Indian Institute of Information Technology, Pune

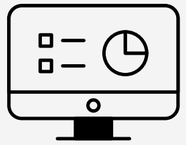
BTP Project
AI Storytelling Web-App

An AI Storytelling App for kids with language assesment

Group Supervisor: Dr. Bhupendra Singh

Group Members:

Saransh Mehra
Tushar Parsai
Aayush Sharma
Ayush Sinha
Sudhir Sude



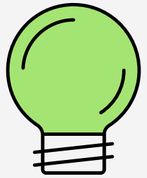
Introduction

Project Summary:

The project aims to develop an AI-driven Storyteller for young children. It will **generate interactive stories** based on user prompts, assess pronunciation and vocabulary through **voice detection**, and **provide a grading system** based on the child’s reading and pronunciation skills. The performance report will be shared with parents for tracking the child's progress.

Key Features:

- Story generation based on text prompts.
- Pronunciation and vocabulary assessment using voice recognition.
- Grading system that evaluates reading and speaking abilities.
- Detailed performance reports for parents.
- Children friendly and Parental controlled system



Motivation

Current Gaps in Educational Tools:

- It's challenging for parents to find engaging and **educational resources for their children**.
- Our app offers AI-generated stories with daily reading goals, helping children **develop a consistent reading habit** while enjoying personalized, captivating narratives.

Lack of Personalized Feedback:

- Traditional reading tools fall short in providing tailored feedback.
- Difficult for parents and educators to track areas for improvement, such as pronunciation or reading fluency.

Difficulty in Assessing Pronunciation and Vocabulary:

- Conventional methods of reading rarely focus on pronunciation and vocabulary development.
- Don't offer real-time analysis or feedback.
- Our app bridges this gap with AI-powered speech recognition.

Problem Statement

Core Problem:
How can we create an AI-driven educational platform that not only engages kids but also accurately assesses their pronunciation, vocabulary, and reading skills, while providing feedback and reports to parents?

Sub-problems:

- How to generate age-appropriate, dynamic stories that engage children.
- How to design an effective pronunciation assessment system for kids.
- How to grade children’s reading skills fairly and provide meaningful reports for parents.

Challenges:

- Developing a robust voice recognition system that works well with children’s unique speech patterns.
- Designing a user interface that is simple and fun for kids to use.
- Creating an AI grading system that is both accurate and flexible for different age groups.

Literature review

1. Children's speech recognition is challenging:
Automatic speech recognition (ASR) systems have difficulty recognizing children's speech, with **error rates up to 275% higher** than those for adults due to unique acoustic characteristics of children's voices.

2. Vocal tract size and bandwidth affect ASR performance:
Our study Study have shown that reducing the bandwidth from 8 kHz to 4 kHz disproportionately
increases **word error rates (WERs)** for children's speech, highlighting the importance of designing specialized ASR systems for children.

3. Lack of child-specific data hinders advancements in child ASR:
Unlike adult speech recognition, which has benefited from large-scale transcribed datasets and end-to-end models, child ASR is hindered by a lack of child-specific data, making it challenging to adapt existing models to recognize child speech.

Literature review

4. Transfer learning and audio augmentation can improve child ASR: Approaches like transfer learning and audio augmentation have emerged as promising methods for adapting adult ASR models to recognize child speech, particularly in the context of non-native child speakers.

5. Targeted fine-tuning on child speech data is key: Recent research has shown that pretraining on large-scale adult datasets and fine-tuning on child speech data can significantly improve child speech recognition, offering promising directions for future research in this area using models like wav2vec2, even with only 10% of the available child speech data.

Proposed Solution

1. Personalized AI-Generated Stories

The app uses OpenAI's LLM model to generate engaging, child-friendly stories with accompanying illustrations.
GPT Script (.gpt file) is utilized to ensure that all stories are appropriate for children, following strict content guidelines.
The app provides daily reading goals for children, promoting consistent learning and habit-building.

