Preprints.org

Article

Integrating Generative AI-Based Script Writing with Story Visualization: A Comprehensive Approach to Automated Narrative Creation

Zhen Bin It [*] , Jovan Bowen Heng , Tee Hui Teo

Article

Integrating Generative AI-Based Script Writing with Story Visualization: A Comprehensive Approach to Automated Narrative Creation

Zhen Bin It *, Jovan Bowen Heng and Tee Hui Teo

Singapore University of Technology and Design, Singapore

* Correspondence: zbhienn0920@gmail.com

Abstract

The fusion of generative AI and advanced visual synthesis technologies has opened new frontiers in automated storytelling. While large language models (LLMs) have achieved remarkable proficiency in generating coherent and emotionally engaging narratives, a consistent challenge lies in bridging the semantic gap between textual scripts and their visual interpretation. This paper presents a comprehensive framework that unites generative AI-based script writing with high-quality story visualization. We delve into cutting-edge techniques in narrative generation, explore semantic abstraction methods, and detail visual rendering pipelines powered by diffusion and multi-modal models. Our integrated architecture emphasizes semantic alignment, temporal coherence, and narrative consistency throughout the storytelling process. Experimental evaluations and qualitative case studies validate the effectiveness of the approach across diverse genres. This work aims to serve as a foundational model for the next generation of storytelling systems, paving the way for applications in entertainment, education, and interactive media.

Keywords: generative artificial intelligence; script writing

Introduction

Storytelling is one of the oldest and most universal human activities, serving as a vessel for cultural transmission, moral instruction, entertainment, and identity formation. Traditionally, storytelling involved oral narration or textual composition supported by illustrations or performance arts. In recent years, technological advances have transformed storytelling into an interactive and multimodal experience through digital media [2]. Artificial intelligence (AI), particularly generative models, now plays a pivotal role in automating and enhancing narrative creation, offering novel possibilities for creativity and engagement.

The emergence of Large Language Models (LLMs) such as OpenAI's GPT series, Meta's LLaMA, and Google's Gemma marks a watershed in natural language generation [1]. These models are trained on massive corpora of text and can generate coherent, contextually relevant, and stylistically diverse narratives spanning numerous genres. Simultaneously, advances in generative image synthesis — including models like DALL·E, Stable Diffusion, and MidJourney — have enabled the creation of photorealistic or stylistically rich images from textual prompts, expanding AI's role into visual storytelling.

Despite these successes, a significant challenge remains: these two modalities (text and image) largely operate in isolation, lacking unified frameworks for consistent cross-modal storytelling. For example, a generated script might describe a protagonist with distinctive attributes and emotional states, but the accompanying visuals may fail to capture these nuances, leading to a disjointed experience that diminishes immersion and user satisfaction [5]. This discrepancy arises due to semantic gaps, temporal misalignment, and differing contextual understanding between language and vision models.

This paper addresses these limitations by proposing a holistic, integrated framework for generative AI-driven storytelling that harmonizes script writing and story visualization. Our approach encompasses semantic abstraction, event-level representation, cross-modal alignment, and temporal coherence to produce narratives that are rich, immersive, and visually faithful. Through a detailed analysis of state-of-the-art generative techniques and an exploration of multimodal architectures, we chart a path toward automated storytelling systems capable of engaging diverse audiences in interactive and meaningful ways.

Literature Review

Automated Story Generation: Evolution, Architectures, and Ethical Frontiers

1. Historical Trajectory and the Emergence of Generative Models

Automated story generation, as a computational discipline, has evolved substantially over the past few decades. Early approaches to narrative automation were largely grounded in rule-based paradigms, utilizing symbolic logic, handcrafted ontologies, and rigid syntactic templates. Notable systems such as Tale-Spin and MEXICA exemplified these methods, relying on predefined world models and plot schemas to simulate storytelling processes. These systems afforded a high degree of interpretability and narrative structure but were often constrained by brittleness and a lack of generative flexibility. Retrieval-based methods followed, introducing case-based reasoning where existing narrative components were reused or reassembled [5]. Although retrieval-based systems addressed some generative limitations, they struggled with producing truly novel or adaptive storylines, often failing to account for nuanced semantic variation and long-term narrative progression.

The advent of deep learning and large-scale language models (LLMs), marked a transformative shift in this domain. Models such as OpenAI's GPT-2, GPT-3, and GPT-4, Meta's LLaMA, Google's Gemma, and DeepSeek-7B have demonstrated a remarkable capacity for generating coherent, context-sensitive, and stylistically rich narratives [38]. These generative models are trained on vast corpora encompassing diverse textual forms—from fiction and journalism to scientific prose— thereby capturing latent syntactic and semantic structures conducive to storytelling.

