



STORYTELLING UTILIZING GENERATIVE AI TO FOSTER INCLUSION OF INDIVIDUALS WITH DISABILITIES

CUENTACUENTOS BASADO EN IA GENERATIVA PARA PROMOVER LA INCLUSIÓN DE PERSONAS CON DISCAPACIDADES

Keren Mitsue Ramírez Vergara^{1,*} , Asdrúbal López-Chau¹ ,
 Rafael Rojas Hernández¹

Received: 27-03-2024, Received after review: 05-06-2024, Accepted: 18-06-2024, Published: 01-07-2024

Abstract

This article presents the comprehensive design and evaluation of a digital storytelling system tailored for Latin American children aged 4 to 6, leveraging generative artificial intelligence. Tests were conducted to assess the system's functionality, content diversity, generation times, and voice quality, including intonation, speed, and pronunciation. The results substantiate the system's operational efficacy and user-friendly interface. The stories generated demonstrate substantial diversity, as indicated by Jaccard indices calculations, which reveal a maximum value of 0.2 derived from evaluating 30 distinct stories. As expected, there was a proportional increase in story generation times relative to their length. 'Onyx' from OpenAI's text-to-speech (TTS) was identified as the most appropriate voice for storytelling. Nonetheless, pronunciation inaccuracies were observed across all tested TTS model voices. The analysis demonstrated that the system generates a variety of stories that foster value formation in Spanish-speaking children, thereby promoting the importance of including individuals with disabilities. Notably, all content within the stories was found to be suitable for children, with no inappropriate material detected in any of the narratives.

Keywords: ChatGPT, Storytelling, Disability, AI Generative, Inclusion

Resumen

En este artículo se presenta el diseño completo y la evaluación de un sistema cuentacuentos digital destinado a niños de entre 4 y 6 años en Latinoamérica. Este sistema está basado en inteligencia artificial generativa. Se realizaron pruebas que abarcaron el funcionamiento del sistema, la diversidad de contenidos, los tiempos de generación, la evaluación de voz, entonación, velocidad y calidad de pronunciación. Los resultados confirman que el sistema funciona correctamente y es intuitivo. Las historias generadas muestran un alto grado de diversidad, ya que al calcular los índices de Jaccard, el valor máximo encontrado fue de 0,2 en las evaluaciones de treinta cuentos analizados. Como era de esperarse, los tiempos de generación aumentan conforme se incrementa la longitud de los cuentos. Se identificó que la voz que mejor se adapta para contar los cuentos es Onyx de la TTS de OpenAI. Sin embargo, se observaron errores de pronunciación en todas las voces del modelo TTS. De acuerdo con el análisis realizado, el sistema crea historias diferentes, que promueven valores en los niños de habla hispana, fomentando la importancia de la inclusión de personas con discapacidad. Cabe destacar que en ningún cuento se encontró contenido no apto para niños.

Palabras clave: ChatGPT, cuenta cuentos, discapacidad, IA generativa, inclusión

^{1,*}Centro Universitario UAEM ZUMPANGO, Universidad Autónoma del Estado de México. Zumpango, Estado de México. México. Autor para correspondencia ✉: kramirezv003@alumno.uaemex.mx.

Suggested citation: Ramírez Vergara, K. M.; López-Chau, A. and Rojas Hernández, R. "Storytelling Utilizing Generative AI to Foster Inclusion of Individuals with Disabilities," *Ingenius, Revista de Ciencia y Tecnología*, N.º 32, pp. 101-113, 2024, DOI: <https://doi.org/10.17163/ings.n32.2024.10>.

1. Introduction

Generative Artificial Intelligence (GAI) constitutes a notable breakthrough in artificial intelligence (AI), offering the capability to generate diverse content types, including texts, images, source code for various programming languages, scenario designs, legal arguments, and high-definition videos.

Large Language Models (LLMs) facilitate the automated generation of text, crafting original documents by leveraging extensive textual data harvested from the Internet. This generation is enabled through contemporary reinforcement learning architectures incorporating human feedback and deep learning technologies [1], specifically transformers.

ChatGPT has become a potent tool across multiple domains, including education, marketing, finance, and customer service. In the educational sector, the deployment of generative text-based AI has witnessed a significant global increase in the digital era [2]. This technology assists in elucidating concepts through simplified explanations, facilitates problem-solving by demonstrating diverse methodologies, and aids in developing reading and comprehension skills among early-year students, among other applications.

This innovative technology propels the shift toward an educational paradigm that is more immersive, dynamic, participatory, and inclusive, underscoring the pivotal roles of teachers and students as agents of change [3]. Integrating systems like ChatGPT into educational frameworks is poised to augment human capabilities, contribute to reducing inequalities, and foster the promotion of core values.

One effective method to cultivate these values is through storytelling to children. This approach not only aids in comprehending the world but also stimulates the imagination and facilitates conflict resolution. Storytelling can convey significant messages, expand horizons, and encourage active engagement with the environment. Furthermore, it enhances communication, debate, and interpretation skills essential for holistic development [4].

As an educational tool, storytelling significantly enhances teaching and learning by making knowledge acquisition engaging and enjoyable. Storytelling facilitates reflection and moral consideration and is crucial in stimulating children's cognitive and intellectual development from an early age [4].

Therefore, enriching the storytelling experience with human values early in a child's life is vital. Accordingly, the system proposed in this article is designed for children aged 4 to 6 years, a critical period during which children start to become acquainted with written text.

The system's emphasis on inclusive education aligns with the increasing importance attributed in recent decades to eliminating discrimination against

vulnerable groups. This initiative aims to mitigate such issues, particularly in Mexico, where inclusive educational practices are limited.

