Zero2Story: Novel Generation Framework for Anyone

Chansung Park[1,2], Youngbin Lee[2], Sangjoon Han[2], Jungue Lee[2]
[1]ETRI, [2]MODULABS, South Korea
chansung18@etri.re.kr, {dudqls4019, jphan32, melodysdreamj}@gmail.com

Abstract

This paper explores a novel approach to collaborative storytelling where an AI generates paragraphs and provides branching options for human authors to shape the narrative in a turn-based style. It also discusses the integration of three generative AI technologies for text, image and audio within a unified platform, offering new creative possibilities across multiple media formats.

1 Introduction

The advent of large language models has profoundly transformed the landscape of text composition, enabling individuals with varying backgrounds and skill sets to engage in tasks previously reserved for domain experts. This technological advancement has democratized content creation by bridging the gap between novices and professionals[1][2].

However, especially in the context of novel writing, the barrier to entry is still high for who don't have prior writing experience[3]. Zero2Story[1] is a framework to lower the barrier by fostering enjoyment throughout the collaborative creative process between human writers and AIs in turn-based role-playing games(TRPGs) style using text, image, and audio generation technologies.

2 Method

The Zero2Story introduces a structured three-step approach to narrative generation. Initially, background information encompassing genre, locale, and mood is selected, each with ten distinct items. Based on this, up to four characters can be created with the traits of name, age, MBTI, personality, and job. Optionally, it employs Stable Diffusion(CIVIT.AI[2] 129896 and 23906 checkpoints) to generate character images aligned with the background and character information, enriching the narrative ambiance and writing experience.

Upon those information, the Zero2Story crafts the opening paragraphs with selectable narrative paths. When one of the paths is chosen by a human writer, subsequent narrative segments are generated that seamlessly dovetail with the preceding text. Also, it optionally suggests background images and audio based on the generated stories with Stable Diffusion(CIVIT.AI 93931 and 65728 checkpoints) and MusicGen to give vitality in each turn. This interplay continues until the human writer decides to end the writing and export the works. Exported works are organized in a single static HTML with appropriately placed stories and media including character images for rich reading experience.

Additionally, randomness in generative AI sometimes outputs unsatisfied contents, so Zero2Story has a regeneration control that lets human writer to recurrently refine or regenerate the generated contents with and without custom prompts from the human writer. Above all, the PaLM API[3]'s large language

---

[1]https://github.com/coding-pot/Zero2Story
[2]https://civitai.com
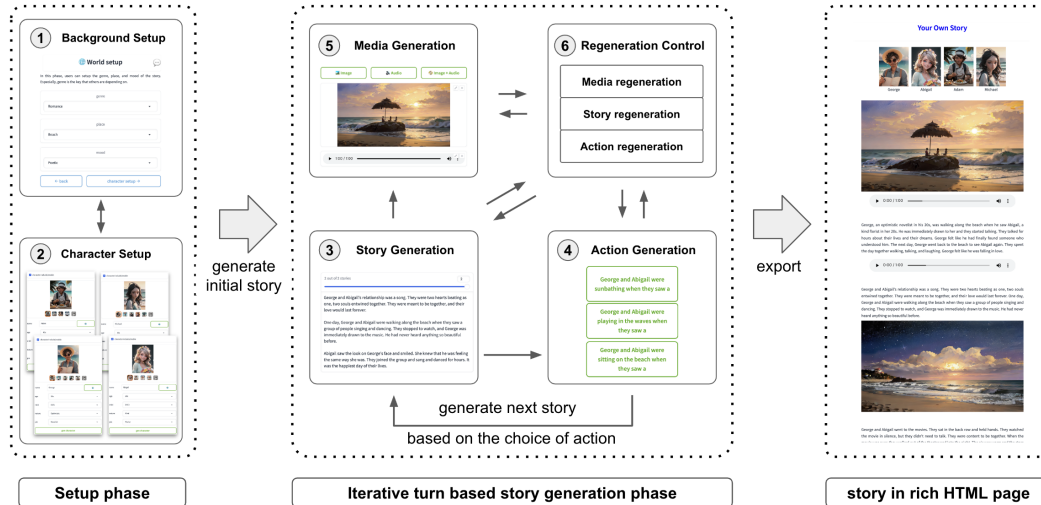[3]https://developers.generativeai.google/products/palm

Figure 1: Zero2Story framework overview.

