

AIStory: design implication of using generative arts AI for visual storytelling

ARIEL HAN*, University of California, Irvine, USA

ZHENYAO CAI, University of California, Irvine, USA

SEUNGMIN JEONG, Korea Telecom (KT), South Korea

SUN MYUNG CHOI, Hammie hamster Labs, Inc, USA

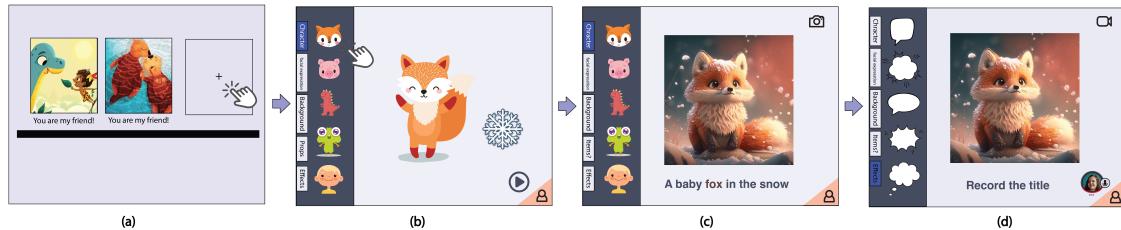


Fig. 1. (a) main page where children can review their stories and move on to workspace, (b) children can choose characters, backgrounds, and props for story creation from the stickers, (b) the prompt (textural data) as captions of the stickers will be sent to AI art generator API and generate emergent images, (c) description of the images will be displayed, (d) children can narrate their stories and they can move on to the next page.

The study examines the design implication of leveraging generative AI for literacy development for children. We sought to elicit insights on the applicability of generative AI for educational purposes from various stakeholders (i.e., parents, therapists, teachers, and AI researchers). We recruited nine participants to elicit their opinions on designing a visual narrative app with generative AI. We examined the opportunities and limitations of the current generative AI tools. Using the implications from our evaluation, we propose AIStory, an AI-powered visual storytelling application prototype that can be used for children's creative expression, storytelling, and literacy development.

CCS Concepts: • Computer systems organization → Embedded systems; Redundancy; Robotics; • Networks → Network reliability.

Additional Key Words and Phrases: Child-centred AI, AI interaction, Creativity, Storytelling

ACM Reference Format:

Ariel Han, Zhenyao Cai, Seungmin Jeong, and Sun Myung Choi. 2023. AIStory: design implication of using generative arts AI for visual storytelling. In *CHI 2023 Workshop on Child-centred AI Design: Definition, Operation and Considerations*, April 23, 2023, Hamburg, Germany. ACM, New York, NY, USA, 4 pages.

1 INTRODUCTION

As AI/ML models have evolved, Generative AI has become a hot topic for producing creative digital content from qualitative datasets (i.e., audio, video, and images). Generative AI includes various modalities, from text to image, text

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2023 Copyright held by the owner/author(s).

Manuscript submitted to ACM

53 to video, and text to audio; they receive text input as prompts to generate unique content [9]. Recently, Text to image
54 AI - art generators have increased in attention. DALL-E released their API [7], MidJourney published their service [6],
55 and Stable Diffusion released open-source codes public for free [4] in the summer of 2022, which can disrupt creative
56 industries such as gaming and animation [10]. These generative AI tools afford easier and faster ways of producing
57 creative content (e.g., music, animation, story, and visual arts), which can be effective in the creative processes for
58 creators [12]. However, due to the fast-paced, rapidly changing nature of the technology, the usability of generative AI
59 tools and their applicability in educational settings has been under-explored [1].
60

61 In this study, we aim to explore opportunities for incorporating generative AI tools for educational purposes.
62 Specifically, our goal is to design personalized, adaptive visual storytelling tools for young learners powered by
63 generative AI. To that end, we set out to understand how learners interact and collaborate with AI-based creative tools
64 to design safe, accessible, and user-friendly interactions and interfaces for visual storybook creation with AI systems.
65