According to 2019 OECD reports, Mexico ranks among the countries with the lowest levels of educational inclusion, with only 2.85% of students with disabilities receiving education, despite approximately 15% of the student population having some form of disability [5]. Furthermore, the World Bank noted in 2021 that although there are approximately 85 million people with disabilities in Latin America and the Caribbean [6], progress in enhancing this demographic's employment, education, and healthcare programs has been minimal.

This article outlines the design and implementation of a system that leverages generative artificial intelligence, specifically ChatGPT, among other technologies, for the automatic generation of stories targeted at Spanish-speaking children. The system is driven by two primary objectives: a) to instill values such as respect, tolerance, and empathy towards individuals with disabilities through storytelling, and b) to enhance reading and comprehension skills among Spanish-speaking children.

The principal contributions of this article are summarized as follows:

1. It proposes a novel approach to address the challenge of enhancing the inclusion of people with disabilities in Latin America, utilizing a cutting-edge technological solution.
2. It details the comprehensive design of a generative artificial intelligence (GAI)-based software system specifically tailored to create and narrate stories for Spanish-speaking children.
3. It illustrates the development of a specifically crafted prompt that generates diverse stories, which cultivate values of empathy and respect towards individuals with disabilities.
4. It assesses the system's performance in terms of the stories' diversity and the intonation, speed, and pronunciation quality of the storytelling narratives.
5. It makes the complete source code of the system and supplementary files available for non-commercial use through a GitHub repository. [7].

1.1. Literature Review

A comprehensive literature review was conducted using the IEEE Explore, Science Direct, and Scopus electronic databases. Searches were carried out in both English and Spanish. The inclusion criteria were restricted to journal articles, books, and conference papers published between 2021 and 2024. This specific

date range was chosen because generative text artificial intelligence became globally accessible starting in 2021. 91 documents were collected through the review process, comprising 85 articles, 5 books, and 1 manual.

The searches were conducted using the following keywords and logical operators:

- "*inclusiveness AND AI AND education*",
- "*disability AND AI AND education*",
- "*storytelling AND inclusiveness AND education*",
- "*apps AND inclusiveness AND AI*"

The analysis of sources retrieved from well-established databases indicated that Scopus contributed 18 documents, ACM Digital Library 4, ScienceDirect 47, and IEEE Explore 4 documents. Articles not directly relevant to the current study were excluded, resulting in a focused selection of 23 pertinent articles.

The state-of-the-art review centered on four core categories delineated as the primary research objectives:

- The role of digital storytelling in child development.
- The utilization of artificial intelligence in enhancing creativity.
- The impact of digital storytelling on child development.
- The potential of digital storytelling to promote inclusive education.

1.1.1. Digital Storytelling in Child Development

The significant influence of digital storytelling on child development has been extensively documented by various scholars. Bratitsis *et al.* define a story as a sequence of sentences that narrate events or experiences, typically involving central characters [8].

Through storytelling, themes such as compassion, solidarity, and empathy are prominently featured and explored. Additionally, Juppi observes that a digital story typically integrates elements such as text, music, sound effects, or the author's own recorded voice and recommends that the duration of a digital story should ideally range between 2 and 4 minutes [9]. Juppi further elucidates that digital narratives are often designed to empower individuals by fostering personal growth, enhancing control over their lives, and enabling them to act as informed citizens. This empowerment is facilitated through developing technical and creative skills in expression and communication, which are hallmark

features of digital narratives. Juppi also advocates for educational institutions across various academic levels to leverage digital storytelling to promote inclusive education, empathy, respect, civic engagement, and democratic participation.

The impact of digital storytelling on children has been extensively explored in prior research. Bratitsis and Ziannas investigated the development of social empathy in children over 6 years old through interactive digital storytelling, utilizing the tale "The Sad Little Chick" created in the Scratch programming environment [8]. The study engaged 25 sixth-grade early childhood education students who read the story and engaged in interactive activities designed to elicit their emotional responses. The outcomes were encouraging, demonstrating heightened interest and sensitivity towards the main character and an enhanced understanding of empathy and its practical application in daily situations. This research underscores how digital stories can effectively foster children's comprehension of inclusive values and empathy.

Conversely, Tseng *et al.* developed PlushPal, which utilizes machine-learning techniques to transform plush toys into interactive digital objects. PlushPal enables children to digitalize their stuffed animals, allowing them to recognize gestures and produce personalized sounds [10]. Furthermore, it integrates storytelling techniques to animate the toys, imbuing them with capabilities that foster connections with positive memories and previous experiences.

1.1.2. The utilization of artificial intelligence in enhancing creativity

Creativity is commonly defined as the capacity to generate novel, unique, and valuable ideas or artifacts [11]. This capacity is augmented by GAI systems. For instance, Haase and Hanel suggest that chatbots, equipped with expansive databases, can recombine ideas to produce outputs that exhibit levels comparable to everyday human creativity [11].

Li views AI as an invaluable resource for human writers, facilitating the enhancement and expansion of increasingly complex ideas, thereby fostering a divergent thought process [12]. He anticipates a future where AI and human collaboration will be dynamic, interactive, and participatory. Conversely, Habib *et al.* assess the creativity of GAI systems within educational settings by examining the flexibility, elaboration, and originality of responses via user acceptance tests [1].

Their findings emphasize [13] improvements in divergent thinking and the introduction of varied perspectives. The research underscores the importance of carefully integrating GAI into creative education to promote a symbiotic relationship between human creativity and AI. Additionally, Li supports the notion that the ethical application of ChatGPT could enhance

inclusivity and diversity in educational contexts [12].

