AI Personalized Interactive Fiction for Young Children

Jasmine Lesner[a], Luke Murayama[a], Tony Guizar[a], Poom Phunjamaneechot[a] and Daniel Shapiro[a]

{jlesner,lmurayam,aguizaro,sphunjam,dgs}@ucsc.edu
[a]University of California, Santa Cruz

Abstract.
This paper presents a novel high-fidelity prototype designed to enable young children to create personalized, interactive stories using advanced AI technologies. Our approach addresses a gap in existing AI-driven storytelling applications by focusing specifically on young children. The prototype allows children to select story topics using emojis and engages them with vivid multimedia elements including AI-generated illustrations, sounds, and music. Our goal is to foster critical thinking through choice-based narratives that allow children to influence story outcomes. Parents can review and provide feedback on AI-generated media to ensure age-appropriate content. The system integrates large language models (LLMs) for story generation with text-to-image, text-to-sound, and text-to-music AI models to deliver an engaging multimedia storytelling experience. The paper describes the prototype system in conjunction with a story generated by it. Our prototype demonstrates the potential of child-centric AI storytelling tools to positively impact children's learning and development.

1 Introduction

'AI Dungeon' [43] and 'NovelAI' [7] are innovative, AI-driven text adventure apps that adapt stories in real time based on user inputs. They offer a wide range of narrative paths in various genres. However, these apps are not suitable for young children due to their mature content, advanced reading level, and lack of an educational focus. Recognizing this gap and the potential of generative AI to create engaging, choice-based stories for young children, we wanted to empower children with AI for crafting custom stories and enable parents to review and refine these stories.

Creating stories allows children to benefit from the novelty bias, where fresh narratives capture their attention, improving learning and memory retention [10, 16]. The IKEA effect and endowment effect mean that children value the stories they create, making them personally meaningful [28, 12]. Additionally, children engage more deeply with personalized stories, leading to improved learning outcomes like better vocabulary acquisition [24, 14].

These insights guided our design, such as having children choose story topics using emojis. The colorful, familiar, and visual nature of emojis makes them appealing and engaging, while their well-defined image sets are easily recognized as UNICODE inputs by most large language models [48].

Choice-based stories are supported with branching narratives. To keep story branches organized and coherent, we use XML to represent each story and apply XSLT [46] to develop each branch independently through LLM prompt chains. After the narrative is complete, descriptions generated by the LLMs are used to create illustrations, sounds, and music with secondary AI models. The final story XML is then transformed into Twee [19], a narrative scripting syntax that compiles into an interactive, web-based format using the Twine engine [20]. Figure 1 illustrates the prototype's interaction flow[1].

Figure 1. Interaction Flow: (1) Child chooses topics for their story. (2) AI generates the story and media galleries. (3) Parent reviews galleries, gives AI feedback. (4) AI updates story guided by parent. (5) Child enjoys their new custom story.

1.1 Contributions

This research presents a high-fidelity prototype that empowers young children to create personalized, interactive stories using advanced AI technologies. Our primary contributions include:

1. **Novel Application**: We introduce a unique application designed specifically for young children, allowing them to create stories that capture their preferences and imagination. This distinct focus on young users differentiates our approach from existing AI-driven storytelling tools aimed at older audiences.

[1] Video about the prototype [27]: https://youtu.be/TaVGem3nFrk

2. **Child-Centric Story Creation**: Our prototype puts young children at the heart of the storytelling process. By using emojis to choose story topics and making decisions that influence outcomes, kids are encouraged to think critically and practice decision-making skills while being deeply engaged in the experience.
3. **Multimedia Integration**: By incorporating AI models for text-to-image, text-to-sound, and text-to-music generation, our prototype offers a rich multimedia storytelling experience that resonates with young audiences.
4. **Parental Oversight Mechanism**: Understanding the need for appropriate content, our prototype provides parental oversight features. Our story-lines and their pictures, sounds and music are fixed after being generated, with children choosing plot developments at pre-specified junctures. We choose this design so parents can review and regenerate AI-generated elements, ensuring stories align with their expectations and educational goals.
5. **Scalable and Efficient System Design**: Our prototype is designed for horizontal scaling, utilizing a pool of AI model workers that operate on suitable GPUs. This setup accelerates content creation, making it feasible to handle multiple users simultaneously.
6. **Accessibility and Open Source Contribution**: The prototype is freely available online[2], encouraging broad accessibility and community participation. With the source code[3] released under an open-source license, other researchers and developers can build upon our work, fostering further innovation in AI-driven educational tools.

1.2 Organization

The rest of this paper is organized as follows: We begin by examining progress in automated story generation and interactive storytelling, and how these fields shape our approach. We then showcase our child-centered design with screenshots and a sample story. We describe how our prototype aids children in crafting stories, focusing on the narrative structure and our system's use of XML for representation. We provide a detailed look at how the four main components of our prototype integrate, emphasizing the Story Production process. We conclude with an evaluation of the prototype's limitations and suggest avenues for further research.

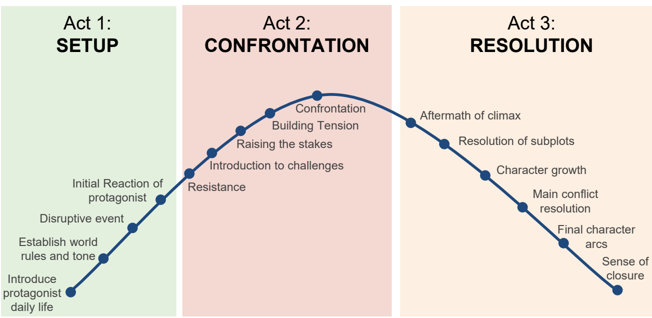

Figure 2. A three act fifteen scene story plan [9].

