MITHRIL: A Multi-modal and Immersive Generative AI Approach for Roleplaying Tabletop Games

[1]Michael Atkins,* [1]William Denman,* [1]Tuna Han Salih Meral, [2]Doga Dogan, [1]Pinar Yanardag
[1]Virginia Tech
[2]Adobe Research
{pinary}@vt.edu

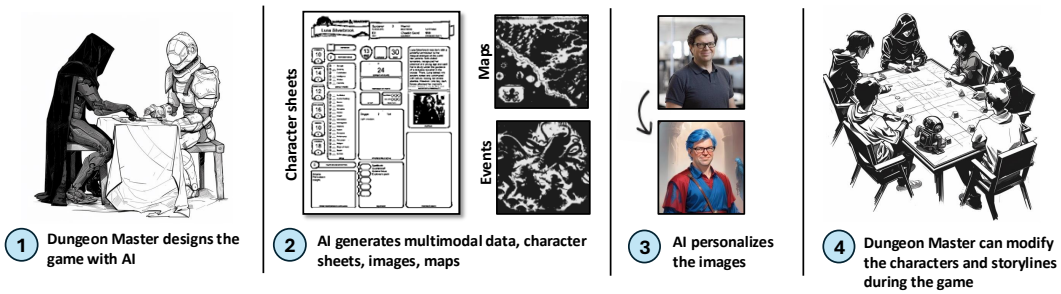

Figure 1: Dungeon Masters use MITHRIL to craft characters and campaign elements. MITHRIL's multi-modal capabilities enable the generation of detailed images and backstories for characters, outlining their skills and attributes. Additionally, MITHRIL offers personalization options, allowing characters to be visually customized to resemble the players themselves.

Abstract

Tabletop Role-Playing Games (TTRPGs) like Dungeons & Dragons (D&D) offer rich, collaborative storytelling experiences where Dungeon Masters (DMs) play a crucial role in creating narrative arcs, designing encounters, and bringing the game world to life. However, the demands of crafting engaging stories, developing complex characters, and managing the intricate details of gameplay can be overwhelming for DMs, often requiring significant effort and time. This paper introduces a novel system that assists DMs in generating multimodal content, including both text descriptions and images for characters, campaigns, and story arcs, with advanced personalization features. Leveraging state-of-the-art AI capabilities, the system enables DMs to automatically generate detailed character backstories, skill sets, and high-quality images, creating a more immersive and visually enriched narrative experience. Unlike existing tools that focus solely on text generation, our system integrates both textual and visual modalities and allows for a high degree of customization, such as embedding players' faces onto character avatars. This unique approach addresses the growing demand for personalized and dynamic content in TTRPGs, reducing the DM's workload while enhancing creativity and player engagement. We detail the design and implementation of our system, evaluate its usability and effectiveness, and discuss how it sets a new standard for multimodal content generation and personalized storytelling in TTRPGs.

*Equal contribution.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: The First Workshop on Generative and Protective AI for Content Creation.

(a) A game session with 8 players

(b) Personalized characters of the players

(c) Character Sheet

Figure 2: We conducted several game sessions (a) to evaluate the effectiveness of the proposed system. Participants received personalized character sheets (c) designed to resemble their likeness (b). These character images were also projected onto a screen to facilitate easy viewing among participants. Additionally, an interactive map, created using our tool, was also shared with the participants, enabling them to navigate their characters on the map using their laptops.

1 Introduction

Tabletop Role-Playing Games, such as Dungeons & Dragons (7), have become increasingly popular as collaborative storytelling experiences where players create characters and navigate through fantastical worlds led by a Dungeon Master (DM). The DM is responsible for crafting narrative arcs, designing encounters, and bringing the game world to life by portraying non-player characters (NPCs) and orchestrating the game's progression. This role is highly demanding, requiring a delicate balance between creativity and rule-based decision-making, as DMs must manage not only the overarching story but also fine-grained details of character interactions, environment descriptions, and plot developments. While DMs are the cornerstone of the D&D experience, they often face challenges in creating rich, immersive content that caters to the specific interests and styles of their player groups. This paper introduces a novel system designed to assist DMs in generating multimodal content, including text descriptions and images for characters, campaigns, and story arcs, while also providing personalization features.