1.1.3. Artificial Intelligence in Child Development

In a notable study, Kalantari *et al.* [14] explored the effects of AI on early childhood education. This investigation utilized a qualitative exploratory approach involving children aged 6 to 7 years and their parents to assess a software application named "Kids Story Builder." The findings from this study indicated that the technology not only enhances children's understanding of and connection with themselves and their families but also promotes narrative thinking during the story creation process.

Jiahong and Yang [15] conducted an exploratory review that assesses, synthesizes, and highlights recent literature on the application of AI in early childhood education. However, their study touches only superficially on the specific uses of AI within these contexts.

1.1.4. Artificial Intelligence to Promote Inclusive Education

Artificial intelligence (AI) technologies and emerging technological tools significantly influence society and are progressively being incorporated into educational settings [16].

Consequently, numerous scholars advocate for appropriately integrating these tools within educational frameworks. While the literature reviewed does not reveal specific implementations of Generative AI (GAI) in inclusive education, it does offer guidance, advice, and recommendations designed to ensure a beneficial impact. These technologies can transform education by altering students' experiences within and beyond the classroom [16].

Yu [17] underscores that the core educational value of ChatGPT resides in its ability to facilitate access to knowledge, create content, and promote educational inclusion. He also emphasizes that ethical management, transparency, and accountability represent significant challenges AI introduces in educational contexts. ChatGPT should augment human capabilities and contribute ethically, steering towards a more immersive, dynamic, participatory, and inclusive educational experience.

Li and Lan [18] concur with Salas-Pilco *et al.* [16] and offer a framework for adequately adopting technology, underscoring the importance of promoting social inclusion.

2. Materials and Methods

The GAI-based system designed to generate stories that promote the inclusion of individuals with disabilities was developed using the Kanban methodology.

This approach enhanced continuous task delivery by allowing for the visualization of progress across various sections and tracking pending tasks. The choice of this methodology was informed by its emphasis on continuous delivery, whereby team members work on tasks as they arise without rigidly assigned roles. Any team member could assume new tasks from the list as required. Moreover, a suite of modern technologies was employed to ensure the system's robustness and operational efficiency.

TypeScript was selected as the programming language for its capability to develop robust web applications. It compiles code into JavaScript, enabling it to run across any browser, platform, or operating system [19]. Beyond its technical capabilities, TypeScript is an open-source language that enhances JavaScript syntax, ensuring compatibility with various browsers, servers, and operating systems. The decision to use TypeScript was also supported by its seamless integration with Angular and other libraries, which aids in the development and scalability of the application [20].

Complementing TypeScript, the Angular framework generated the system's user interfaces. Angular is renowned for its efficient Document Object Model (DOM) management and its capability to create scalable web applications in conjunction with TypeScript. Furthermore, Angular is acclaimed for its user-friendly learning curve and capacity to boost development productivity, making it an optimal choice for this project.

ChatGPT was integrated into the application to facilitate the story generation functionality. Its creative collaborator role derives from its capability to ideate dramatic content, develop characters, and craft storylines. ChatGPT has been effectively utilized in interactive storytelling and gaming, enabling users to create dynamic and personalized narrative experiences where the plot adjusts to users' preferences and experiences. Additionally, it can enhance stories or poems with a diverse array of realistic words, emotions, and characters [21]. These attributes render ChatGPT particularly valuable and attractive for generating inclusive content for individuals with disabilities.

The implementation of the application was structured into three distinct phases: developing user interfaces, generating prompts for the model, and integrating the OpenAI API. The overall architecture of the developed system is graphically summarized in Figure 1.

2.1. Prompt Structure Design

During the initial phase of system development, the significance of establishing a general prompt structure to guide the system in generating coherent, relevant, precise, and appropriate results for Spanish-speaking children was recognized. To achieve this, three fundamental elements of this structure were identified, which users can customize to facilitate effective story generation. These elements are as follows:

- three types: micro-story, flash story, and short story, consisting of approximately 300, 750, and 2000 words, respectively.

- In block ④ of Figure 1, the prompt is generated based on parameters the user selects, which directs the model to produce precise and child-appropriate results. The established structure specifies ChatGPT’s role in generating the stories and detailed instructions on the format for the returned response. This structure is illustrated in Figure 2.

In this figure, placeholders within braces, such as *{tipo_de_cuento}*, *{número_de_palabras}*, etc., are substituted with specific values selected by the user (child), corresponding to blocks ①, ② and ③ in the architecture depicted in Figure 1. Furthermore, the characteristics of the stories were defined to achieve the anticipated outcomes. The observations were as follows:

1. *Storyteller Definition.* The model is a storyteller promoting values in individuals with disabilities.
2. *Response Format.* The response must be in JSON format, and its structure must be explicitly defined..

Eres un cuentacuentos creativo que inculca valores en sus cuentos para niños de latinoamérica, los cuentos son divertidos e interesantes, y a veces misteriosos.

La respuesta debe ser en formato JSON con la siguiente estructura:

```
{
  'titulo': string,
  'personajes': [nombre del personaje],
  'contenido': string
}
```

Crea un cuento de {tipo_de_cuento} con {número_de_palabras} palabras sobre el tema {tema_a_tratar}, en el que participen los siguientes personajes:

- {nombre_del_personaje_uno}, un {ocupación_del_personaje_uno} de nacionalidad {nacionalidad_del_personaje_uno} con las siguientes características físicas: {características_físicas_del_personaje_uno}.

La personalidad de {nombre_del_personaje_uno} es {personalidad_del_personaje_uno}.