With autoregressive decoding strategies and transformer-based architectures, these models maintain context over longer spans, adapt dynamically to prompts, and mimic stylistic features of human-authored literature. However, while they offer unprecedented scalability and creativity, persistent challenges remain. These include maintaining thematic consistency over extended narratives, managing character development, preserving causal relationships, and aligning outputs with ethical and genre-specific expectations. Furthermore, the "black-box" nature of LLMs poses significant interpretability concerns, which complicate their integration into user-facing storytelling tools.

2. Fine-Tuning Strategies and Domain-Specific Optimization

One of the most promising directions in automated storytelling involves the fine-tuning of pre-trained LLMs on task-specific or domain-specific corpora. Fine-tuning enables model behaviour to be aligned more closely with desired narrative qualities, such as coherence, fluency, emotional tone, and age-appropriateness. Supervised Fine-Tuning (SFT) and Reinforcement Learning with Human Feedback (RLHF) have emerged as leading methodologies in this regard.[18]

A compelling example is the fine-tuning of the DeepSeek-7B model for children's storytelling. This work incorporated a dual-objective optimization: improving narrative structure and integrating a content moderation layer to filter inappropriate themes. By leveraging a child-oriented dataset annotated for lexical simplicity, moral themes, and emotional arcs, the model was fine-tuned to generate content suitable for young readers. Evaluative metrics including ROUGE-1, ROUGE-2,

METEOR, and BERTScore exhibited significant improvements post-fine-tuning, reflecting gains in both linguistic fidelity and semantic appropriateness.

Importantly, the inclusion of a content moderation subsystem represents a paradigm shift from purely generative concerns to responsible generation. This enables real-world applications where automated narratives must conform to pedagogical goals, cultural sensitivities, and platform-specific content policies. In educational settings, for example, such models can serve as personalized story generators that adapt to curriculum objectives while ensuring safety and inclusivity.

Beyond SFT, unsupervised pretraining on filtered domain-specific corpora—such as folklore, mythologies, or science fiction—has been explored to endow LLMs with genre-specific conventions. Curriculum learning and multi-phase training strategies have also been adopted, where models first learn basic narrative constructs before being exposed to complex plot devices and genre tropes. These techniques enhance a model's capacity to internalize both the structural scaffolding and thematic stylization necessary for authentic storytelling.

3. Event-Centric Representations and Narrative Structure

While LLMs demonstrate impressive raw generation capabilities, they often lack explicit narrative control, resulting in digressions or inconsistencies. To counteract this, a significant body of research has focused on incorporating structured event representations as intermediate abstractions in the storytelling pipeline. This approach decomposes narratives into sequences of events that follow a controlled semantic schema, providing a scaffold for coherent plot development.[23]

Martin et al. introduced a structured representation wherein each sentence is abstracted into an event tuple comprising <subject, verb, object, modifier>. This formulation simplifies complex sentences into their core semantic components, reducing data sparsity and enhancing generalization. For instance, the sentence "The camera tracks a Coelophysis through the woods" becomes <camera, track, Coelophysis, ∅>, abstracting narrative content into an interpretable and manipulable structure.

These event tuples are often enriched using linguistic resources such as VerbNet for verb classification, Named Entity Recognition (NER) for character tracking, and WordNet for synonym expansion. Such enhancements improve the model's capacity to maintain referential integrity and logical progression across scenes. Moreover, this approach allows for modular generation pipelines wherein event-to-event transitions model plot dynamics, and event-to-sentence realizations convert abstract events into fluent prose.

Empirical findings reveal that training on 2-gram event sequences, rather than full narrative graphs, yields better results in preserving local coherence without overfitting to global narrative arcs. Additionally, input-reversed unidirectional RNNs have shown surprisingly strong performance compared to more complex architectures like bidirectional LSTMs or transformers. These findings suggest that certain narrative dynamics may be better captured by simpler, sequential models—especially when training data is limited or when interpretability is prioritized.

This stream of research underscores the value of hybrid architectures that combine the linguistic richness of LLMs with the logical rigor of structured representations. Such integrations are particularly beneficial in applications requiring high-level control, such as branching storylines in interactive fiction or adaptive educational narratives.

4. Conditional Generation and the StoryGenAI Framework

To further enhance user control and genre adaptability, conditional story generation has emerged as a critical innovation. Conditional models accept auxiliary inputs—such as keywords, genre tags, desired length, or sentiment indicators—which are used to guide the generation process.