2 Related Work

Automatic Story Generation (ASG) has evolved significantly since its early days with rule-based systems like TALE-SPIN [30] and cognitive model-based systems such as MINSTREL [38]. Recent developments have shifted focus towards leveraging transformer models [41], such as Google's BERT [6] and OpenAI's GPT series [3], which use vast amounts of data to improve narrative generation capabilities. These models have catalyzed numerous ASG efforts [39, 1, 36, 29], demonstrating remarkable advancements in text generation that sometimes rival human writers.

For long format stories, sophisticated techniques have been developed to construct plot outlines and hierarchical representations of story plots and characters, to assist in tracking their motivations and states [21, 47, 49]. These methods ensure the story remains consistent and coherent, especially important as long format stories exceed the usable context window limits of LLMs [35, 13]. However, for children's stories, which are typically brief enough to fit within LLMs' context windows, these complex techniques are less necessary. Instead, the focus shifts to keeping the narrative simple and engaging for children through vivid visuals, lively dialogue, and character actions.

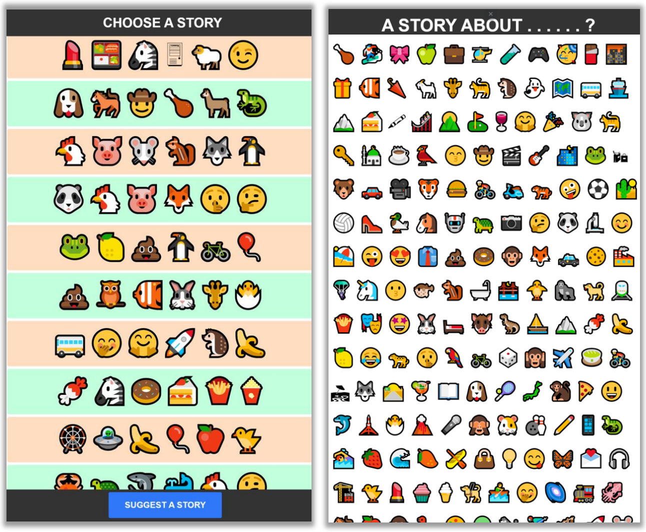

Figure 3. **Story List View** (left) contains a list of previously generated stories and offers the option to 'SUGGEST A STORY'. **Story Suggest View** (right) allows the user to pick story topics to generate a new interactive story

Our approach draws on theatrical plays and concepts like 'storylets' which facilitate dynamic storytelling suitable for young audiences [23]. Storylets are small, self-contained narrative segments that activate based on interactions or previous events, offering a flexible narrative structure. They differ from traditional scenes in a play, which follow a fixed sequence. Our work builds on these ideas by ensuring the choices offered within our stories uniquely influence the story's direction, although, like in a play, the scene order in every narration path we generate is predetermined; from the first scene and after each narration branch all our sequences of scenes follow each other in a fixed order (Figure 2) creating a 'Time Cave' structure which we will discuss in Section 4.1.

3 Child-Centered Experience

Our prototype is best understood using a walkthrough of screenshots. Figure 1 shows the five steps of our **child-centered experience**:

[2] High-fidelity prototype [26]: http://www.ufafu.com
[3] Source code [25]: https://github.com/jlesner/aipif

1. **Child chooses topics for their story.** There are hundreds of emoji topics to choose from (Figure 3 right).
2. **AI generates the story and media galleries.** The story includes illustrations, sounds, music and decisions . The galleries give parents an overview of these elements (Figure 5) .
3. **Parent reviews galleries, gives the AI feedback.** This is to be sure that AI generated media elements are suitable for their child.
4. **AI updates story guided by parent.** AI models remake unapproved media – as many times as requested.
5. **Child enjoys their new story.** The child engages with custom multimedia elements based on their selected story topics and their story time decisions (Figure 4).

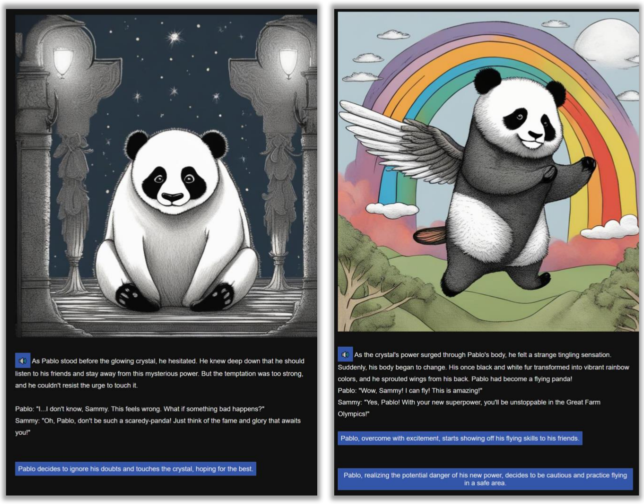

Figure 4. **Story Page View** flips through each story scene showing illustrations and text. One or more large blue buttons advance the story to the next scene determined by each button's text.

Now imagine a young child wants to create a new story for themselves. Here is the story the child created with their selected topics:

🐼🐔🐷🦊🤫🤔

SCENE 1: *Once upon a time in the animal kingdom, there was a mysterious farm where all the animals could talk and had their own unique abilities. One day, as the sun rose over the farm, the animals gathered for their morning meeting.* ***Pablo the mischievous panda****: "Good morning, everyone! What's on the agenda today?"* ***Chickie the clever chicken****: "Cluck-cluck! Today, we're going to test our animal abilities in the Great Farm Olympics!"* ***Piggy the playful pig****: "Oink-oink! I can't wait to show off my incredible mud-splashing skills!"* ***Foxy the sly fox****: "Hehe, I'll be using my cunning tricks to outsmart everyone!"* ***ACTION****: Pablo decides to participate in the Great Farm Olympics, even though he hasn't quite mastered any special abilities.* ***PICTURE****: A clever chicken with vibrant feathers and a mischievous smile.* ***MUSIC****: A lively, upbeat tune plays, creating an atmosphere of excitement and anticipation.*