- {nombre_del_personaje_dos}, un {ocupación_del_personaje_dos} de nacionalidad {nacionalidad_del_personaje_dos} con las siguientes características físicas: {características_físicas_del_personaje_dos}.

La personalidad de {nombre_del_personaje_dos} es {personalidad_del_personaje_dos}.

- {nombre_del_personaje_tres}, un {ocupación_del_personaje_tres} de nacionalidad {nacionalidad_del_personaje_tres} con las siguientes características físicas: {características_físicas_del_personaje_tres}.

La personalidad de {nombre_del_personaje_tres} es {personalidad_del_personaje_tres}.

Consideraciones para el cuento:

- El cuento debe ser original, por favor, esfuérzate en que no se repitan los cuentos.

- Si hay personajes con alguna discapacidad, asegúrate de que sean los protagonistas. Enfoca la atención en ellos para que la trama del cuento resalte el valor de {tema_a_tratar} hacia los personajes con discapacidad.

- Evita colocar en el cuento la descripción física y de personalidad de los personajes, solo menciona implícitamente cual es la discapacidad del personaje, si así la tuvieran.

- El contexto en el que se desarrollan los personajes debe ser ambientado en la nacionalidad del protagonista o los protagonistas, ya sean lugares turísticos, pueblos, costumbres, cultura, ideologías, economía, etc. Debes ser específico en el lugar en el que se encuentran, puedes mencionar nombre de los lugares.

- Puedes elegir el estilo narrativo entre monólogo, narración en segunda persona, narración epistolar, narración como testigo, narración omnisciente, etc. además de cambiar la atmósfera, el ambiente y el simbolismo; la idea es crear un cuento original, creativo y único o inédito que inculque valores a los niños.

- Es importante que los cuentos no sean repetidos en cada una de las iteraciones que se realicen, debe de cambiar el contexto, ambiente, el nombre del cuento, su trama y la forma de narración.

- Es importante que el cuento deje un mensaje o impacto sobre los valores. Este mensaje no debe ser explícito en el cuento, debe ser tratado con una intención didáctica (el niño debe ser capaz de comprender el mensaje o valores a lo largo del desarrollo del cuento).

Figure 2. Prompt for Story Generation

- 3 *Story Creation.* The system is programmed to craft a story utilizing the parameters enclosed in braces, which are to be replaced by values selected by the user.
- 4 *Story Considerations.* This section specifies that the story should be original and center on characters with disabilities who will serve as the protagonists. The narrative aims to portray these individuals' real-life challenges and conflicts and illustrate how they can overcome daily barriers and challenges by embodying values. To ensure the setting is realistic and authentic, it is required that the story's backdrop incorporates elements such as places, ideas, customs, or stories that are characteristic of the characters' nationality, with a specific focus on Latin American and Caribbean contexts.

2.2. Connection with ChatGPT

The third phase of the development process entails integrating the OpenAI API into the system. This API offers a comprehensive suite of services, including natural language processing, speech synthesis, and text generation. However, certain functionalities provided by the API do not apply to the objectives of the proposed system.

Access to OpenAI's services necessitates authentication using an API key supplied by the platform, as depicted in block (5) of Figure 1. It is crucial to acknowledge that usage of this API incurs fees, which vary based on the volume and nature of the requests submitted.

The system employed the GPT-3.5-turbo model for story generation, as indicated in block (6) of the architecture illustrated in Figure 1. The configuration of parameters for this model was as follows:

1. *Message.* This parameter takes an array of message objects, which may function in the roles of system, user, or assistant, each with its specific content [24]. For the system developed, the message structure is detailed in Figure 2.
2. *Model.* The GPT-3.5-turbo-instruct model was utilized.
3. *Number of Tokens.* This parameter establishes the maximum number of tokens the generator can produce in a single request. For this system, the limit was set at 2048 tokens.
4. *Temperature.* This parameter influences the generated text's variability and originality level. A higher temperature setting results in more diverse and creative outputs and increases the likelihood of generating incoherent or irrelevant responses. Consequently, a temperature setting of 0.5 was chosen for optimal performance in this application.

The model outputs the generated content in JSON format. The application then processes this response, which extracts the content and displays it within the user interface (refer to block (7) of the architectural diagram). Consequently, the child views the generated story textually, accompanied by images of the featured characters.

In addition to displaying the story in text form, the developed system is equipped to narrate the generated stories using synthetic voices. This feature was implemented to accommodate the application's primary users: children learning to read. The OpenAI Text To Speech (TTS) API was employed for this purpose, as illustrated in block (6A), Figure 1). This API offers a selection of six integrated voices: alloy, echo, fable, onyx, nova, and shimmer, which support narrations in various languages, including Spanish [25].

2.3. Measuring Story Similarity

One of the principal attributes of the storytelling system is its capability to generate distinct stories with each execution. To quantify this diversity, the Jaccard index was employed.

The Jaccard index is a statistical tool used to assess the similarity and diversity between two sets, calculated according to Equation (1).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

Where A and B represent the sets being compared, and the values of the Jaccard index range from 0 to 1. A value of 0 indicates no similarity between the sets, while a value of 1 signifies that the compared sets are identical.

In the analysis of similarity between stories, pre-processing was conducted, comprising the following stages:

- *Text transformation.* All characters were converted to lowercase to ensure uniformity across the dataset.
- *Text cleaning.* All punctuation marks, character names, nationalities, numbers, and any non-alphabetic characters were removed. Additionally, extra spaces were eliminated to ensure text consistency.
- *Lemmatization.* This process involves converting words to their base form, or lemma. Different words with similar meanings can be treated as the same entity, enabling the system to recognize different verb conjugations as the same base

word. The spaCy library was utilized for lemmatization, employing the pre-trained model designated as `es_core_news_md`, tailored explicitly for processing the Spanish language. The lemmatization process was systematically applied to all the words within the document.