In recent years, generative AI has gained significant attention for its ability to create text, images, audio, and even video content based on user input or existing data (11; 19; 8; 12; 13). In the context of TTRPGs, where storytelling, creativity, and immersion are central, generative AI presents a unique opportunity. By leveraging the power of AI to automate and enrich the content generation process, DMs can overcome many of the traditional challenges associated with campaign preparation. Generative AI can assist in developing detailed backstories, creating unique character portraits,

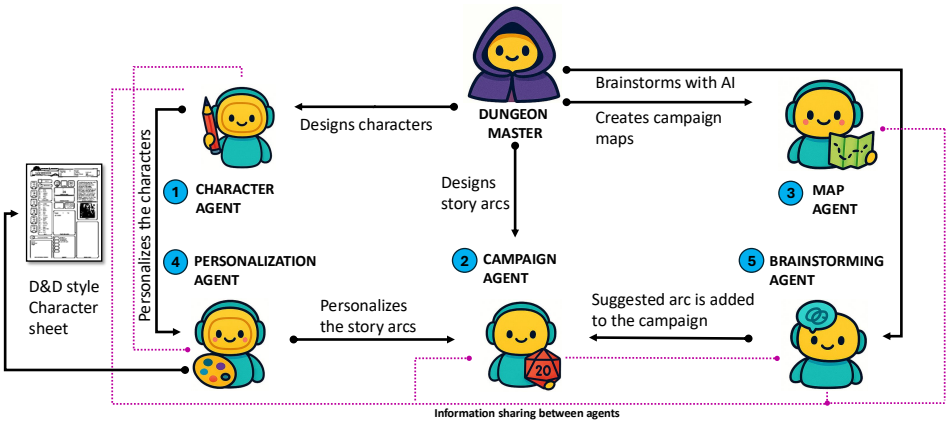

Figure 3: Our framework consists of four interconnected agents: (1) Character Agent (2) Campaign Agent (3) Map Agent (4) Personalization Agent (5) Brainstorming Agent.

generating dynamic descriptions of environments and encounters, and even adapting the narrative on the fly based on DM's actions. This capability not only saves time but also allows DMs to focus more on the interactive and improvisational aspects of game mastering, enhancing the overall player experience. Our proposed system generates detailed character backstories, skill sets (such as proficiencies and abilities), and high-quality images that visualize both characters and story arcs. Unlike existing systems that primarily focus on text-based content generation (14; 20), our approach integrates both textual and visual modalities, allowing DMs to create a richer, more immersive narrative experience. For example, our system can generate an image depicting characters entering *an enchanted forest and facing a giant spider*, simultaneously providing text descriptions of the scene, NPC dialogues, maps, and character sheets (see Fig. 1). Furthermore, our system offers unique personalization features, such as embedding players' faces onto generated character avatars, enhancing engagement and immersion. This multimodal, personalizable approach addresses a significant gap in current TTRPG tools by providing DMs with a platform for campaign and character design.

2 Methodology

Our framework (see Fig. 3) consists of four interconnected agents: (1) Character Agent generates characters, NPCs, and enemies based on user input, providing detailed profiles and attributes for each; it also interacts with the Personalization Agent to create customized character images. (2) Campaign Agent develops story arcs and narrative content based on user prompts and integrates character details provided by the Character Agent to maintain narrative consistency. (3) Map Agent generates detailed images of campaign locations. (4) Personalization Agent personalizes character images and event scenes by incorporating player-uploaded photos and text descriptions, working closely with the Character Agent to ensure visual consistency with the generated characters. (5) Brainstorming Agent collaborates with both the Campaign Agent and Character Agent to propose alternative story ideas in friendly, neutral, and antagonistic tones, ensuring diverse narrative paths while maintaining alignment with established characters and storylines. Together, these agents form a cohesive system that enables dynamic and immersive storytelling for Dungeon Masters and players in TTRPGs.