The StoryGenAI framework exemplifies this approach by treating storytelling as a conditional text-to-text generation task. Users input a desired word count, target genre, and a list of keywords, and the model generates a narrative aligned with these constraints. The architecture employs a compact GPT-2 variant with 12 decoder layers, pre-trained on a broad web corpus for linguistic

competence. For fine-tuning, genre-specific datasets were annotated with 10–20 keywords using Term Frequency-Inverse Document Frequency (TF-IDF) tokenization.[18]

Sampling strategies play a crucial role in balancing creativity and coherence. StoryGenAI integrates top-k sampling (which limits the next-token selection to the k most probable candidates) with nucleus sampling (which selects tokens from a dynamic probability mass). This hybrid approach mitigates issues such as token redundancy, repetition, and thematic drift, which are common pitfalls in naïve sampling techniques.

Evaluation results demonstrate a BLEU score of 0.704 for narratives up to 500 words—an impressive metric given the diversity of genres and input conditions. Moreover, the absence of data augmentation during validation ensures that evaluation metrics reflect genuine model generalization rather than memorization.

StoryGenAI represents a shift toward user-steerable storytelling systems, offering valuable applications in interactive fiction, game dialogue generation, creative writing education, and narrative-based therapy. It demonstrates how conditioning can be leveraged not only to enforce genre conventions but also to reflect user intent, thereby expanding the creative agency of human-AI collaboration.

5. Philosophical and Pedagogical Considerations

The rise of AI-generated narratives raises important philosophical questions about the nature of creativity, authorship, and aesthetic value. Traditionally, storytelling has been viewed as a uniquely human endeavor—intertwined with culture, emotion, and introspection. The increasing capability of machines to generate engaging and stylistically nuanced narratives challenges this notion.

In literature and media studies, debates are emerging around what constitutes originality when AI can emulate canonical authors or synthesize new literary forms. Questions of intellectual property also arise—who owns an AI-generated story? The developer? The user? The model itself?

From a pedagogical perspective, AI storytelling systems are being incorporated into classroom settings as tools for teaching writing, literature, and critical thinking. Case studies indicate that AI-generated prompts can stimulate student creativity, foster engagement, and offer exposure to diverse narrative structures. For instance, AI-generated analogies or plot twists can be used to teach figurative language or plot mechanics.[35]

Rather than replacing human creativity, AI serves as a co-creative partner, expanding the horizon of what is possible in narrative art. This aligns with constructivist educational theories that emphasize learner agency and multimodal exploration. In this context, AI becomes a collaborator in the meaning-making process rather than a deterministic author.

However, ethical guardrails are essential. Concerns include algorithmic bias, harmful stereotypes, unauthorized style emulation, and the potential erosion of traditional literary values. Implementing ethical frameworks—such as value alignment, transparency, and audience-specific filtering—ensures that AI systems contribute positively to the cultural and educational landscape.

Related Works

Generative AI in Storytelling and Multimodal Narrative Visualization

1. Evolution of Automated Story Generation

Generative Artificial Intelligence (GenAI) has radically transformed the landscape of storytelling by enabling machines to generate rich, coherent, and contextually adaptive narratives. Early automated story generation systems were predominantly rule-based, depending heavily on manually curated grammars and templates derived from classical storytelling frameworks [30]. Systems like Tale-Spin and MINSTREL followed predefined rules based on literary theories such as Vladimir Propp's 31 narrative functions or Joseph Campbell's monomyth, also known as the Hero's Journey. These systems provided high degrees of interpretability and structural control, but they

were often criticized for producing repetitive, predictable, and emotionally flat stories. The rigidity of their templates limited their capacity to adapt to new narrative genres, audience preferences, or dynamic input.

The emergence of neural network-based approaches, particularly Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, marked a substantial improvement. These models introduced a data-driven paradigm, learning to predict sequences from large narrative datasets. This enabled more flexible story progression and improved local coherence. However, RNNs and LSTMs struggled with maintaining long-range dependencies—an essential aspect of extended storytelling where plot threads must span multiple paragraphs or chapters. Consequently, although more adaptable than their rule-based predecessors, these models remained constrained in generating truly engaging and thematically consistent long-form narratives.

A paradigm shift occurred with the advent of transformer architectures, exemplified by models such as OpenAI's GPT series (GPT-2, GPT-3, GPT-4) and Meta's LLaMA models. These Large Language Models (LLMs) are trained on vast corpora of text and excel at understanding and generating human-like language across a broad range of topics and styles. The attention mechanism underlying transformers enables them to consider context over long text spans, significantly enhancing the coherence and depth of generated narratives. These models can also perform zero-shot and few-shot learning, allowing them to generate domain-specific stories or mimic literary styles without extensive retraining. However, even state-of-the-art LLMs face challenges in preserving narrative arcs, character consistency, and thematic progression in very long or branching texts—areas that are actively being addressed through reinforcement learning with human feedback (RLHF), retrieval-augmented generation (RAG), and advanced prompt engineering strategies.

2. Integration of Classical Narrative Frameworks

To enhance the quality, emotional resonance, and cultural relevance of AI-generated stories, researchers are increasingly integrating classical narrative theories into generative models. These frameworks offer structural blueprints that guide story development and help ensure coherent narrative arcs.