3. Results and Discussion

A functionality test was conducted on the system to ensure accurate story generation. Following this, the diversity of the generated content was evaluated, and an analysis of the voices used in story production was undertaken. The test of the storytelling system was executed in a web browser on a computer running the Windows 10 operating system.

3.1. Functionality Tests

The system functionality test entailed inputting parameters selected by a user. For demonstration purposes, the following steps were executed:

1. A micro-story was selected via the interface, as depicted in Figure 3.



Figure 3. Story Type Selection Interface

2. Three characters –Andrés, Fernanda, and Omar– were chosen for the story, as illustrated in Figure 4.

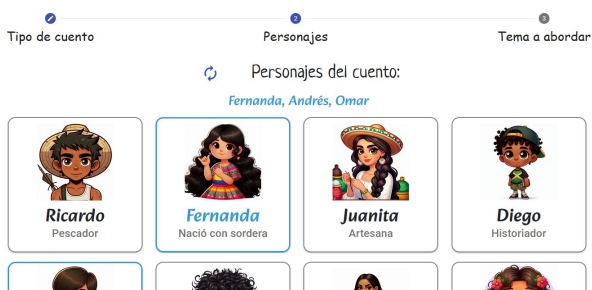


Figure 4. Character Selection Interface

3. The theme of tolerance was selected, as indicated in Figure 5.

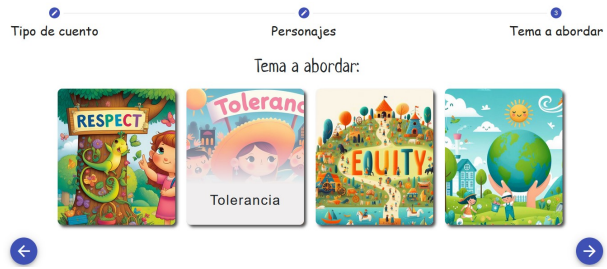


Figure 5. Theme Selection Interface

Once the parameters were set, the story depicted in Figure 6 was generated. This story generation process took approximately 4816 milliseconds.

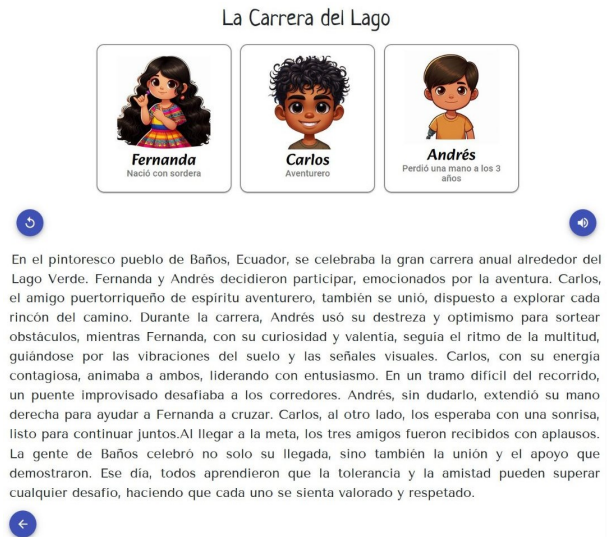


Figure 6. Example of Generated Story

Within the system interface, the user can initiate the playback of the story by selecting the audio button in the upper right corner. The audio, generated by the TTS-1-HD model and narrated using the Alloy voice, lasts 1:01 minutes. The creation process for this audio took 10.79 seconds.

3.2. Diversity of Generated Content

Two experiments were conducted to assess the diversity of content in the generated stories, precisely the extent to which the stories produced by the system differ.

In the first experiment, the characters, themes, and story types were held constant, and the system was tasked with generating stories using these fixed parameters. While the generated stories exhibited similarities,

none of the thirty executions produced identical (repeated) stories. The contents generated in this test are available for public access in the GitHub repository.

Figure 7 summarizes the words frequently used in the titles of the 30 stories generated by the system. Notably, some words not explicitly included in the designed prompts still adhered to the instructions to craft engaging stories.



Figure 7. Word Cloud of the Titles

Conversely, the stories generated by the system prominently feature words that promote tolerance and collaboration, as evidenced by the word cloud in Figure 8, which is derived from the content of the stories.



Figure 8. Word Cloud of the Story Contents

All Jaccard indices between pairs of stories were calculated, and these values are graphically depicted in Figure 9. The analysis reveals minimal similarity among the stories, with the highest Jaccard index value recorded at 0.2137 between micro-stories 23 and 29. These stories are available in the previously mentioned repository.

In the second experiment, stories were generated by randomly selecting characters, themes, and types of stories. After producing more than thirty stories, substantial variety was observed, all by the instructions specified in the designed prompt. Figure 10 graphically displays the Jaccard indices between the stories. The average Jaccard index recorded was 0.0359, with a standard deviation of 0.0427. The highest Jaccard index observed was 0.2.

Upon reviewing the generated stories, it was observed that each narrative incorporates a message that reinforces values and respect towards characters with disabilities.