3 Conclusion

We believe that our approach sets a new standard for the integration of multimodal content generation and personalization in TTRPGs, paving the way for more dynamic and engaging storytelling frameworks. Our demo is accessible to public by developing a web-based software prototype with an intuitive interface that does not require any specialized hardware or technical skills: `https://dungeons-dragons-brown.vercel.app`.

References

[1] firebase. https://firebase.google.com/docs/firestore, accessed: 2024-09-01

[2] nextjs. https://nextjs.org/, accessed: 2024-09-01

[3] vercel. https://vercel.com/, accessed: 2024-09-01

[4] Ang, C., Cortel, L.R., Santos, C.L., Ong, E.: Fable reborn: Investigating gameplay experience between a human player and a virtual dungeon master. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–7 (2023)

[5] Callison-Burch, C., Tomar, G.S., Martin, L.J., Ippolito, D., Bailis, S., Reitter, D.: Dungeons and dragons as a dialog challenge for artificial intelligence. arXiv preprint arXiv:2210.07109 (2022)

[6] Friedrich, F., Brack, M., Struppek, L., Hintersdorf, D., Schramowski, P., Luccioni, S., Kersting, K.: Fair diffusion: Instructing text-to-image generation models on fairness. arXiv preprint arXiv:2302.10893 (2023)

[7] Gygax, G., Arneson, D.: Dungeons & dragons, vol. 19. Tactical Studies Rules Lake Geneva, WI (1974)

[8] Ho, J., Chan, W., Saharia, C., Whang, J., Gao, R., Gritsenko, A., Kingma, D.P., Poole, B., Norouzi, M., Fleet, D.J., et al.: Imagen video: High definition video generation with diffusion models. arXiv preprint arXiv:2210.02303 (2022)

[9] IJsselsteijn, W.A., De Kort, Y.A., Poels, K.: The game experience questionnaire (2013)

[10] Kelly, J., Mateas, M., Wardrip-Fruin, N.: Towards computational support with language models for ttrpg game masters. In: Proceedings of the 18th International Conference on the Foundations of Digital Games. pp. 1–4 (2023)

[11] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)

[12] Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., Aberman, K.: Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 22500–22510 (2023)

[13] Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E.L., Ghasemipour, K., Gontijo Lopes, R., Karagol Ayan, B., Salimans, T., et al.: Photorealistic text-to-image diffusion models with deep language understanding. Advances in neural information processing systems **35**, 36479–36494 (2022)

[14] Sturdee, M., Gamboa, M., Heron, M.: Ttrpg ux: Requirements & beyond. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–9 (2023)

[15] Triyason, T.: Exploring the potential of chatgpt as a dungeon master in dungeons & dragons tabletop game. In: Proceedings of the 13th International Conference on Advances in Information Technology. pp. 1–6 (2023)

[16] Wang, Q., Bai, X., Wang, H., Qin, Z., Chen, A., Li, H., Tang, X., Hu, Y.: Instantid: Zero-shot identity-preserving generation in seconds. arXiv preprint arXiv:2401.07519 (2024)

[17] Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect: the panas scales. Journal of personality and social psychology **54**(6), 1063 (1988)

[18] You, X., Taveekitworachai, P., Chen, S., Can Gursesli, M., Li, X., Xia, Y., Thawonmas, R.: Dungeons, dragons, and emotions: A preliminary study of player sentiment in llm-driven ttrpgs. In: Proceedings of the 19th International Conference on the Foundations of Digital Games. pp. 1–4 (2024)

[19] Zhang, C., Zhang, C., Zhang, M., Kweon, I.S.: Text-to-image diffusion models in generative ai: A survey. arXiv preprint arXiv:2303.07909 (2023)

[20] Zhu, A., Martin, L., Head, A., Callison-Burch, C.: Calypso: Llms as dungeon master's assistants. In: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment. vol. 19, pp. 380–390 (2023)