Dungeon Automation: Rolling the Dice for DMs

Gopal Nambiar
University of California
Davis, USA
gnambiar@ucdavis.edu

Mahek Jain
University of California
Davis, USA
mhkjain@ucdavis.edu

Ruthuvikas Ravikumar
University of California
Davis, USA
rvravikumar@ucdavis.edu

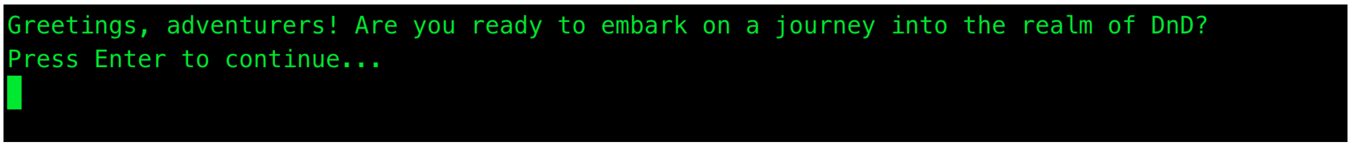

Figure 1. Welcome prompt players are greeted with

Abstract

Automated systems have been explored in various domains, but their use in Dungeons and Dragons (D&D) campaigns is still emerging. This research paper investigates the effectiveness of three main modules: the automated dungeon master (ADM); the Summarizer, the module that summarizes the generated campaigns; and Viz, the module that generates visualizations of key plot points.

ADM attempts to answer the research question of whether the current state-of-the-art LLM can run a D&D campaign. The Summarizer evaluates the accuracy of summarizing a D&D campaign, while Viz generates visualizations of key plot points to provide players with tangible artwork.

This paper presents the results of our experiments that aim to address the research questions of how accurately and effectively these modules can run, summarize, and generate visualizations of a D&D campaign. Our experiments show that ADM can effectively run a D&D campaign, Summarizer can accurately summarize a campaign, and Viz can generate appropriate visualizations.

This research can potentially revolutionize how D&D campaigns are run, making them more accessible and engaging for players. It can also have implications for other domains where automated systems can assist human decision-making.

CCS Concepts: • **Computing methodologies** → *Discourse, dialogue and pragmatics*; *Natural language generation.*

Keywords: AI, Dungeons and Dragons, Text summarization, Deep learning

ACM Reference Format:
Gopal Nambiar, Mahek Jain, and Ruthuvikas Ravikumar. 2024. Dungeon Automation: Rolling the Dice for DMs. In *Proceedings of ECS289G (Project Report)*. UC Davis, Davis, CA, USA, 8 pages.

1 Introduction

Dungeons and Dragons (D&D) is a popular tabletop role-playing game that has been played for over 45 years. The game involves creating characters and going on adventures, where the dungeon master (DM) guides the players through a world of their creation. In recent years, D&D has seen a surge in popularity, with online platforms such as Roll20 and D&D Beyond making it easier to play remotely. This growing popularity has increased demand for DMs, which can be challenging for those new to the game.

While DMing can be a fun and rewarding experience, it is also a complex task that requires creativity, quick thinking, and the ability to adapt to unpredictable situations. Automating this process is an intriguing prospect, as it could potentially reduce the workload on the DM and make the game more accessible to newcomers.

However, creating an AI that can effectively run a D&D campaign is daunting. D&D is a game that relies heavily on storytelling, improvisation, and the ability to adapt to unforeseen circumstances. These are all tasks that are notoriously difficult for AI systems to accomplish. Furthermore, D&D campaigns can be highly subjective, with different DMs and players having different interpretations of the game's rules and lore.

In this paper, we explore the use of automated systems in running and summarizing D&D campaigns. Specifically, we present three main modules: ADM, the automated dungeon master; Summarizer, the module that summarizes the generated campaigns; and Viz, the module that generates visualizations of key plot points. We aim to investigate the effectiveness of these modules and their potential to revolutionize the way D&D campaigns are run.