SCENE 2: *As the animals prepared for the Great Farm Olympics, a mysterious figure appeared out of nowhere. It was a sneaky snake named Sammy, known for his tricks and pranks. Sammy slithered up to Pablo and whispered, "Psst! I heard there's a secret shortcut to victory in the Olympics. Follow me if you want to win!" Curiosity got the better of Pablo, and he decided to follow Sammy, unaware of the trouble that awaited him.* ***Pablo****: "What's this secret shortcut you're talking about, Sammy?"* ***Sammy****: "Oh, Pablo, it's a shortcut that will make you the champion of the Olympics! But remember, it's a secret, so don't tell anyone!"* ***ACTION****: Pablo decides to trust Sammy and follows him, ignoring the warnings from his friends.* ***PICTURE****: A sneaky snake with a mischievous grin, slithering towards Pablo.* ***MUSIC****: A suspenseful tune, hinting at the impending danger.*

SCENE 3: *As the sneaky snake Sammy led Pablo deeper into the farm, they stumbled upon a hidden cave. Inside, they discovered a glowing crystal that seemed to radiate with power. The air crackled with anticipation as Sammy whispered, "Pablo, this crystal has the ability to grant any animal a special superpower. With it, you can become the greatest athlete in the Great Farm Olympics!" Pablo's eyes widened with excitement and he couldn't resist the temptation.* ***Pablo****: "Wow, Sammy! This crystal is incredible! I can't wait to see what superpower it grants me!"* ***Sammy****: "Yes, Pablo! Just touch the crystal and your power will be revealed!"* ***ACTION 1****: Pablo decides to touch the crystal, hoping to gain an extraordinary superpower.* ***ACTION 2****: Pablo hesitates, realizing that touching the crystal might have unknown consequences.* ***PICTURE****: A mysterious cave with a glowing crystal in the center, surrounded by shadows.* ***MUSIC****: A mystical melody filled with wonder and excitement, as if something magical was about to happen.*

Before this story ends about a dozen more scenes follow (Figure 2). Figure 4 shows screenshots of what the child sees. **ACTION** elements appear as large blue buttons that advance the story. **PICTURE**, **SOUND**, **MUSIC** elements are sent to text-to-image, text-to-sound and text-to-music AI models so that each scene has its own unique illustrations, sounds and music. The screenshots in Figure 4 are the two scenes that follow when the child selects **ACTION 1**. Figure 1 and 5 show more illustrations from this same story.

4 Story Structure

To explain how our prototype aids children in creating stories, we must first discuss story structure. Recall from Section 2 that our methodology incorporates elements from theatrical plays and leverages 'storylets' to support dynamic storytelling for young audiences.

A theatrical play consists of multiple scenes, each propelling the

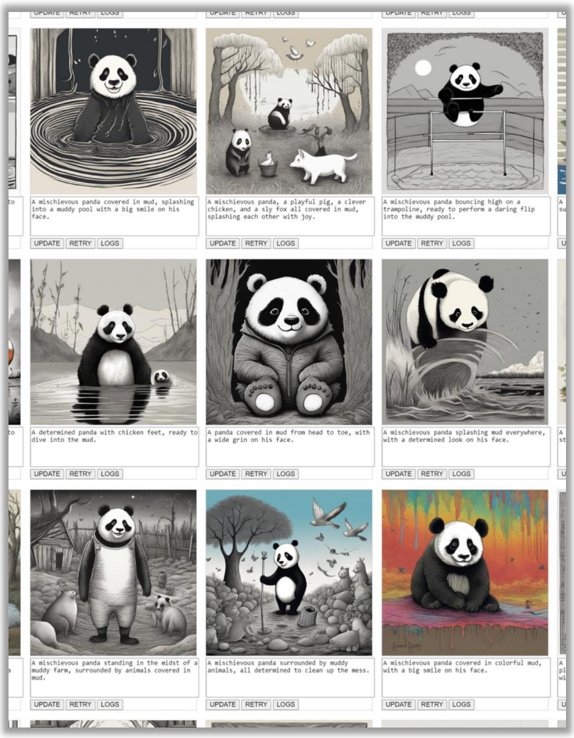

Figure 5. **Picture Gallery View** is one of three gallery views which let parents inspect illustrations, sounds, and music, and to request generation be re-attempted with the same prompt or with an updated prompt. These views are accessed from the bottom of each story page to hide them from children.

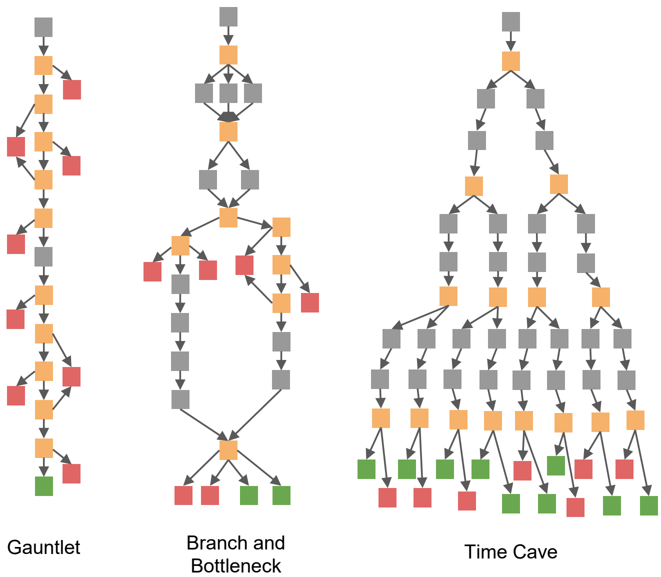

Figure 6. Branch structures for choice-based stories.

narrative through dialogue and action within a specific setting. Similarly, our stories unfold through a series of scenes that include unique illustrations, sounds, music, dialogue, and character actions, some of which children can influence. To guide the development of the narrative, we use 'story plans' (Figure 2). These predetermined plans instruct the LLM in crafting subsequent scenes. Scenes and story plans are structured in XML. In these XML representations the number of ACTION elements that a scene should offer children is controlled by a `branch_count` attribute which enables interactive, choice-driven narratives.