For instance, Freytag's Pyramid—a five-act structure comprising exposition, rising action, climax, falling action, and resolution—has been employed in systems like Dynamic StarCraft to maintain narrative pacing and emotional build-up. This structural embedding allows AI models to modulate tension and resolution, generating more satisfying story experiences. Likewise, Propp's morphology of the folktale, which identifies recurring character archetypes (e.g., the hero, the villain, the donor) and narrative functions (e.g., departure, struggle, return), helps models maintain culturally familiar and emotionally resonant storylines.

By blending these classical models with modern generative techniques, AI storytelling systems can strike a balance between creativity and structure. This is particularly beneficial in domains such as children's literature, where the use of familiar narrative patterns aids comprehension and emotional engagement. Incorporating age-appropriate themes, character archetypes, and moral dilemmas also helps align AI-generated content with educational and developmental goals.

3. Personalized and Adaptive Storytelling for Education

One of the most promising applications of GenAI storytelling lies in personalized education. Adaptive storytelling tools, such as TinyTeller AI [8] and BookBot, tailor narratives to individual users based on variables like reading level, language proficiency, cognitive development stage, and thematic preferences. These systems dynamically adjust vocabulary complexity, sentence structure, and thematic elements to meet the unique needs of learners, supporting inclusive and differentiated instruction models.

Such personalization enhances both engagement and educational efficacy. For early readers, for example, stories can be simplified with repetitive sentence patterns, phonetic emphasis, and visual

cues. For older students, narratives can incorporate cross-curricular elements—such as science concepts, historical events, or ethical dilemmas—thereby facilitating interdisciplinary learning.

Adaptive storytelling has also been explored in the context of special education. AI systems can generate social stories for children with autism spectrum disorders (ASD) [11], helping them navigate social norms and emotional regulation. By modeling social interactions in narrative form, these systems provide contextualized learning that is both accessible and emotionally grounded.

The modular architecture of modern AI storytelling platforms supports this flexibility. Story engines can plug into learning management systems (LMSs), track user progress, and adapt content over time. This enables long-term personalization and facilitates assessment integration, making GenAI storytelling an increasingly viable tool for scalable, individualized education.

4. Multimodal Storytelling: Text, Audio, and Visuals

The convergence of GenAI with multimodal AI systems has significantly enriched narrative experiences by incorporating not just text, but also speech, imagery, animation, and even interactive environments. This fusion facilitates immersive storytelling that engages multiple sensory modalities, thereby increasing emotional impact and cognitive engagement.

Advancements in Text-to-Speech (TTS) technologies, such as Microsoft's Custom Neural Voice or ElevenLabs, allow for expressive voice synthesis that reflects character emotion and narrative tone. This is especially effective in audiobooks and educational applications, where emotional nuance supports comprehension and retention.[20]

Simultaneously, diffusion-based text-to-image generation models—such as DALL·E, Midjourney, and Stable Diffusion—have empowered creators to produce high-fidelity illustrations based on narrative input. These visuals are not only aesthetically compelling but also contextually aligned with story content, reinforcing reader understanding and immersion. In interactive storytelling applications like AIsop, generated images are dynamically synchronized with the narrative text and voiceover, forming a cohesive, audiovisual storytelling experience.

Platforms integrating these technologies in Virtual Reality (VR) or Augmented Reality (AR) settings are pushing the boundaries further [4]. Systems like AIsopVR embed users within AI-generated story worlds, enabling them to interact with characters and influence plot developments in real-time. User studies indicate that such multimodal narratives enhance empathy, comprehension, and memory retention—especially in educational settings or therapeutic applications.

Despite their promise, multimodal storytelling systems face technical challenges in ensuring temporal and semantic coherence across different modalities. Aligning generated visuals with plot pacing, maintaining voice consistency across characters, and dynamically adapting multimedia content to user input require advanced synchronization and narrative planning mechanisms.

5. Narrative-Based World Generation and Procedural Content

Generative AI is also revolutionizing procedural content generation (PCG) in gaming and interactive storytelling environments by embedding narrative logic into the creation of virtual worlds. Rather than populating game environments with arbitrary or repetitive assets, systems like StoryViz translate textual stories into spatial representations—e.g., generating Minecraft-style settlements that reflect the thematic elements and emotional tone of a story.

This narrative-grounded PCG is achieved using optimization techniques such as swarm intelligence or genetic algorithms, which balance aesthetic and functional aspects of world-building. For example, a story set in a post-apocalyptic wasteland might algorithmically generate barren landscapes, dilapidated structures, and survivalist NPC behaviour, enhancing narrative immersion.