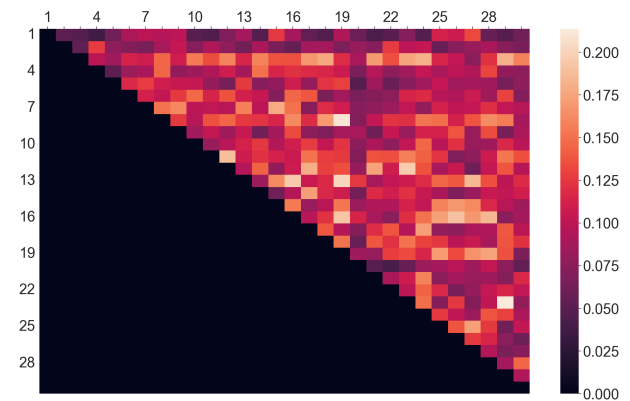


Figure 9. Jaccard Indices Between Stories with Fixed Characters, Type, and Theme

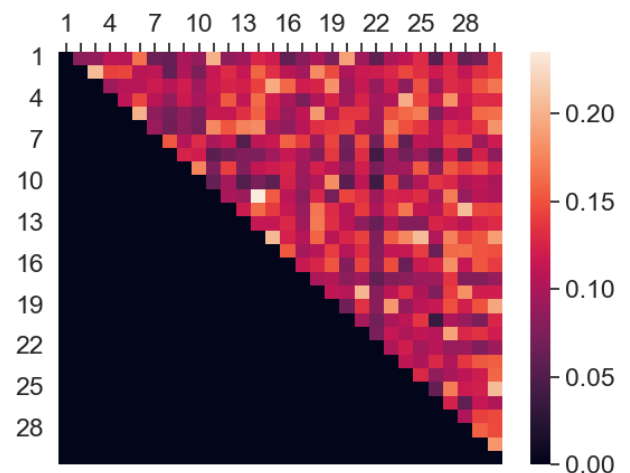


Figure 10. Jaccard Indices Between Stories with Randomly Selected Types, Characters, and Themes

3.3. Analysis of Voices in Story Playback

To evaluate the different voices offered by OpenAI, it was crucial to analyze parameters such as story generation time, voice tone, intonation variation, speech speed, and pronunciation clarity. The primary objective was to identify the most suitable voice to ensure a clear, natural, and emotionally engaging listening experience for children.

3.3.1. Generation Times

The system generated thirty stories, with the type, characters, and themes randomly assigned for each

narrative. The time ChatGPT took for each request to produce the corresponding audio file was meticulously recorded. Eight audio files were generated for micro-stories, eleven for flash fiction, and eleven for short stories. Table 1 summarizes the creation times and durations for each story type.

Table 1. Comparison of Creation Times and Duration of Stories

	Micro-story	Flash fiction	Short story
Creation (average)	9.31 s	11.82 s	12.49 s
Creation (std)	1.03 s	2.41 s	1.82 s
Duration (average)	58 s	66 s	74 s
Duration (std)	13 s	10 s	14 s

As expected, the average time required to create audio files for narration increases with the length of the story. However, the difference in average audio generation times between flash fiction and short stories is less compared to micro-stories. Additionally, it is noted that the generation of audio files for flash fiction exhibits a higher standard deviation compared to the other story types.

It is confirmed that the narration audio duration corresponds with the story type being told. Notably, the duration of the short stories exhibits a larger standard deviation than the other story types.

3.3.2. Voice Tone Evaluation

The tone of voice in a recording plays a crucial role in assessing the end user’s listening experience, as it significantly influences the narratives’ comprehension, empathy, and persuasive power. Upon analyzing the tones in the generated audios, it was observed that three of the six voices, Alloy, Nova, and Shimmer, exhibited a friendly and pleasant tone while narrating the stories, as detailed in Table 2.

Table 2. Comparison of Voice Tones in Story Playback

Voice	Description
Alloy	It projects a friendly tone.
Fable	It is characterized by a serious tone without evoking any specific emotion.
Echo	It exhibits a serious tone with a narrative approach.
Onyx	It exhibits a deep and measured tone, making it ideal for storytelling.
Nova	While possessing a friendly tone, it is perceived as somewhat unnatural.
Shimmer	It features a friendly and natural tone, ideal for storytelling.

Additionally, the voice of Onyx, characterized by its serious tone, was noted to be particularly well-suited for storytelling. This is attributed to its formal quality, which effectively complements this type of narrative.

3.3.3. Intonation Variation

Assessing the variation in intonation within the audio can significantly influence the emotional and persuasive tone of the narrative. In the context of these stories, it plays a crucial role in effectively transmitting the values and messages intended for the children. Table 3 illustrates the intonation variations for the voice types used in OpenAI’s TTS.

Appropriate intonation captures the listener’s attention, elicits emotions, and enhances understanding of the topics discussed. An analysis of this aspect in the audio reveals that while most voices demonstrate consistent intonation variation, the Onyx voice is distinguished by its significant intonation variation. This voice incorporates appropriate pauses, enhancing the narrative of the story. In contrast, the other voices feature very brief pauses. Specifically, the Nova voice demonstrates less effective intonation, resulting in a monotonous and unvaried narration. This factor could potentially diminish children’s interest in listening to the audio.

Table 3. Intonation Variation in Story Playback

Voice	Description
Alloy	It exhibits a constant and precise intonation, closely adhering to punctuation such as commas and periods, thereby contributing to a pleasant and fluid reading experience.
Fable	It maintains a consistent and appropriate intonation in relation to punctuation.
Echo	While the intonation remains consistent and appropriate concerning punctuation, the less noticeable pauses detract from the clarity and narrative flow of the story.
Onyx	It achieves effective intonation through correctly timed pauses, enhancing the story’s narrative.
Nova	The intonation is ineffective, resulting in a monotonous, linear narration.
Shimmer	Despite lacking notable pauses in the narration, it provides good intonation within paragraphs.