4.1 Choice-Based Stories

Choice-based stories engage children by requiring them to concentrate, assess situations, and make decisions that affect the story's outcome [4, 32, 22]. Such stories often serve as more effective educational tools than linear narratives [44, 45, 31].

We visualize choice-based stories as directed graphs with scenes as nodes and decisions as edges. Key configurations within these graphs include (Figure 6): **Gauntlet Stories** typically follow a main trajectory with slight detours that may lead to abrupt endings, such as a character's demise. **Branch and Bottleneck Stories** feature several important paths that merge at critical moments, necessitating an awareness of prior choices to ensure continuity. **Time Cave Stories** offer various significant, equally impactful paths, promoting exploration and multiple replays through diverse conclusions. Additional structures are Quest Stories, Open Map Stories, Sorting Hat Stories, Floating Modules Stories, Loop and Grow Stories, Trapped in Time Stories, and Hub and Spoke Stories[4] [2].

While Gauntlet stories tend to conclude sharply, often in loss or tragedy, and may not suit young children, Time Cave stories are ideal for repeated engagement due to their expansive narrative possibilities and straightforward, regular structure.

5 System Design

Children use our high-fidelity prototype to create custom choice-based stories. Figure 7 illustrates the components of our prototype.

5.1 User Interface

Children can interact with three views: **Story List View** lets them select an existing story or suggest a new one (Figure 3 left). **Story Suggest View** enables children to pick topics for a custom story by selecting and deselecting emoji images until six are chosen (Figure 3 right). **Story Page View** displays the scenes of a selected story using Twine (Figure 4).

Access buttons for the next four views are hidden at the bottom of every Story Page View and require scrolling to be seen. These views are intended for parental oversight of story content: **Picture Gallery View** shows illustrations and allows parents to read and edit the AI-generated descriptions and request new illustrations (Figure 5 left). **Music Gallery View** presents music elements, enabling parents to listen to the music, read descriptions, and request a retry (Figure 5 right). **Sound Gallery View** displays sounds for each story scene

[4] Can our XML story plan representation support all of these story structures? In structures where multiple paths converge at the same scene, it becomes essential to track the specific path taken to ensure the scene accurately reflects the player's journey. Stories for young children do not need such tracking so it is not a part of our prototype. For stories with loops, incorporating this tracking is advisable to maintain an engaging narrative.

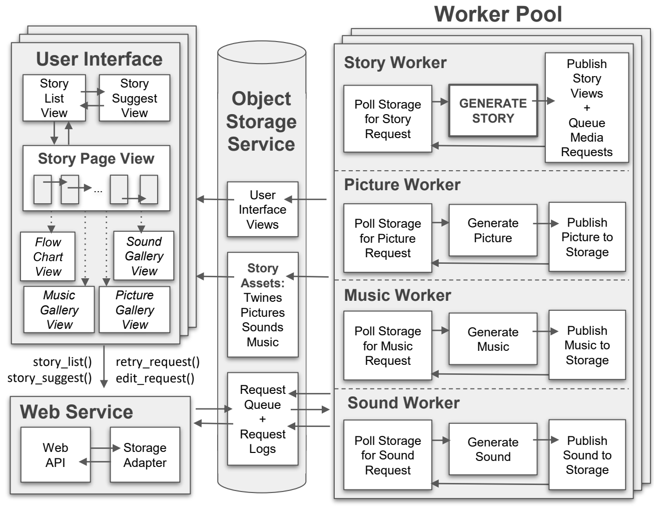

Figure 7. How the components of our prototype work together

with options for parents to listen, read AI-generated descriptions, and request retries.

5.2 *Web Service*

The user interface communicates with our web service to fetch or suggest stories. The `story_list()` function retrieves available stories, `story_suggest()` registers new story ideas, `retry_request()` triggers AI to regenerate a story element, and `edit_request()` updates media based on parental inputs.

5.3 *Object Storage*

We store three types of objects: (1) User interface elements such as story views. (2) Story-specific assets including pages, charts, and media files. (3) Logs and requests for story and media generation.

Initially, we used a file system shared among workers for Object Storage. To accommodate globally distributed GPU workers (e.g., from Vast.ai and RunPod.io), we migrated to AWS S3.

5.4 *Worker Pool*

The worker pool uses GPU workers to run AI models. Each worker continually checks storage for new requests, processes them, and returns the results to storage. When a child submits a story idea, it is stored as a request. Retrieving this request starts the 'Story Production' process (Section 6), which includes requests for pictures, sounds, and music. This asynchronous and parallel processing greatly reduces the time it takes to create a story. Consider a story with 32 narrative paths and 15 scenes each, totaling 528 AI model requests for all generated elements. Without fast GPUs operating simultaneously, generating such a story could take hours. By using multiple NVIDIA RTX3090 / RTX4090 GPUs, we can cut this time down to minutes.

6 Story Production

This section describes the story production process (Figure 8) as executed by the Story Worker component in Figure 7.