Further innovations involve multi-agent creative systems, where individual AI agents handle specific narrative or environmental elements—such as terrain design, cultural elements, character arcs, and story progression. These agents interact to co-construct evolving narrative environments that respond to player actions and adapt in real-time, supporting emergent storytelling. This model

is particularly compelling for educational simulations and serious games, where the environment must evolve based on learner decisions or instructional objectives.

6. Educational Applications and Narrative-Based Learning

Beyond traditional storytelling, GenAI is increasingly used to enhance pedagogy through narrative-driven learning. By embedding instructional content within compelling story arcs, AI can improve motivation, comprehension, and knowledge retention. For example, systems like ConceptualTales craft stories that explain STEM concepts using analogies and characters from popular culture, such as superheroes or fantasy universes. A chemistry lesson might be framed as a magic duel between wizards representing different elements, while a lesson on gravity could involve a superhero navigating planetary forces.[20]

Cognitive psychology supports this approach: narrative engagement activates brain regions related to memory, empathy, and decision-making, making abstract concepts more relatable and easier to understand. Furthermore, narrative learning encourages higher-order thinking by prompting students to reflect on ethical, social, or scientific dilemmas within the story world.

This approach also supports collaborative and project-based learning. Students may co-author stories with GenAI, engage in role-playing scenarios, or build interactive worlds using story-driven game design tools. These activities cultivate creativity, critical thinking, and digital literacy—core competencies in 21st-century education.

Discussion

The literature on automated story generation reveals a rapidly evolving field characterized by increasing sophistication in both model architecture and application scope. From rule-based systems to transformer-based LLMs, the trajectory illustrates a persistent pursuit of narrative coherence, user control, and creative authenticity. Techniques such as fine-tuning, event abstraction, and conditional generation have significantly expanded the functional capabilities of generative models.

At the same time, the field continues to grapple with challenges related to long-range coherence, style transferability, ethical responsibility, and interpretability. There is also growing recognition of the need to balance model complexity with transparency and resource efficiency—especially for educational and interactive applications.

Future research should explore multimodal integration (e.g., combining text and visual storytelling), longitudinal studies on human-AI co-creation, and more robust frameworks for culturally inclusive and ethically sound narrative generation. Additionally, cross-disciplinary collaboration—spanning computer science, literary theory, ethics, and education—will be essential in shaping the future of AI-generated storytelling.

Despite the remarkable progress in generative storytelling, significant challenges remain. One key issue is maintaining long-term narrative coherence, especially in adaptive or interactive stories that span multiple sessions or respond dynamically to user input. Even the most advanced LLMs can introduce inconsistencies in plot, character behaviour, or setting continuity over time.

Multimodal synchronization also presents complex technical hurdles. Seamlessly aligning text, speech, imagery, and user interaction requires temporally aware generation models and robust narrative planning architectures. Tools for real-time co-creation with users—such as interactive fiction engines or game narrative editors—must evolve to provide both creative flexibility and structural guidance.

Ethical considerations are paramount, particularly in applications targeting children. Mitigating bias, ensuring cultural inclusivity, and maintaining content appropriateness are ongoing concerns. Transparent content moderation pipelines and explainable AI frameworks are necessary to ensure accountability. Moreover, involving educators, psychologists, and ethicists in the design of storytelling systems is crucial to ensure developmental alignment and social responsibility.

Looking ahead, research is focused on enhancing memory-augmented generation, integrating user emotion and intent into adaptive story engines, and developing collaborative AI that treats users

as co-authors rather than passive consumers. Ultimately, the goal is to create generative storytelling systems that are not only capable of producing entertaining and educational narratives but also empower users to explore, create, and inhabit rich, meaningful narrative worlds.

Conclusion

The integration of generative AI-based script writing with high-quality story visualization marks a transformative step in the evolution of automated narrative creation. By bridging the divide between language and vision models, this research presents a cohesive framework that ensures semantic alignment, temporal consistency, and narrative coherence across modalities. The proposed approach moves beyond isolated generative outputs, enabling rich, immersive storytelling experiences that combine textual depth with visual expressiveness.

As demonstrated in this work, the convergence of large language models and advanced visual synthesis tools can facilitate the automated generation of narratives that are not only engaging but also responsive to audience needs and genre expectations. Our architecture, built upon semantic abstraction and cross-modal alignment, represents a scalable and adaptable foundation for future developments in AI storytelling. The experimental validations underscore its potential across diverse domains, including entertainment, education, digital marketing, and interactive fiction.

However, this convergence also brings forth new challenges—ranging from maintaining long-range coherence to addressing ethical concerns and cultural biases. These issues underscore the need for continued innovation in model interpretability, user co-creation interfaces, and socially responsible AI development. Multidisciplinary collaboration will be essential to refine these systems for real-world deployment, ensuring that AI-generated narratives remain inclusive, meaningful, and safe.