3.3.4. Speech Rate

Speech rate is a critical factor that directly impacts story comprehension and the listening experience of child audiences. As detailed in Table 4, the analysis reveals that the Nova voice is inefficient due to its inconsistent rhythm, which complicates comprehension for children. Conversely, the Fable and Echo voices maintain a constant speed, but their narrative styles are overly simplistic and do not align well with the storytelling format. In contrast, the Onyx voice excels with its speech rate, ideally suited for storytelling and offers a more immersive and engaging listening experience.

Table 4. Comparison of Speech Rate in Story Playback

Voice	Description
Alloy	Due to its constant speed, the audio remains clear and understandable.
Fable	It is characterized by a rapid tempo; although pauses are included in the narration, the fast pace diminishes the expressiveness of the audio narrative.
Echo	The speed is consistent, yet it lacks sufficiently long pauses, yielding a presentation more closely resembling plain reading rather than storytelling.
Onyx	The speed is steady and calm, making it arguably the most suitable voice for narrating stories and tales.
Nova	The rhythm varies, featuring periods of quick pronunciation interspersed with slower segments.
Shimmer	It maintains a consistent speed throughout the audio.

3.3.5. Pronunciation Quality

According to OpenAI's documentation, the Text-to-Speech (TTS) model supports voices in various languages, including Spanish. Upon analysis of the thirty audio files, mispronunciations were detected in the narrations. As detailed in Table 5, the Nova voice exhibited notable inefficiencies in word pronunciation and occasionally switched languages, demonstrating instability in maintaining Spanish as the default language. Mispronunciations of character names were observed across all voices. However, despite difficulties pronouncing character names, the Onyx voice exhibited the most accurate pronunciation.

Table 5. Comparison of Pronunciation Clarity in Story Playback

Voice	Description
Alloy	The pronunciation of certain words exhibits a foreign accent, with an elongation of specific sounds. Notably, words such as "Respetar," "perspectiva," and "Ximena" are not articulated correctly in Spanish.
Fable	Pronunciations were detected in several words incorrectly pronounced in Spanish, including "Isabella" and "respeto." Additionally, there were instances where certain words, such as "igualdad," were not wholly articulated.
Echo	Pronunciation errors were noted in several Spanish words, including "hábil," "Ximena," "rumoreaba," "braille," "barrera," and "Sofía." Despite these inaccuracies, this voice exhibits a less foreign accent compared to others.
Onyx	The pronunciation of words is generally clear and precise, yet challenges arise with specific character names, such as "Isabella" and "Ximena."
Nova	The pronunciation lacks precision, exhibiting significant errors, especially in character names. Additionally, in the analysis of five stories narrated with this voice, four exhibited language switches during playback.
Shimmer	This Spanish voice exhibits clear and precise pronunciation, although it struggles with specific character names, such as "Isabella" and "Ximena".

4. Conclusions

AI-based storytelling systems are increasingly recognized as valuable tools in addressing a broad spectrum of challenges. Among these, disability stands out as a significant concern in Latin America and the Caribbean, where it affects over 85 million individuals. These systems offer innovative approaches to inclusivity and accessibility within these communities.

This article proposed designing a storytelling system tailored for Spanish-speaking children, leveraging AI-based generative technologies. This system enables the creation of personalized stories that feature child characters with disabilities, fostering a more inclusive narrative environment.

Including characters with disabilities within the narratives is a critical element that significantly enhances the promotion of diversity and equality from an early age. By leveraging the capabilities of the AI-based generative system, unique stories are generated with each request, even when the same parameters are used. This ensures a diverse and enriching experience with each interaction, underlining the system's effectiveness in fostering inclusivity.

In the various tests conducted on the system, a quantitative comparison of the diversity of the generated stories was performed using the Jaccard index criterion. The results confirmed the low similarity between the analyzed stories, indicating high content variability.

Additionally, the intonation, speed, and pronunciation quality of the synthetic voices were evaluated, highlighting opportunities for improvement in state-of-the-art audio generation technologies for storytelling.

Despite current challenges, such as resource consumption during text-to-speech conversion, lack of natural voice quality, pronunciation errors, and system latency, the proposed system is innovative and holds potential utility for educators seeking to implement inclusive educational tools in their classrooms.

Ongoing efforts are being made to develop additional systems to bridge gaps and enhance accessibility and quality of life for individuals with disabilities. Future research will collaborate with experts from disciplines such as psychology and education to critically assess the content and messages conveyed in the generated stories. Further analysis will also be conducted to determine the suitability and comprehensibility of these stories for children in Latin America.

Acknowledgements

The authors thank the Autonomous University of the State of Mexico for the support provided through project 7018/2024CIB.