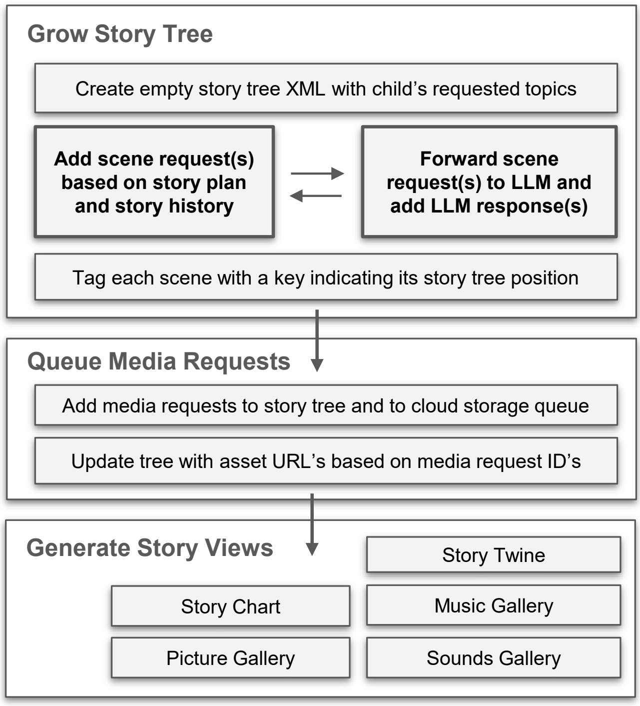

Figure 8. Three phases of story production.

Inputs include: (1) Child-selected story topic emojis. (2) Story plan XML to direct scene generation. (3) An LLM adapter for scene-by-scene writing. (4) Text-to-image model adapter for illustrations. (5) Text-to-sound model adapter for audio effects. (6) Text-to-music model adapter for background music.

Outputs are: (1) Story tree XML with all narration branches, and descriptions of illustrations, sounds, and music. (2) Generation requests for story and media, along with their logs. (3) Story Page View, Story Chart View, and media gallery views. (4) Generated illustrations, sounds, and music from the respective workers.

6.1 *Phase 1: Grow Story Tree*

We employ forward chaining to grow story trees as follows:

1. Create an empty story tree XML with the desired story plan and child's selected topics. Populate the story XML with an LLM request (Figure 9 top) to generate the initial scene.
2. Send the request via an API bridge[5], attaching the resulting scene description back into the story tree (Figure 9).
3. Use XSLT rules to insert LLM requests for subsequent scenes into the story tree (Figure 10). Parallel growth occurs if earlier scenes branched.
4. Send out pending requests via an API bridge and update the story tree XML with responses.

Steps 3 and 4 repeat to grow each branch until it reaches the last scene in the configured story plan (Figure 2). Scenes that offer decisions use a branched scene template shown in Figure 11 with responses that fork the story . The story branches because the LLM is instructed

[5] Our **API bridge** scans XML for recognizable requests, issues them to the appropriate API, and embeds the responses back into the XML. When XML contains multiple requests, we can issue them concurrently to reduce overall latency.

to "Please continue the story from the last `protagonist_reaction` above.." as shown in Figure 10.

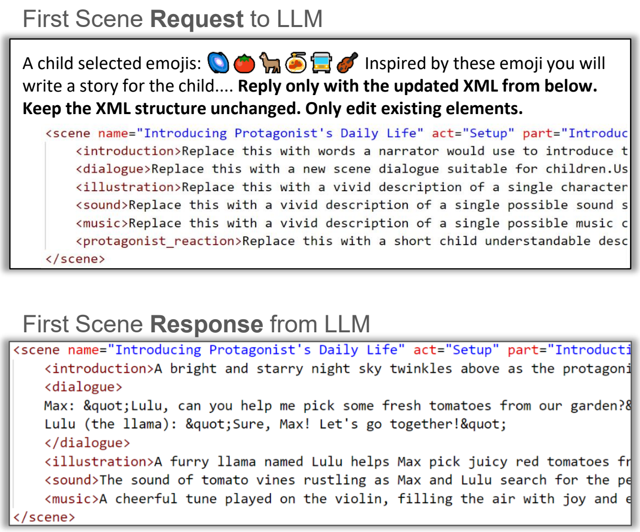

Figure 9. LLM creates the story's first scene.

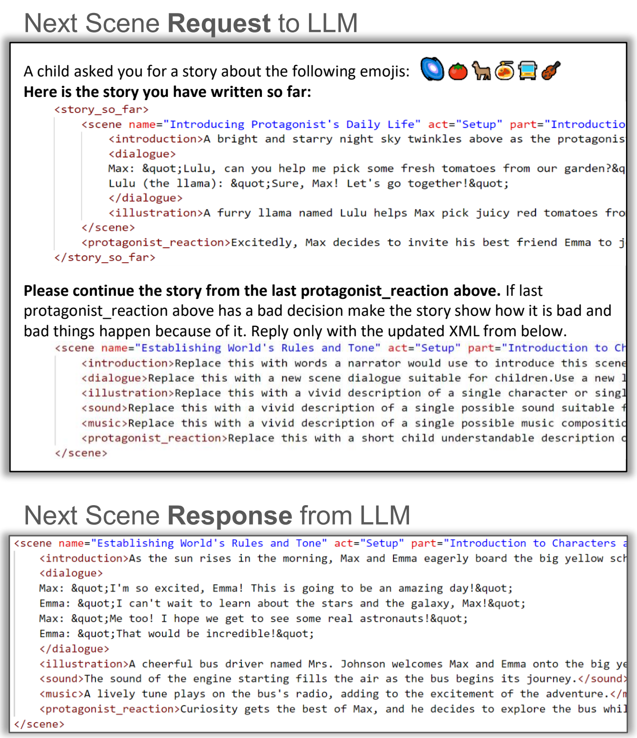

Figure 10. Subsequent scene creation by the LLM, The `story_so_far` accumulates the story down the story branch for coherence.

6.2 *Phase 2: Queue Media Requests*

Each story scene includes an illustration, sound, and music. The LLM describes these in text, which subsequent AI models render as actual media files. This is accomplished asynchronously by a pool of workers (Section 5.4), each request receiving a unique ID used in the URL for the published media.