Once setting the basic information of ① and ②, Zero2Story continuously suggests ③ stories , ④ actions, and ⑤ media in each turn while giving ⑥ regeneration control to refine the works interactively.

model is the most important component because it generates not only the stories but also appropriate prompts for image and audio generation by understanding the underlying narratives.

Lastly, Zero2Story provides a chat interface thoroughly in overall steps. Since general purpose large language model such as PaLM has a universal knowledge for various tasks, a human writer could develop ideas about the basic writing principles, background and character setups, and generated stories through the chat interface. The important point of having a chat interface is that the control difficulties that arise when generative AI is integrated as part of the system for automation can be resolved through chat interaction with generative AI itself.

3   Results

The collaborative aspect of the Zero2Story framework provided real-time narrative guidance to writers. It simplified the storytelling process, making it an enjoyable endeavor akin to a game with selectable paths and rich media, thereby democratizing story creation for ordinary individuals. Furthermore, the recurrent refinement mechanisms mitigated randomness with preliminary feedback from a user-centric evaluation underscoring the framework's efficacy in streamlining the narrative generation process while retaining creative control.

We have experimentally created few dozens of stories with Zero2Story, and we found that it adeptly generates short story text without human intervention. However, it struggles with lengthy stories and more complex settings. For instance, it sometimes does not make third and fourth characters appear in the stories or tends to end stories in the very early stage. Also, it tends to produce stories that are very cliché such as "The prince and princess lived happily ever after". Hence, collaboration between AI and human writer by giving feedback and guidance each other is crucial to control the stories interesting and dynamic.

4   Conclusion

In conclusion, the Zero2Story framework significantly simplifies story creation, making it accessible and enjoyable, especially for novices. While it adeptly generates short stories and allows for human-AI collaboration, its challenges with complex or lengthy narratives underline the importance of continued synergy between human and AI to enhance storytelling quality and diversity. Through recurrent refinements and multimedia enrichment, Zero2Story holds promise as a foundation for future endeavors in democratizing creative storytelling.

5 Ethical concerns

The emergence of AI in literature brings a ethical dilemma centered on authorship and attribution. As it crafts rich narratives, the conventional boundaries of authorship begin to blur. The quandary arises in discerning the true creator of a novel: whether it is the human authors of the texts behind the training data, developer behind Zero2Story, the individual who fine-tuned its intricacies, or Zero2Story itself. These ambiguities not only muddle the process of assigning credit, but also unsettle our long-held beliefs about creativity and originality. A pragmatic solution could be to clearly define AI's role in the artistic process. By discerning if Zero2Story is a tool or a primary creator, attribution becomes clearer. This ensures rightful credit for both human and AI contributions, addressing the ethical issues in AI-generated literature.

Additionally, employing AI in storytelling introduces ethical concerns, mainly potential bias and discrimination within AI models, that require careful examination. The adoption of AI-driven turn-based storytelling and the integration of generative AI technologies raises several ethical considerations that merit careful scrutiny. First and foremost is the potential for bias and discrimination embedded within AI models. As AI generates content and provides branching options, it may inadvertently reflect and perpetuate biases present in the training data, leading to biased narratives that reinforce stereotypes or marginalize certain groups. Ensuring fairness and inclusivity in AI-generated content is paramount, demanding ongoing vigilance in data selection and model training.

6 Acknowledgments

References

[1] Mirowski, Piotr and Mathewson, Kory W and Pittman, Jaylen and Evans, Richard. (2023) Co-Writing Screenplays and Theatre Scripts with Language Models: An Evaluation by Industry Professionals. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp.1-34.

[2] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. (2022) Wordcraft: Story Writing With Large Language Models. In 27th International Conference on Intelligent User Interfaces. ACM, Helsinki Finland, 841–852.

[3] Flower, L. and Hayes, J.R. (1981) A Cognitive Process Theory of Writing. College Composition and Communication, 32, 365-387.