In conclusion, this work lays the groundwork for next-generation storytelling systems that harness the strengths of generative AI to automate, augment, and personalize narrative experiences. As technologies mature, the vision of AI as a creative collaborator—empowering users to tell stories, visualize dreams, and engage with content in unprecedented ways—is becoming an achievable reality.

References

1. M. P, P. Velvizhy and P. S. Sherly, "Optimized Story Generation using DeepSeek LLM with Supervised Fine-Tuning," 2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India, 2025, pp. 1-5, doi: 10.1109/RMKMATE64874.2025.11042725.

2. S. Bai, D. E. Gonda and K. F. Hew, "Write-Curate-Verify: A Case Study of Leveraging Generative AI for Scenario Writing in Scenario-Based Learning," in IEEE Transactions on Learning Technologies, vol. 17, pp. 1301-1312, 2024, doi: 10.1109/TLT.2024.3378306.

3. Kok, C.L.; Ho, C.K.; Aung, T.H.; Koh, Y.Y.; Teo, T.H. Transfer Learning and Deep Neural Networks for Robust Intersubject Hand Movement Detection from EEG Signals. Appl. Sci. 2024, 14, 8091. https://doi.org/10.3390/app14178091.

4. E. Gatti, D. Giunchi, N. Numan and A. Steed, "Around the Virtual Campfire: Early UX Insights into AI-Generated Stories in VR," *2025 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*, Lisbon, Portugal, 2025, pp. 136-141, doi: 10.1109/AIxVR63409.2025.00027.

5. J. Demke, R. Morain, C. Wilhelm and D. Ventura, "Multi-agent Story-based Settlement Generation," *2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, Washington, DC, USA, 2021, pp. 1149-1153, doi: 10.1109/ICTAI52525.2021.00182.

6. C. Salge, C. Guckelsberger, M. C. Green, R. Canaan, and J. Togelius, "Generative design in Minecraft: Chronicle challenge," in Proceedings of the Tenth International Conference on Computational Creativity, 2019, pp. 311–315.

7. H. Zhang, C. Fu, H. Zhao, and B. Wu, "Adapting Large Language Models for Content Moderation: Pitfalls in Data Engineering and Supervised Fine-tuning," arXiv, abs/2310.03400, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2310.03400.

8. M. Kim, T. Kim, A. Nguyen, E. N. Gomez and J. Jin, "TinyTeller AI, an AI-based Adaptive Storytelling Application," *2024 Artificial Intelligence x Humanities, Education, and Art (AIxHEART)*, Laguna Hills, CA, USA, 2024, pp. 44-51, doi: 10.1109/AIxHeart62327.2024.00017.

9. M. Garber-Barron and M. Si., "Towards Interest And Engagement, A Framework for Adaptive Storytelling " Intelligent Narrative Technologies: 2012 AIIDE Workshop Technical Report, vol. 8 No. 2.

10. Han, A., Cai, Z. (2023). Design implications of generative AI systems for visual storytelling for young learners. IDC '23: Proceedings of the 22nd Annual ACM Interaction Design and Children Conference.

11. D. Baradari, H. Han, J. Xia and C. A. Strelecki, "Beyond Imagination: Leveraging Generative AI to Enhance Learning Through Story World Analogies," *2024 IEEE Frontiers in Education Conference (FIE)*, Washington, DC, USA, 2024, pp. 1-8, doi: 10.1109/FIE61694.2024.10892813.

12. J. Radesky, H. M. Weeks, A. Schaller, M. Robb, S. Mann, and A. Lenhart, "Constant Companion: A Week in the Life of a Young Person's Smartphone Use," *Common Sense*, 2023.

13. R. M. Oikarinen, J. K. Oikarinen, S. Havu-Nuutinen, and S. Pöntinen, "Students' collaboration in technology-enhanced reciprocal peer tutoring as an approach towards learning mathematics," *Education and Information Technologies*, 2022. https://doi.org/10.1007/s10639–021-10799-3.

14. M. Wang, "Artificial intelligence narration in virtual reality," in *2023 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. IEEE, 2023, pp. 376–380.

15. H. Shakeri, C. Neustaedter, and S. DiPaola, "Saga: Collaborative story-telling with gpt-3," in *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing*, 2021, pp. 163–166.

16. Kok, C.L.; Tan, T.C.; Koh, Y.Y.; Lee, T.K.; Chai, J.P. Design and Testing of an Intramedullary Nail Implant Enhanced with Active Feedback and Wireless Connectivity for Precise Limb Lengthening. *Electronics* 2024, *13*, 1519. doi: 10.3390/electronics13081519

17. "OpenAI GPT-4 Turbo ", https://help.openai.com/en/articles/8555510-gpt-4-turbo-in-the-openai-api, 2024. OpenAI.