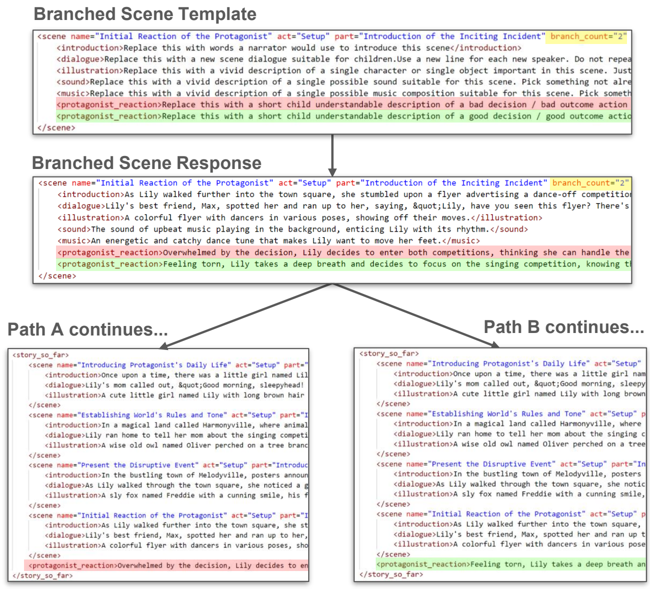

Figure 11. Control of branching in the story using the `branch_count`, `story_so_far` and `protagonist_reaction` elements.

6.3 *Phase 3: Generate Story Views*

In the final phase, story views are created:

1. Story Page View (Figure 4) is a SugarCube [8] styled Twine generated using Twee [19].
2. Story Chart View is a Mermaid.js[37] flow chart which shows parents a summary of the story and its branching pattern.
3. Story Gallery Views (Figure 5) are vanilla HTML and enable parents to examine story media (pictures, sounds, music) and request AI to retry specific ones.

After this last phase the result is a fully-formed interactive story with text, illustrations, sounds, and music. Now parents can review the generated content, and children can enjoy it.

7 Limitations and Future Work

Our prototype has limitations and areas for future improvement:

Story Quality - In Section 3, we demonstrated story quality using GPT-3.5 instead of the more advanced GPT-4. Future work will upgrade our prototype with the latest advancements in language models and multimedia integration, including text-to-image, text-to-sound, and text-to-music technologies. Further, we want to explore the potential of AI to enhance stories with rhymes, riddles, quizzes and AI-generated characters [33].

Converging Branches - Our LLM currently writes each story branch independently, leading to different branches occasionally converging into common story tropes, an issue that persists even with GPT-4. To maintain distinct narratives in each branch, we need to refine our LLM prompts to consider the story developments in other branches. This branch convergence is visible in Story Chart views which show a flow chart of all story branches but it takes multiple readings of a story with different choices to spot the convergence.

Narration Vocabulary - Stories with too many unfamiliar words can hinder word learning [15]. GPT-4 can use an appropriate vocabulary for young children, but GPT-3.5 and weaker models can struggle. Ongoing research into lexical simplification could improve LLM outputs for young audiences [40], making our prototype an ideal testbed for these advancements.

Parental Control - Parents can request retries of specific illustrations, sounds, or music. Extending this control to story text edits in a publicly hosted prototype designed for children raises safety and vandalism concerns. To evolve the prototype in this direction requires the addition of user accounts so each family has their own set of stories.

Content Safety and Bias - Large language models reproduce societal biases and pose a risk of inappropriate themes, stereotypes, or values for young audiences. While parental oversight features allow for content review and regeneration, we need more robust strategies to ensure age-appropriate and culturally diverse content. We expect future work to examine psychological impacts of AI-generated stories on children in collaboration with child psychologists and education experts.

Sound Generation - Integrating child-friendly sound effects has been challenging. Available text-to-sound models like Suno Bark [18] may generate unsettling sounds, and no model reliably produces suitable sounds like animal noises. Our prototype implements an LLM-powered search for royalty-free sounds [11] but that is also unreliable. Given this gap our makeshift interim solution for scene sounds involves having Suno Bark narrate scene introductions.

Music Looping - Unlike scene sounds that play once, music loop continuously and transition smoothly. Attempts with LLM-generated ABC music notation [42] were unsuccessful. MusicGen [5] offers variety but its clips are difficult to loop seamlessly. We currently use fading techniques to blend transitions.

Style and Aesthetics - Children select story topics, and parents may request AI retry specific story elements, but our prototype does not allow user control over story plans or style or aesthetics. These settings are globally configured; using multiple configurations simultaneously requires running several instances with different URLs.

Instruction Adherence - GPT-3.5 and weaker models need prompt tuning to reliably respond, for example to use a desired format such as XML. We have verification checks, a cleanup process, and a retry mechanism to handle their failures. This approach allows us to use faster but less reliable LLMs for testing. Exploring token-level constraints to restrict the LLM decoder from generating off-target tokens is a potential future direction.

Character Depiction - Consistent character portrayal using text-to-image models is challenging. Our makeshift solution to illustrate each character only once is limiting and does not always work. We are exploring fine-tuning our own LoRAs [17] for Stable Diffusion XL [34], which involves generating character training data and requires an additional ~ten minutes of GPU time per story.

Load Balancing - Our public prototype provides access to existing stories. However, the generation of new stories is restricted to the hours of scheduled demonstrations. To evolve our prototype, we need load balancing. This will enable us to dynamically adjust worker pools based on demand and cost variations, particularly when GPU rental prices drop below a specified threshold.

Prototype Evaluation - Our prototype shows the potential of AI-driven interactive storytelling for children, but rigorous user evaluation is essential for validating its effectiveness and guiding future improvements. We plan to enhance stories with comprehension questions with the aim to measure learning and gather data on which story elements and story types are more effective.

8 Conclusion

This study introduces a prototype for AI-driven interactive fiction designed for young children. It lets children craft their own interactive stories using text, images, sounds, and music, thanks to advanced AI technologies. The design focuses on young users, allowing them to use choices to shape story outcomes enhancing their creativity and critical thinking.