2. Speech Recognition for Reading Assessments

Using Hugging Face for speech recognition and evaluation offers a more modern and effective approach due to the advancements in transformer-based models and self-supervised learning. Here's a deeper look into the technical aspects of using Hugging Face for speech-related tasks:

Transformer-Based Models for Speech Recognition

Hugging Face provides access to state-of-the-art transformer-based models like Wav2Vec 2.0, Whisper, and HuBERT. These models leverage transformers to **process raw audio data and recognize speech with remarkable accuracy**, making them particularly suited for tasks involving children's speech, which can be more variable in terms of pronunciation, pitch, and fluency.

Wav2Vec 2.0:

Wav2Vec 2.0 is built on a self-supervised learning architecture. It learns contextualized representations directly from raw audio by masking parts of the speech input and predicting the missing audio features. This is similar to the way models like BERT predict masked words in NLP tasks.

Proposed Solution

3. Progress Tracking and Parental Insights

The app tracks daily reading assessments, allowing parents to monitor their child's improvement.
Real-time feedback on fluency and pronunciation helps parents understand where their child excels and where additional focus is needed.
The dashboard displays progress metrics in an easy-to-understand format, promoting active involvement in the child’s learning journey.

4. Safe, Age-Appropriate Content

Every story generated in the app is curated to be safe, educational, and suitable for young readers. By utilizing the GPT Script, we ensure that the content adheres to strict child-safety standards.

5. Reading Evaluation

Our app evaluates a child's reading skills based on various factors such as Pronunciation accuracy, Fluency,Inflection and stress, Pacing, Expression and Engagement
We also assess comprehension and understanding by asking relevant questions, providing a comprehensive picture of the child's reading abilities.

Technologies used

Frontend (UI)	HTML, CSS, React JS.
Backend	Node.js, Express, Django, OpenAI
Database	MongoDB, Cloudinary
AI/ML	Python, Tensorflow, Keras, Hugging face