18. S. Jo, Z. Yuan and S. -W. Kim, "Interactive Storyboarding for Rapid Visual Story Generation," *2022 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, Yeosu, Korea, Republic of, 2022, pp. 1-4, doi: 10.1109/ICCE-Asia57006.2022.9954679.

19. Li, Mengtian, "Photo-sketching: Inferring contour drawings from images." *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2019.

20. M. Mohammed Abass Hamza, "Artificial Intelligence and Literary Creativity: Examining New Frontiers in English Literature," *2024 International Conference on IoT, Communication and Automation Technology (ICICAT)*, Gorakhpur, India, 2024, pp. 773-778, doi: 10.1109/ICICAT62666.2024.10923368.

21. Isola, Phillip, "Image-to-image translation with conditional adversarial networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

22. B. Patil, G. B. Yadav, A. Buchade, S. Borkar, S. Bhosale and V. Honbute, "Dynamic StarCraft: Multi-Agent Generative AI for Immersive Experiences," *2025 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, 2025, pp. 1-7, doi: 10.1109/ESCI63694.2025.10987974.

23. Hiroshi Tanaka, Mei Li, "Augmented Reality, TTS, and Motion Capture for Text-Driven Animation ", International Conference on Augmented Reality and Interactive Storytelling. 2022

24. Alice Wang, Gregory Stevens, "Text-to-Speech and Animation in Educational Storytelling ", *Journal of Educational Technology & Animation.* 2022

25. H. Ahuja, D. Gupta, M. Kumar and J. Chugh, "Exploring AI Ability to Generate Artistic Content, Music Literature, and Other Creative Works," *2024 International Conference on Integrated Circuits, Communication, and Computing Systems (ICIC3S)*, Una, India, 2024, pp. 1-6, doi: 10.1109/ICIC3S61846.2024.10602992.

26. L. P. Khan, V. Gupta, S. Bedi and A. Singhal, "StoryGenAI: An Automatic Genre-Keyword Based Story Generation," *2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*, Greater Noida, India, 2023, pp. 955-960, doi: 10.1109/CISES58720.2023.10183482.

27. J. Wu, W. Gan, Z. Chen, S. Wan, and H. Lin, "Ai-generated content (aigc): A survey," *arXiv preprint* arXiv: 2304.06632, 2023.

28. M. Shidiq, "The use of artificial intelligence-based chat-gpt and its challenges for the world of education; from the viewpoint of the development of creative writing skills," *Proceeding of International Conference on Education, Society and Humanity*, vol. 1, no. 1, pp. 353–357, 2023.

29. W. Du and Q. Han, "Research on application of artificial intelligence in movie industry," in 2021 International Conference on Image, Video Processing, and Artificial Intelligence, vol. 12076. SPIE, 2021, pp. 265–270.

30. A. Pradeep, A. Satmuratov, I. Yeshbayev, O. Khasan, M. Iqboljon and A. Daniyor, "The Significance of Artificial Intelligence in Contemporary Cinema," *2023 Second International Conference on Trends in Electrical, Electronics, and Computer Engineering (TEECCON)*, Bangalore, India, 2023, pp. 111-116, doi: 10.1109/TEECCON59234.2023.10335867.

31. J. Kong, L. Siek and C. L. Kok, "A 9-bit body-biased vernier ring time-to-digital converter in 65 nm CMOS technology," 2015 IEEE International Symposium on Circuits and Systems (ISCAS), Lisbon, Portugal, 2015, pp. 1650-1653, doi: 10.1109/ISCAS.2015.7168967.

32. M. M. Chan, H. R. Amado-Salvatierra, R. Hernandez-Rizzardini and M. Rosales, "Assessing Student Perceptions of Video Quality and Effectiveness From AI-Enhanced Digital Resources," *2024 IEEE Digital Education and MOOCS Conference (DEMOcon)*, Atlanta, GA, USA, 2024, pp. 1-5, doi: 10.1109/DEMOcon63027.2024.10747930.

33. R. Selvaraj, A. Singh, S. Kameel, R. Samal and P. Agarwal, "Vidgen: Long-Form Text-to-Video Generation with Temporal, Narrative and Visual Consistency for High Quality Story-Visualisation Tasks," *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, Pune, India, 2024, pp. 1-8, doi: 10.1109/I2CT61223.2024.10544050.

34. P. Ammanabrolu, E. Tien, W. Cheung, Z. Luo, W. Ma, L. Martin, M. Riedl, "Story Realization: Expanding Plot Events into Sentences," 2019. [Online]. Available: arxiv.org/abs/1909.03480.

35. H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, S. Bhosale et al., "Llama 2: Open foundation and fine-tuned chat models," arXiv preprint arXiv:2307.09288, 2023.

