

Multimodal Story Generation and Interaction System Using Generative AI

Vidyashree N V¹, Koushik M², Dr. T Vasudev³

^{*1,2}Students, Department of Computer Applications

^{*3}Faculty, Department of Computer Applications

Maharaja Institute of Technology, Mysore, Mandya, Karnataka

Abstract: *Storytelling is a fundamental form of communication used for education, entertainment, and knowledge sharing, enabling ideas and experiences to be conveyed in an engaging and meaningful way. With the emergence of Generative Artificial Intelligence, it has become possible for machines to automatically create original and context-aware stories using advanced natural language processing techniques. Generative AI models can understand user inputs and generate coherent narratives, making automated storytelling more interactive and scalable. However, most AI-based storytelling systems are limited to text or audio outputs, restricting accessibility for users with hearing impairments. To address this challenge, the proposed StoryBot project focuses on generating stories using Generative AI and presenting them in an inclusive manner through text, voice, and Indian Sign Language. By adopting a multimodal storytelling approach, the system aims to make AI-generated stories accessible to a wider and more diverse audience.*

Keywords: *Storytelling*

1. INTRODUCTION

Storytelling has always been a powerful medium for education, entertainment, and cultural preservation. With the rapid advancement of Generative Artificial Intelligence (GenAI), automated story generation systems have emerged that can dynamically create narratives based on user inputs. However, most existing AI-based storytelling platforms are designed primarily for text or audio delivery, making them inaccessible to individuals with hearing impairments.

According to the World Health Organization (WHO), more than 430 million people globally experience disabling hearing loss, and this number is expected to rise significantly in the coming decades. In India alone, nearly 63 million people are affected by auditory impairment. Despite this, AI-driven storytelling systems rarely incorporate sign language as a medium of story presentation. This lack of inclusivity limits access to creative and educational content for the Deaf and Hard of Hearing (DHH) community, particularly children in rural and underserved areas.

The proposed system, StoryBot, focuses on AI-based story generation and delivers stories through multiple modalities—text, voice, and Indian Sign Language (ISL). By combining generative AI with multimodal output, the system ensures that storytelling becomes universally accessible while remaining interactive, engaging, and educational.

Scope:

The scope of the StoryBot project extends across education, entertainment, and public communication domains where storytelling plays a vital role. The system can be widely used in **inclusive e-learning platforms**, especially for **Deaf and Hard of Hearing (DHH) children**, to deliver educational stories, moral lessons, and interactive learning content in text, voice, and Indian Sign Language (ISL). Schools for the Deaf, special education institutions, and NGOs working with hearing-impaired learners can adopt this system to improve comprehension, engagement, and learning outcomes. In rural and underserved regions, where access to trained ISL educators is limited, StoryBot can act as a virtual storytelling assistant, enabling self-paced and visually rich learning experiences.



Beyond education, the project has strong applicability in the **digital entertainment industry**, including children's storytelling apps, interactive storybooks, and educational games. Content creators, publishers, and ed-tech startups can use the platform to generate inclusive story content that reaches both hearing and non-hearing audiences simultaneously. The system can also be integrated into **libraries, museums, and cultural centers** to present folk tales, historical narratives, and cultural stories in an accessible multimodal format.

Additionally, StoryBot can be deployed in **public information and awareness systems**, such as interactive kiosks in railway stations, bus terminals, airports, and government offices, where story-based communication is used for public education, safety instructions, or social awareness campaigns. Researchers and developers working in **assistive technology, human-computer interaction, and inclusive AI** can further extend the system by incorporating emotional expressions, adaptive storytelling, and support for other regional or international sign languages. Overall, the project serves as a scalable and adaptable solution that promotes inclusive access to AI-generated storytelling for children, educators, content creators, and the Deaf and Hard of Hearing community.

Problem Statement:

Despite significant advancements in AI-based story generation, current systems fail to provide an inclusive storytelling experience for the Deaf and Hard of Hearing (DHH) community. With millions of individuals affected by hearing impairment in India, there is a pressing need for storytelling platforms that go beyond text and audio to include sign language. The absence of Indian Sign Language in automated storytelling limits access to educational and recreational content, particularly for Deaf children who rely on visual communication.