66 Accordingly, we designed a prototype, AIStory, a web-based app where users can create a visual story with the AI
67 art generator (see Figure 1). In the app, there are built-in stickers that are categorized for story creation (i.e., characters,
68 backgrounds, props, speech bubbles, and effects). The character section, for example, has options like animals, dinosaurs,
69 monsters, humans, and fairytales. Each sticker has a caption as a prompt input that will be sent to generative arts AI
70 API to create high-quality images for the story creation (see Figure 1).
71

72 **2 APPLICABILITY OF LEVERAGING AI ART GENERATORS**

73 We first reviewed the current AI art generators on the market (DALL-E, Stable Diffusion, and Midjourney). Then we
74 interviewed nine participants (i.e., a speech-language pathologist, an art therapy practitioner, a mother of a child with
75 a communication disorder, a mother of a child with autism, a computer science student in college, two elementary
76 school teachers, and two AI researchers). We examined the functionality and applicability of the AI art generators for
77 educational purposes and discussed the strengths and difficulties of designing visual narrative apps.
78

79 A majority of participants expressed interest in exploring these new tools, but also expressed concerns about
80 their educational potential. First, the current AI art generators (Dall.E, Stable Diffusion) lack context from a novice's
81 perspective, especially for kids, which highlights the importance of creating kid-friendly interfaces. The second
82 consideration is AI ethics, the potential for generating images that are biased and inappropriate, which underscores the
83 necessity of providing protection against inappropriate image generation. The third problem is a lack of ownership of
84 the digital outcome because AI generates everything for the user, which emphasizes providing customization options
85 for users' creation processes.
86

87 **3 CONCEPT DESIGN**

88 With this implication, we designed a prototype called AIStory with kid-friendly interfaces for youths to create a visual
89 story easily by following the steps (see Figure 2). To mitigate the pitfalls of current AI arts tools, we propose sticker-based
90 interfaces where users can pick stickers to customize their stories. After users choose a sticker, the system then sends
91 curated prompts as captions of the stickers (i.e., thorough description of the image) with highlighted keywords and
92 styles (i.e., digital arts, 3d arts) to the back end as a safeguard against generating inappropriate image [5](see Figure 2).
93

94 To achieve this, we stored predefined keywords (i.e., characters, props, settings) that closely tie to the story's context
95 and a specific art style that the AI generator can proceed (i.e., digital arts, pure background, wide angles, natural light,
96 oil painting, 8K, artist names). In the back end, we added a system that combines the user's input with the stored
97

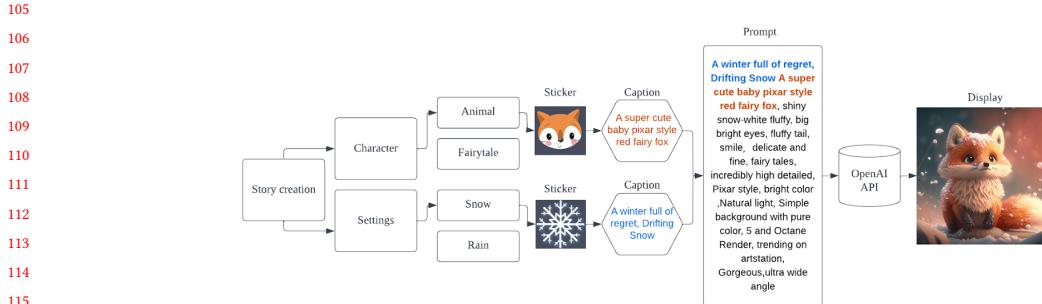


Fig. 2. Image creation workflow: once the user enters story creation mode, users can choose stickers under each category, the sticker captions will be merged as one coherent prompt based on prompt engineering that we pre-built in the database that will be sent to AI art generator API to generate unique, high-quality images for users to use for their visual storytelling.

keywords for consistent scenes by sending the combined prompt to the AI generator API so that the AI generator can proceed with the thorough description of the images and send the final output to the front-end (see Figure 2).

