SMART CULTURAL STORY TELLER

(**Project report by Hamsalekha Sridhara-iitrprai\_24081055**)

Indian mythology is a vast treasure of stories filled with valuable lessons for children. These tales have the power to teach important life values in a fun and engaging way. Unfortunately, many children today have limited access to them. Through my project, I aim to bring these stories to young audiences by transforming them into short, engaging video stories and sharing them via social media.

## **Problem Statement**

Use AI to preserve and retell cultural stories in engaging, interactive, and accessible ways, especially for younger generations

## **Objectives**

## Transform mythological stories into short narrated video clips.

* Ensure accessibility through free-tier APIs.
* Support multilingual narration (English, Kannada).

## **Methodology / Tech stack**

**Control Flow of the system**

1. **Story Input**
   * User provides a story prompt (e.g., from Indian mythology).
2. **Segmentation**
   * **Gemini-2.5-Flash API** splits the story into 5–7 scenes.
   * Generates narration text and image prompts.
3. **Image Generation**
   * Scene visuals generated using **Imagen-3.0-generate-002 API**.
4. **Scene Processing**
   * Narration audio created via **gTTS** (English, Kannada) or **ElevenLabs**.
   * Scene image animated using **MoviePy** (pan/zoom).
   * Subtitles overlaid using **PIL (Python Imaging Library)**.
5. **Final Assembly**
   * All scenes combined into a complete narrated video using **MoviePy**.

## **Implementation / Development cycle (** [Link for Git Repository](https://github.com/S-Hamsalekha/AI-story-teller)**)**

* **Text-to-Image (T2I) Experiments**
  + Conducted initial trials with Hugging Face models such as **Flux Schnell** and **Stable Diffusion** on **Google Colab** and **Kaggle**.
  + Encountered **memory constraints** and **high runtimes**, limiting large-scale experimentation.
* **Stable Diffusion v1.5**
  + Selected SD V1.5 for its relatively faster inference.
  + Performed adequately on **simple prompts**, but failed to generate **high-fidelity Indian mythological characters**.
* **LLM-Based Scene Structuring**
  + Integrated **Large Language Models (LLMs)** to automatically **segment stories into scenes**.
  + Used **MoviePy** for stitching generated images into a sequence, synchronizing them with narration.
* **API Exploration**
  + Evaluated **Google Gemini 2.0 Flash APIs**, which produced **better-quality images for mythological characters**.
  + Access was discontinued (free tier access was removed post Sept 26).
* **Final Image Generation Pipeline**
  + Migrated to **Google Imagen 3.0 APIs**, which provide: A **free tier** for experimentation.
  + **Consistent, high-quality images** across mythological themes.
* **Narration and Text-to-Speech (TTS)**
  + Added **Google TTS (gTTS)** for narration, supporting:
    - **English (Indian accent)**
    - **Kannada**
  + Integrated **ElevenLabs APIs** to overcome flat, robotic tones of gTTS.
  + Neural voices from ElevenLabs provided **more natural speech quality**.
* **Animation Layer**
  + Implemented cinematic effects using **MoviePy**: Zoom in/out, Pan
* **Subtitles and Captions**
  + Incorporated subtitle/caption overlay with **Python Imaging Library (PIL)** for accessibility and improved user engagement

## **Results & Observations**

* Generated quality videos for 10–12 mythological stories, in an execution time of around 2- 3 minutes, for each story (*\*Ramayana\**, *\*Mahabharata\**).
* Narration with an Indian accent and captions integrates well with visuals.
* Cloud APIs ensure stability, speed, and ease of use.
* Free-tier APIs used, makes it easy for everyone to use.
* Demo and sample outputs shared at: [link](https://drive.google.com/drive/folders/1GdPDWORcDhZj37ju6VLLNIUAqx-zJ4BV?usp=drive_link)

## **Conclusion & Future Work**

* **Summary of Contributions**
  + Developed a system to transform mythological stories into short narrated videos with captions.
  + Integrated image generation, TTS narration (English & Kannada), animation, and subtitles into a unified pipeline.
  + Ensured accessibility by relying on free-tier cloud APIs.
* **Scope for Improvement**
  + Broaden language support and adopt more natural neural voices (e.g., AI4Bharat, Bhashini APIs).
  + Improve character consistency across generated images.
  + Extend testing to lesser-known mythological stories and characters.

## **Acknowledgements**

I sincerely thank Sudarshan Sir, Karthik Sir, and the entire Minor in AI team for their guidance and support in helping me successfully complete this project