The problem lies in designing a system that can generate dynamic stories using AI and present them simultaneously in text, voice, and ISL, ensuring equal access for users with diverse communication needs. Without such a system, a large section of the population remains excluded from modern digital storytelling experiences.

Objectives of the Project:

- To design and develop a Generative AI-based storytelling system capable of producing dynamic and engaging narratives.
- To implement multimodal story presentation using text, voice, and Indian Sign Language (ISL) for inclusive storytelling.
- To enhance accessibility for Deaf and Hard of Hearing (DHH) users by integrating ISLbased visual communication.
- To provide an interactive storytelling platform that supports educational and entertainment-based applications.
- To improve user engagement and comprehension through synchronized audio-visual story outputs.

Existing System:

The Current automated storytelling systems rely largely on text generation and speech synthesis. These systems use natural language processing models to generate stories and present them through written text or audio narration. Some advanced platforms incorporate conversational agents or robots that narrate stories using speech and facial expressions.

In educational contexts, especially for children, storytelling platforms are used to enhance learning outcomes through interactive narratives. However, these systems primarily target users with normal hearing and literacy skills. For the DHH community, access is often limited to static text-based content, which may not be sufficient due to varying literacy levels. **Limitations of Existing System:**

Existing storytelling systems do not integrate Indian Sign Language, making stories inaccessible to Deaf users.

Most systems depend on speech output, excluding users with hearing impairments.

One-dimensional story presentation (only text or audio) reduces immersion and interaction.

Educational storytelling platforms fail to address the needs of DHH children, especially in rural India.

Indian Sign Language is rarely considered in AI-driven narrative systems.



Proposed System:

The proposed system is a Generative AI-based multimodal storytelling platform that focuses primarily on automatic story generation while ensuring inclusive delivery. The system generates dynamic narratives based on user input and presents the stories in three formats: text, audio narration, and Indian Sign Language video. The storytelling logic is handled using Google Dialogflow, which manages narrative flow, characters, and plot progression. A specialized Generative Semiotic-Sign Alignment (GSSA) protocol is used to map story semantics into accurate ISL gestures. This ensures that the generated stories retain their meaning and emotional depth when translated into sign language.

By offering multimodal outputs, the proposed system makes storytelling accessible to both hearing users and the DHH community. The system is suitable for applications such as inclusive e-learning platforms, digital storytelling tools for Deaf children, and interactive public information systems.

Module Description:

Text and Voice Input Module

This module accepts user input in text or voice form. The input is processed and converted into a structured format suitable for story generation.

Story Generation Engine

Using Google Dialogflow, this module generates dynamic stories by managing narrative logic, dialogue flow, and contextual responses.

Indian Sign Language Translation Module

This module applies the Generative Semiotic-Sign Alignment (GSSA) protocol to convert story semantics into accurate ISL gestures, ensuring linguistic and contextual correctness.

Audio-Visual Output Module

The generated story is delivered simultaneously as:

Text output

Audio narration

ISL video output

This enables users to experience stories through their preferred communication mode.

Hardware and Software Requirements:

Hardware Requirements

Component	Minimum Requirement
Processor	Pentium IV 2.2 GHz or above
Hard Disk	20 GB
RAM	4 GB

Software Requirements

Category	Description
Operating System	Windows 10 / 11
Programming Language	Python 3.10
IDE	Visual Studio Code
Front End	HTML5, CSS3, JavaScript
Back End Framework	Django
Database	MySQL



Server	WAMP / XAMPP
Design Tool	Rational Rose

II. CONCLUSION

The StoryBot system demonstrates how Generative AI can be effectively utilized to transform automated storytelling into an inclusive experience for users with diverse communication needs. By focusing on story generation as the core functionality and extending its reach through text, voice, and Indian Sign Language, the proposed approach overcomes the limitations of conventional audio-centric storytelling platforms. The integration of the GSSA protocol ensures linguistic accuracy in ISL translation, while the multimodal output enhances engagement and comprehension for the Deaf and Hard of Hearing community. This project highlights the potential of inclusive AI systems in reducing accessibility gaps in education and digital entertainment, particularly for Deaf children in rural India. Future enhancements may include support for additional regional sign languages, emotional expression in sign animations, and adaptive storytelling based on user preferences, further strengthening the role of AI in inclusive digital storytelling.