Our contributions include: (1) A unique application tailored for young users, distinguishing it from AI storytelling tools intended for older audiences. (2) Effective integration of multimedia elements to create an immersive storytelling experience. (3) Incorporation of parental oversight features to ensure content appropriateness and support educational goals. (4) A scalable design with broad accessibility, with open-source code available for community use and enhancement.

Developing this prototype provided valuable insights: (1) Age-appropriate content generation is crucial, requiring careful prompt engineering and model selection. (2) Balancing story coherence with interactive elements is challenging, especially in branching narratives. (3) Parental oversight is critical in AI-generated content for children, necessitating user-friendly review mechanisms. (4) Multimedia integration enhances engagement but also introduces technical challenges.

While our prototype is a step forward, several areas need refinement: (1) Enhancing story quality and coherence, especially in branching narratives. (2) Improving multimedia integration, focusing on seamless music looping and consistent character depiction. (3) Expanding control over story customization, including style and aesthetics. (4) Implementing robust content safety measures and bias mitigation strategies. (5) Conducting comprehensive user studies to validate effectiveness and to guide future improvements.

We anticipate that systems like ours will play a growing role in children's entertainment and education. By addressing the identified challenges and building on our initial successes, we aim to create a tool that can support the cognitive and creative development of young children.

Author Contributions

J. L. developed the prototype's LLM story features and web app. L. M., T. G., and P. P. evaluated and integrated AI models for illustrations, sounds, and music. J. L. conducted the literature review and wrote the manuscript. All authors reviewed and edited the work. D. S. supervised the project.

References

[1] P. Ammanabrolu, W. Cheung, W. Broniec, and M. O. Riedl. Automated storytelling via causal, commonsense plot ordering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 5859–5867. AAAI Press, 2021. doi: 10.1609/aaai.v35i7.16733. URL https://ojs.aaai.org/index.php/AAAI/article/view/16733. AAAI Technical Track on Humans and AI.

[2] S. K. Ashwell. Standard Patterns in Choice-Based Games. Online, Jan 2015. URL https://heterogenoustasks.wordpress.com/2015/01/26/standard-patterns-in-choice-based-games.

[3] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. Language models are few-shot learners, 2020.

[4] J. Camingue, E. F. Melcer, and E. Carstensdottir. A (Visual) Novel Route to Learning: A Taxonomy of Teaching Strategies in Visual

Novels. In *International Conference on the Foundations of Digital Games (FDG '20)*, page 13, Bugibba, Malta, 2020. ACM. doi: 10.1145/3402942.3403004.

[5] J. Copet, F. Kreuk, I. Gat, T. Remez, D. Kant, G. Synnaeve, Y. Adi, and A. Défossez. Simple and Controllable Music Generation, 2023.

[6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.

[7] E. Dogan. Novelai. Online, 2024. URL https://novelai.net/.

[8] T. M. Edwards. SugarCube. Online, 2013. URL https://www.motoslave.net/sugarcube/2/.

[9] S. Field. *Summary of Screenplay*. Dell Publishing Company, 1994.

[10] N. C. Foley, D. C. Jangraw, C. Peck, and J. Gottlieb. Novelty enhances visual salience independently of reward in the parietal lobe. *Journal of Neuroscience*, 34(23):7947–7957, 2014.

[11] Freesound Team. Freesound, 2024. URL https://freesound.org. Accessed: 2024-05-02.

[12] S. A. Gelman, E. M. Manczak, and N. S. Noles. The nonobvious basis of ownership: Preschool children trace the history and value of owned objects. *Child Development*, 83(5):1732–1747, 2012.

[13] R. Han, H. Chen, Y. Tian, and N. Peng. Go back in time: Generating flashbacks in stories with event temporal prompts, 2022.

[14] J. Horst, K. Parsons, and N. Bryan. Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology*, 2:17, 02 2011. doi: 10.3389/fpsyg.2011.00017.

[15] J. S. Horst. Context and repetition in word learning. *Frontiers in psychology*, 4:35196, 2013. doi: 10.3389/fpsyg.2013.00149.

[16] J. S. Horst, L. K. Samuelson, S. C. Kucker, and B. McMurray. What's new? Children prefer novelty in referent selection. *Cognition*, 118(2): 234–244, 2011. ISSN 0010-0277. doi: j.cognition.2010.10.015.

[17] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen. Lora: Low-rank adaptation of large language models, 2021.

[18] S. Inc. Bark: Text-prompted generative audio model. Online, 2023. URL https://github.com/suno-ai/bark. MIT License.

[19] Interactive Fiction Technology Foundation. Twee 3 specification. Online, 2021. URL https://github.com/iftechfoundation/twine-specs/blob/master/twee-3-specification.md.

[20] Interactive Fiction Technology Foundation. Twine: An open-source tool for telling interactive, nonlinear stories. Online, 2021. URL https://twinery.org.

[21] J. Kelly, A. Calderwood, N. Wardrip-Fruin, and M. Mateas. There and back again: Extracting formal domains for controllable neurosymbolic story authoring. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 19(1):64–74, Oct. 2023. doi: 10.1609/aiide.v19i1.27502. URL https://ojs.aaai.org/index.php/AIIDE/article/view/27502.

[22] K. Khan, K. Nelson, and E. Whyte. Children choose their own stories: the impact of choice on children's learning of new narrative skills. *Journal of Child Language*, 41(4):949—-962, 2014. doi: 10.1017/S0305000913000160.

[23] M. Kreminski and N. Wardrip-Fruin. Sketching a map of the storylets design space. In *Interactive Storytelling: 11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5–8, 2018, Proceedings 11*, pages 160–164. Springer, 2018.

[24] N. Kucirkova, D. Messer, and D. Whitelock. Parents reading with their toddlers: The role of personalization in book engagement. *Journal of Early Childhood Literacy*, 13, 12 2012. doi: 10.1177/1468798412438068.