References

- [1] J. R. Casar Corredera, “Inteligencia artificial generativa,” *Anales de la Real Academia de Doctores*, vol. 8, no. 3, pp. 475–489, 2023. [Online]. Available: <https://is.gd/3kMGMX>
- [2] J. Sanabria-Navarro, Y. Silveira-Pérez, D. Pérez-Bravo, and M. de Jesús-Cortina-Núñez, “Incidences of artificial intelligence in contemporary education,” *Comunicar*, vol. 77, pp. 97–107, 2023. [Online]. Available: <https://doi.org/10.3916/C77-2023-08>
- [3] S. Droubi, A. Galamba, F. L. Fernandes, A. A. de Mendonça, and R. J. Heffron, “Transforming education for the just transition,” *Energy Research & Social Science*, vol. 100, p. 103090, 2023. [Online]. Available: <https://doi.org/10.1016/j.erss.2023.103090>
- [4] S. Iruri Quispillo and C. A. Villafuerte Álvarez, “Importancia de la narración de cuentos en la educación,” *Comuni@cción: Revista de Investigación en Comunicación y Desarrollo*, vol. 13, no. 3, pp. 233–244, Sep. 2022. [Online]. Available: <https://doi.org/10.33595/2226-1478.13.3.720>
- [5] M. Pozas, C. J. G. Trujillo, and V. Letzel-Alt, “Mexican school students’ perceptions of inclusion: A brief report on students’ social inclusion, emotional well-being, and academic self-concept at school,” *Frontiers in Education*, vol. 8, 2023. [Online]. Available: <https://doi.org/10.3389/feduc.2023.1069193>
- [6] Banco Mundial, “Rompiendo barreras - inclusión de las personas con discapacidad en América Latina y el Caribe.” [Online]. Available: <https://is.gd/diWoks>
- [7] K. Ramírez. (2023) Cuentacuentos. GitHub, Inc. [Online]. Available: <https://n9.cl/pt3ls>
- [8] T. Bratitsis and P. Ziannas, “From early childhood to special education: Interactive digital storytelling as a coaching approach for fostering social empathy,” *Procedia Computer Science*, vol. 67, pp. 231–240, 2015, proceedings of the 6th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion. [Online]. Available: <https://doi.org/10.1016/j.procs.2015.09.267>
- [9] P. Juppi, “Engagement and empowerment. Digital storytelling as a participatory media practice,” *Nordicom Review*, vol. 39, 12 2017. [Online]. Available: <https://is.gd/Wo91Bc>
- [10] T. Tseng, Y. Murai, N. Freed, D. Gelosi, T. D. Ta, and Y. Kawahara, “Plushpal: Storytelling with interactive plush toys and machine learning,” in *Proceedings of the 20th Annual ACM Interaction Design and Children Conference*, ser. IDC ’21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 236–245. [Online]. Available: <https://doi.org/10.1145/3459990.3460694>
- [11] J. Haase and P. H. Hanel, “Artificial muses: Generative artificial intelligence chatbots have risen to human-level creativity,” *Journal of Creativity*, vol. 33, no. 3, p. 100066, 2023. [Online]. Available: <https://doi.org/10.1016/j.yjoc.2023.100066>
- [12] R. Li, “A “dance of storytelling”: Dissonances between substance and style in collaborative storytelling with AI,” *Computers and Composition*, vol. 71, p. 102825, 2024. [Online]. Available: <https://doi.org/10.1016/j.compcom.2024.102825>
- [13] S. Habib, T. Vogel, X. Anli, and E. Thorne, “How does generative artificial intelligence impact student creativity?” *Journal of Creativity*, vol. 34, no. 1, p. 100072, 2024. [Online]. Available: <https://doi.org/10.1016/j.yjoc.2023.100072>
- [14] S. Kalantari, E. Rubegni, L. Benton, and A. Vasalou, “When I’m writing a story, I am really good Exploring the use of digital storytelling technology at home,” *International Journal of Child-Computer Interaction*, vol. 38, p. 100613, 2023. [Online]. Available: <https://doi.org/10.1016/j.ijcci.2023.100613>
- [15] J. Su and W. Yang, “Artificial intelligence in early childhood education: A scoping review,” *Computers and Education: Artificial Intelligence*, vol. 3, p. 100049, 2022. [Online]. Available: <https://doi.org/10.1016/j.caeai.2022.100049>
- [16] S. Z. Salas-Pilco, K. Xiao, and J. Os-hima, “Artificial intelligence and new technologies in inclusive education for minority students: A systematic review,” *Sustainability*, vol. 14, no. 20, 2022. [Online]. Available: <https://doi.org/10.3390/su142013572>
- [17] H. Yu, “The application and challenges of chatgpt in educational transformation: New demands for teachers’ roles,” *Heliyon*, vol. 10, no. 2, January 2024. [Online]. Available: <https://doi.org/10.1016/j.heliyon.2024.e24289>
- [18] L. Sijing and W. Lan, “Artificial intelligence education ethical problems and solutions,” in *2018 13th International Conference on Computer Science & Education (ICCSE)*, 2018, pp. 1–5. [Online]. Available: <https://doi.org/10.1109/ICCSE.2018.8468773>

- [19] E. Valverde and P. Hernández, *TypeScript*, 2023. [Online]. Available: <https://is.gd/WICHuR>
- [20] J. Collell and A. Ferry, *CSS3 y Javascript avanzado*. Universitat Oberta de Catalunya, 2023. [Online]. Available: <https://is.gd/JrhbAi>
- [21] A. Nazir and Z. Wang, “A comprehensive survey of ChatGPT: Advancements, applications, prospects, and challenges,” *Meta-Radiology*, vol. 1, no. 2, p. 100022, 2023. [Online]. Available: <https://doi.org/10.1016/j.metrad.2023.100022>
- [22] M. Vicente-Yagüe-Jara, O. López-Martínez, V. Navarro-Navarro, and F. Cuéllar-Santiago, “Writing, creativity, and artificial intelligence. chatgpt in the university context,” *Comunicar*, vol. 77, pp. 47–57, 2023. [Online]. Available: <https://doi.org/10.3916/C77-2023-04>
- [23] J. J. López. ¿Cuántas palabra tiene un cuento o relato corto? [Online]. Available: <https://n9.cl/rdydx>
- [24] OpenAI. (2023) Text generation models. OpenAI Platform. [Online]. Available: <https://is.gd/fkWCFZ>
- [25] ——. (2023) Text to speech. OpenAI Platform. [Online]. Available: <https://is.gd/XskwW5>