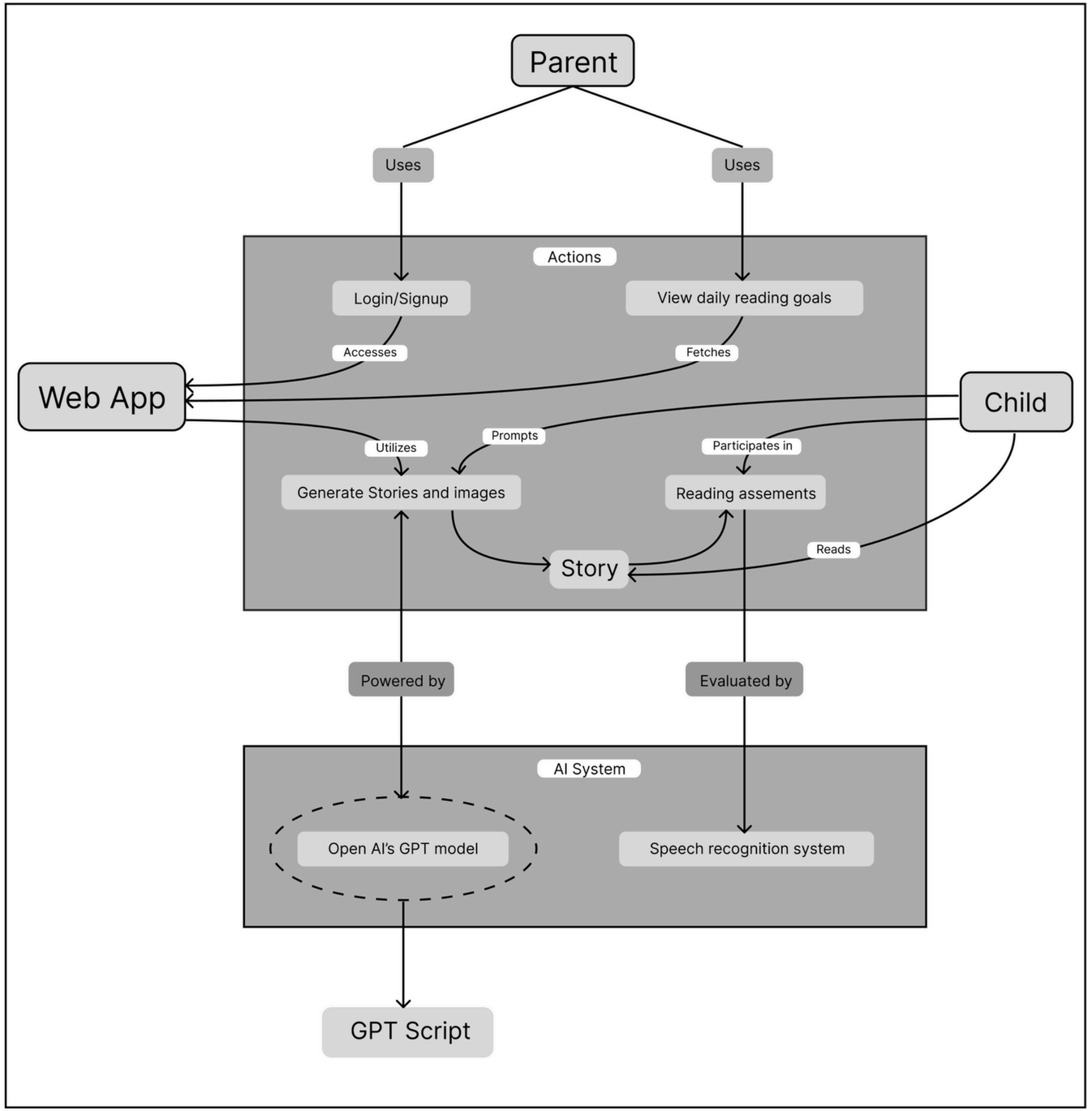

Back to Agenda Page

UseCase

AI-Powered Story Generation: Uses OpenAI's GPT model to create unique stories and illustrations, delivering an engaging and immersive experience for children.

Guideline Compliance with GPT Script: We utilized a GPT Script (.gpt file) to ensure that all stories follow strict content guidelines.

Progress Tracking and Improvement: built-in speech recognition system evaluates child's reading performance.



Future Scope

Multilingual Support:

- The platform can be extended to support multiple languages to cater to diverse linguistic needs.

Gamification:

- Adding rewards, levels, and challenges to make the platform more engaging for kids.

Integrating with Schools:

- Schools could use the platform as a tool to track students' progress in language acquisition and literacy development.

Enhanced Pronunciation Detection:

- Incorporating more advanced speech recognition techniques to improve accuracy with children’s voices.

References

- YAROSLAV GETMAN, NHANPHAN, RAGHEB AL-GHEZI, EKATERINA VOSKOBOINIK, MITTUL SINGH, TAMÁS GRÓSZ, MIKKOKURIMO, GIAMPIERO SALVI SOFIA STRÖMBERGSSON, TORBJØRN SVENDSEN, ANNASMOLANDER, AND SARI YLINEN. **“Developing an AI-Assisted Low-Resource Spoken Language Learning App for Children”,** in IEEE Access, vol. 11, pp. 86025-86037, 2023, doi: 10.1109/ACCESS.2023.3304274.
- R. Jain, A. Barcovschi, M. Y. Yiwere, P. Corcoran and H. Cucu, **"Exploring Native and Non-Native English Child Speech Recognition With Whisper,"** in IEEE Access, vol. 12, pp. 41601-41610, 2024, doi: 10.1109/ACCESS.2024.3378738.
- R. Jain, A. Barcovschi, M. Y. Yiwere, D. Bigioi, P. Corcoran and H. Cucu, **"A WAV2VEC2-Based Experimental Study on Self-Supervised Learning Methods to Improve Child Speech Recognition,"** in IEEE Access, vol. 11, pp. 46938-46948, 2023, doi: 10.1109/ACCESS.2023.3275106.
- Y. Takbiri, A. Amini and A. Bastanfard, **"A Structured Gamification Approach for Improving Children's Performance in Online Learning Platforms,"** 2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Shahrood, Iran, 2019, pp. 1-6, doi: 10.1109/ICSPIS48872.2019.9066006.
- M. Russell, S. D'Arcy and L. Qun, **"The Effects of Bandwidth Reduction on Human and Computer Recognition of Children's Speech,"** in IEEE Signal Processing Letters, vol. 14, no. 12, pp. 1044-1046, Dec. 2007, doi: 10.1109/LSP.2007.906213.