[25] J. Lesner, L. Murayama, T. Guizar, P. Phunjamaneechot, and D. Shapiro. Source Code for AI Personalized Interactive Fiction for Young Children. Online, 2023. URL https://github.com/jlesner/aipif.

[26] J. Lesner, L. Murayama, T. Guizar, P. Phunjamaneechot, and D. Shapiro. Prototype of AI Personalized Interactive Fiction for Young Children. Online, 2023. URL http://www.ufafu.com.

[27] J. Lesner, L. Murayama, T. Guizar, P. Phunjamaneechot, and D. Shapiro. Demonstration of AI Personalized Interactive Fiction for Young Children. Online, 2024. URL https://youtu.be/TaVGem3nFrk. Video.

[28] L. E. Marsh, P. Kanngiesser, and B. Hood. When and how does labour lead to love? The ontogeny and mechanisms of the IKEA effect. *Cognition*, 170:245–253, 2018. ISSN 0010-0277. doi: 10.1016/j.cognition.2017.10.012.

[29] L. Martin, P. Ammanabrolu, X. Wang, W. Hancock, S. Singh, B. Harrison, and M. Riedl. Event representations for automated story generation with deep neural nets. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), April 2018. ISSN 2159-5399. doi: 10.1609/aaai.v32i1.11430.

[30] J. R. Meehan. *The Metanovel: Writing Stories by Computer.* PhD thesis, Yale University, 1976.

[31] E. F. Melcer, K. M. Grasse, J. Ryan, N. Junius, M. Kreminski, D. Squinkifer, B. Hill, and N. Wardrip-Fruin. Getting academical: A choice-based interactive storytelling game for teaching responsible conduct of research. In *International Conference on the Foundations of Digital Games (FDG '20)*, page 12, Bugibba, Malta, 2020. ACM. doi: 10.1145/3402942.3403005.

[32] D. P. Mundy and R. Consoli. Here be dragons: experiments with the concept of 'choose your own adventure' in the lecture room. *Innovations in Education and Teaching International*, 50(2):214–223, 2013. doi: 10.1080/14703297.2012.760869.

[33] P. Pataranutaporn, V. Danry, J. Leong, P. Punpongsanon, D. Novy, P. Maes, and M. Sra. AI-generated characters for supporting personalized learning and well-being. *Nature Machine Intelligence*, 3(12): 1013–1022, 2021. doi: 10.1038/s42256-021-00417-9.

[34] D. Podell, Z. English, K. Lacey, A. Blattmann, T. Dockhorn, J. Müller, J. Penna, and R. Rombach. SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis, 2023.

[35] H. Rashkin, A. Celikyilmaz, Y. Choi, and J. Gao. PlotMachines: Outline-conditioned generation with dynamic plot state tracking. In B. Webber, T. Cohn, Y. He, and Y. Liu, editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4274–4295, Online, Nov 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.emnlp-main.349. URL https://aclanthology.org/2020.emnlp-main.349.

[36] A. See, A. Pappu, R. Saxena, A. Yerukola, and C. D. Manning. Do massively pretrained language models make better storytellers? In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, pages 843–861, Hong Kong, China, November 2019. Association for Computational Linguistics. doi: 10.18653/v1/K19-1079. URL https://aclanthology.org/K19-1079.

[37] C. to Mermaid. Mermaid: Generate diagrams from markdown-like text. Online, 2014. URL https://mermaid.js.org/.

[38] S. R. Turner. *Minstrel: A computer model of creativity and storytelling*. PhD thesis, University of California, Los Angeles, 1992.

[39] J. Urbanek, A. Fan, S. Karamcheti, S. Jain, S. Humeau, E. Dinan, T. Rocktäschel, D. Kiela, A. Szlam, and J. Weston. Learning to speak and act in a fantasy text adventure game. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 673–683, Hong Kong, China, Nov 2019. Association for Computational Linguistics. doi: 10.18653/v1/D19-1062. URL https://aclanthology.org/D19-1062.

[40] M. Valentini, J. Weber, J. Salcido, T. Wright, E. Colunga, and K. von der Wense. On the automatic generation and simplification of children's stories. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 3588–3598. Association for Computational Linguistics, 2023.

[41] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[42] C. Walshaw. The abc music standard 2.1 (dec 2011), 2021.

[43] N. Walton. AI Dungeon. Online, 2019. URL https://aidungeon.com/. Text adventure game using AI to generate content.

[44] S. Watson, N. Vannini, M. Davis, S. Woods, M. Hall, L. Hall, and K. Dautenhahn. Fearnot! an anti-bullying intervention: Evaluation of an interactive virtual learning environment. In *Proceedings of AISB'07*, pages 446–452, Newcastle University, Newcastle upon Tyne, April 2007. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour (AISB).

[45] Weng, Jui-Feng and Kuo, Hsiu-lien and Tseng, Shian-Shyong. Interactive storytelling for elementary school nature science education. In *2011 IEEE 11th International Conference on Advanced Learning Technologies*, pages 336–338. IEEE, 2011.

[46] World Wide Web Consortium. Extensible Stylesheet Language Transformations (XSLT) Version 3.0. Online, 2017. URL https://www.w3.org/TR/xslt-30/.

[47] K. Yang, D. Klein, N. Peng, and Y. Tian. Doc: Improving long story coherence with detailed outline control. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5730–5743. Association for Computational Linguistics, 2023. doi: 10.18653/v1/2023.acl-long.510.

[48] Y. Zhou, P. Xu, X. Wang, X. Lu, G. Gao, and W. Ai. Emojis decoded: Leveraging chatgpt for enhanced understanding in social media communications, 2024.

[49] H. Zhu, A. Cohen, D. Wang, K. Yang, X. Yang, J. Jiao, and Y. Tian. End-to-end story plot generator, 2023.