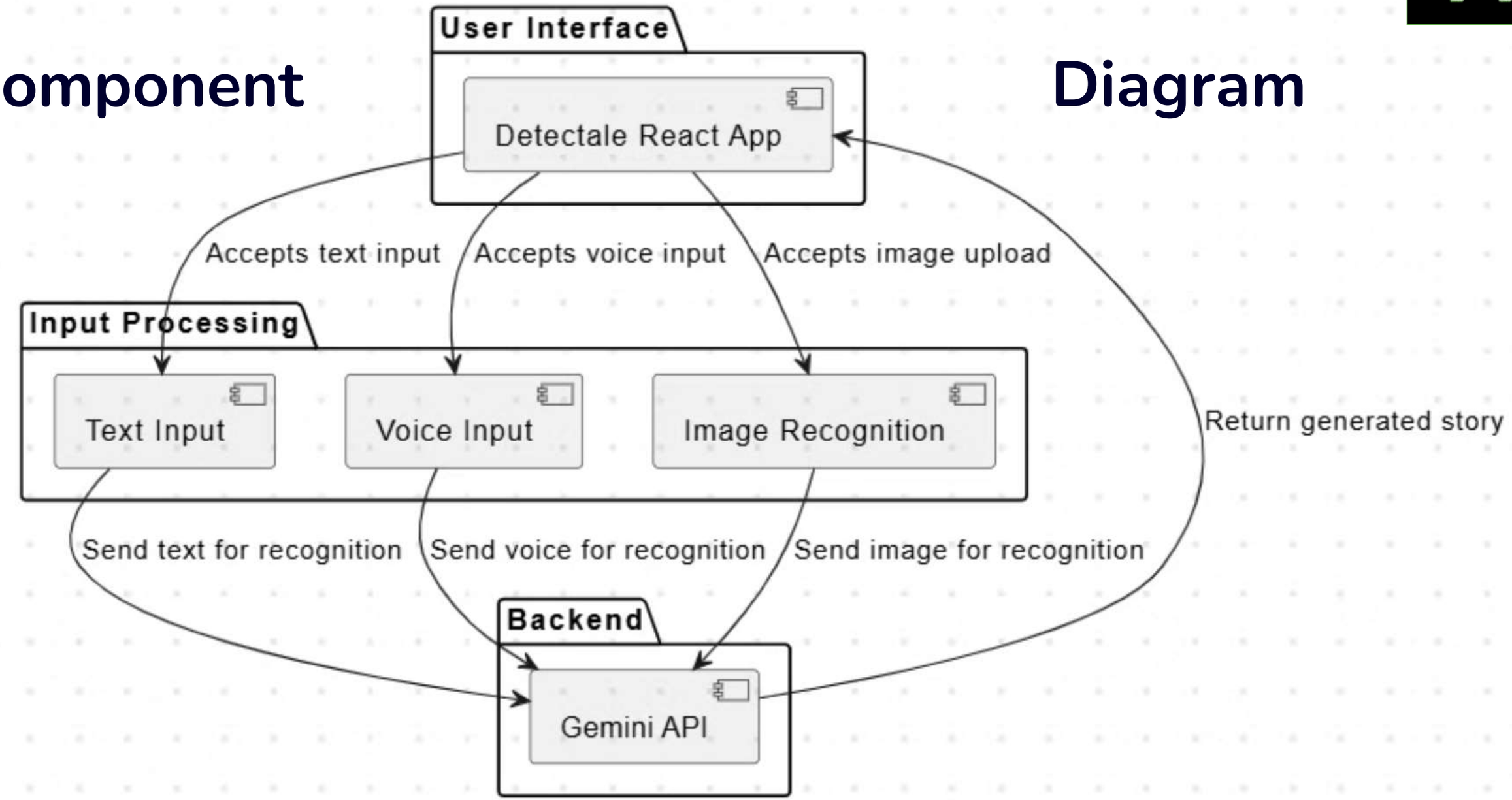


Class Diagram

Sequence Diagram



Component Diagram



The Future of Detectale

Detectale is constantly evolving and expanding its capabilities, there are various future enhancements as follows:

Sharing of Stories

Allow users to share their stories on social media or within the app community.

Offline Mode

Allow users to download stories or use voice recognition offline for improved accessibility.

Audio and Visual Effects

Add soundscapes, background music, or animations to enhance the storytelling experience.

THANK YOU !