4 DISCUSSION AND FUTURE WORK

The preliminary study examined the opportunities and limitations of leveraging text-to-art AI generators for visual storytelling. AIStory web app that we developed allows users to customize their own stories and generate unique images in collaboration with the AI generator. Generative AI showed the potential to provide agencies for children to create digital content as producers that promote their active engagement in the creative processes.

Although employing AI-based generation tools to support learning and creativity in educational settings is still in the early stages of research. AI-based tools' effectiveness, validity, and safety remain unclear [1, 11]. Specifically, with the dynamic of an educational environment, understanding learners' perception of learners' interactions with AI and their strategies to collaborate with it is essential for the successful employment of the systems [3, 8].

Therefore, it is necessary to incorporate educators, parents, and children's perspectives on the application. The study will be conducted with a design-based research (DBR) approach to design, develop, and evaluate AI-powered learning applications. We will include participatory design sessions with educators and parents to validate an age-appropriate convergent curriculum that benefits visual literacy and creative expressions for young learners. The qualitative study will include focus group interviews, observations, and artifact analysis from students' digital creations. We will update features, functions, and interfaces to facilitate a safe, user-friendly system for better interaction and educational outcomes.

New AI tools and applications such as ChatGPT (i.e., dialogue AI for smooth conversation) [2] continue to be released, opening up new opportunities for leveraging them for educational purposes. We aim to explore such tools for implementing better interactivity (child-AI interaction) and educational value for young learners.

The project's contribution will help inform the field about the challenges in leveraging AI in elementary grades for creative storytelling activities. We will conduct studies with children to observe and analyze the learning tools' validity, efficacy, and usability. We will also highlight how children interact with AI agent and their strategies for collaborating with AI tools for their creative processes.

ACKNOWLEDGMENTS

To Robert, for the bagels and explaining CMYK and color spaces.

REFERENCES

- [1] Safinah Ali, Daniella DiPaola, Irene Lee, Jenna Hong, and Cynthia Breazeal. 2021. Exploring Generative Models with Middle School Students. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3411764.3445226>
- [2] ChatGPT. 2022. ChatGPT: Optimizing Language Models for Dialogue. <https://openai.com/blog/chatgpt/>
- [3] Helen Crompton, Mildred V. Jones, and Diane Burke. 2022. Affordances and challenges of artificial intelligence in K-12 education: a systematic review. *Journal of Research on Technology in Education* 0, 0 (Sept. 2022), 1–21. <https://doi.org/10.1080/15391523.2022.2121344> Number: 0 Publisher: Routledge _eprint: <https://doi.org/10.1080/15391523.2022.2121344>
- [4] Stable Diffusion. 2022. Stable Diffusion Public Release. <https://stability.ai/blog/stable-diffusion-public-release>
- [5] Vivian Liu and Lydia B Chilton. 2022. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–23. <https://doi.org/10.1145/3491102.3501825>
- [6] Midjourney. 2022. Midjourney. <https://midjourney.com/>
- [7] OpenAI. 2022. DALL-E 2. <https://openai.com/dall-e-2/>
- [8] Igor Pesek, Novica Nosović, and Marjan Krašna. 2022. The Role of AI in the Education and for the Education. In *2022 11th Mediterranean Conference on Embedded Computing (MECO)*. 1–4. <https://doi.org/10.1109/MECO55406.2022.9797189> ISSN: 2637-9511.
- [9] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. 2022. Hierarchical Text-Conditional Image Generation with CLIP Latents. <https://doi.org/10.48550/arXiv.2204.06125> arXiv:2204.06125 [cs].
- [10] Kevin Roose. 2022. An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy. *The New York Times* (Sept. 2022). <https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html>
- [11] Ben Shneiderman. 2020. Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human-Computer Interaction* 36, 6 (April 2020), 495–504. <https://doi.org/10.1080/10447318.2020.1741118> Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/10447318.2020.1741118>
- [12] James Vincent. 2022. Anyone can use this AI art generator — that's the risk. <https://www.theverge.com/2022/9/15/23340673/ai-image-generation-stable-diffusion-explained-ethics-copyright-data>

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009