36. Y. Li, M. Min, D. Shen, D. Carlson, and L. Carin, "Video generation from text," in Proceedings of the AAAI conference on artificial intelligence, vol. 32, no. 1, 2018.

37. E. Gatti, D. Giunchi, N. Numan and A. Steed, "AIsop: Exploring Immersive VR Storytelling Leveraging Generative AI," *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, Orlando, FL, USA, 2024, pp. 865-866, doi: 10.1109/VRW62533.2024.00229.

38. Kok, C.L. and Siek, L. (2015), Unbalanced input pair zero current detector for DC–DC buck converter. Electron. Lett., 51: 1359-1361. https://doi.org/10.1049/el.2015.0323

39. T. Rahman, H.-Y. Lee, J. Ren, S. Tulyakov, S. Mahajan, and L. Sigal, "Make-a-story: Visual memory conditioned consistent story generation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 2493–2502.

40. F. Brahman and S. Chaturvedi. Modeling protagonist emotions for emotion-aware storytelling. *arXiv preprint* arXiv: 2010.06822, 2020. 1.

41. Y. Li, Z. Gan, Y. Shen, J. Liu, Y. Cheng, Y. Wu, L. Carin, D. Carlson, and J. Gao, "Storygan: A sequential conditional gan for story visualization," 2019.

42. H. Shakeri, C. Neustaedter, and S. DiPaola. Saga: Collaborative story-telling with gpt-3. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing*, pp. 163–166, 2021. 1.

43. J. Shen, C. Fu, X. Deng and F. Ino, "A Study on Training Story Generation Models Based on Event Representations," *2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, Chengdu, China, 2020, pp. 210-214, doi: 10.1109/ICAIBD49809.2020.9137439.

44. H. Kim, J. -H. Choi and J. -Y. Choi, "A Novel Scheme for Managing Multiple Context Transitions While Ensuring Consistency in Text-to-Image Generative Artificial Intelligence," in *IEEE Access*, vol. 12, pp. 150468-150484, 2024, doi: 10.1109/ACCESS.2024.3476933.

45. H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, and D. Metaxas, "StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5908–5916.

46. M. P, P. Velvizhy and P. S. Sherly, "Optimized Story Generation using DeepSeek LLM with Supervised Fine-Tuning," *2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*, Chennai, India, 2025, pp. 1-5, doi: 10.1109/RMKMATE64874.2025.11042725.

47. R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 10674–10685.

48. S. Bai, D. E. Gonda and K. F. Hew, "Write-Curate-Verify: A Case Study of Leveraging Generative AI for Scenario Writing in Scenario-Based Learning," in *IEEE Transactions on Learning Technologies*, vol. 17, pp. 1301-1312, 2024, doi: 10.1109/TLT.2024.3378306.

49. A. S. Rao *et al*., "Moral Storytelling Model Using Artificial Intelligence-Driven Image-to-Text Synthesis," *2024 International Conference on Data Science and Network Security (ICDSNS)*, Tiptur, India, 2024, pp. 01-07, doi: 10.1109/ICDSNS62112.2024.10691066.

50. T. Brooks, J. Hellsten, M. Aittala, T.-C. Wang, T. Aila, J. Lehtinen, M.-Y. Liu, A. Efros, and T. Karras, "Generating long videos of dynamic scenes," *Advances in Neural Information Processing Systems*, vol. 35, pp. 31769–31781, 2022.

51. Kok, C.L.; Heng, J.B.; Koh, Y.Y.; Teo, T.H. Energy-, Cost-, and Resource-Efficient IoT Hazard Detection System with Adaptive Monitoring. Sensors 2025, 25, 1761. https://doi.org/10.3390/s25061761

52. L. Yao, N. Peng, R. Weischedel, K. Knight, D. Zhao, R. Yan, "Planand- write: Towards better automatic storytelling," In AAAI '19 Proc. of the 33rd AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, Feb. 2019, pp. 7378–7385.

53. M. F. Mridha, A. A. Lima, K. Nur, S. C. Das, M. Hasan, and M. M. Kabir, "A survey of automatic text summarization: Progress, process and challenges," *IEEE Access*, vol. 9, pp. 156043–156070, 2021.

54. X. Li et al., "A novel voltage reference with an improved folded cascode current mirror OpAmp dedicated for energy harvesting application," 2013 International SoC Design Conference (ISOCC), Busan, Korea (South), 2013, pp. 318-321, doi: 10.1109/ISOCC.2013.6864038.

55. R. Maggio, "The anthropology of storytelling and the storytelling of anthropology," *Journal of comparative Research in Anthropology and Sociology*, vol. 5, no. 02, pp. 89–106, 2014.

56. Yao, Lili, et al. "Plan-and-write: Towards better automatic storytelling." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. No. 01. 2019.