Project Report, Mar 2023, Davis, CA, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00

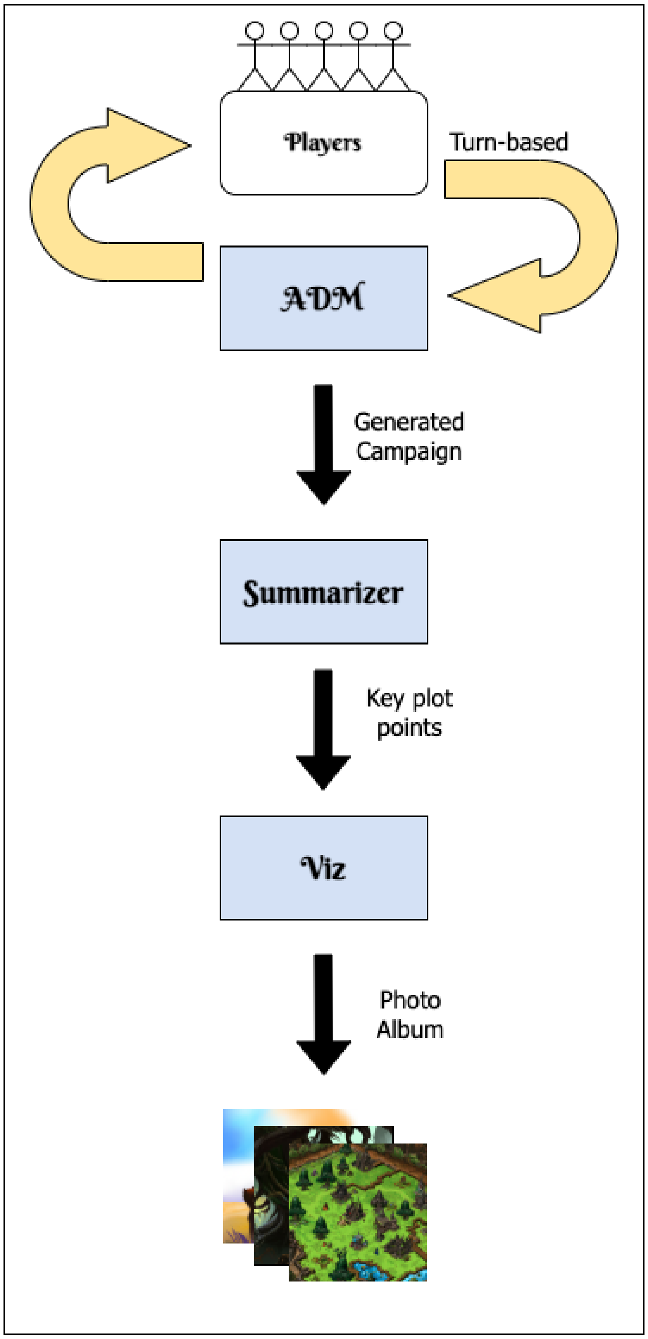

Figure 2. Application Architecture Diagram

2 Related Work

In recent years, automated game design has been a popular topic of interest among researchers. One of the seminal works in this field was introduced by [6], who proposed the concept of automated game design. This work inspired the current study to explore the use of procedural content generation in Dungeons and Dragons campaigns.

Several related works have also explored procedural content generation in games. For instance, [2] presented the use of procedural content generation in games, which was relevant to the current study's exploration of this concept. Moreover, [12] proposed the procedural generation of dungeons, which was relevant to the current paper's focus on using procedural content generation to create game levels.

To contextualize the current study, [9] provided an overview of procedural content generation in games and discussed its potential applications. Additionally, the book edited by [10] was consulted to gain a comprehensive understanding of the topic and its related concepts.

Automated methods for game design have also been explored by researchers. For instance, [3] presented a method for an automated game design using conceptual expansion, which like the current study, explores automated methods for game design.

Finally, [7] discussed a case study in teaching software engineering for computer games courses using a code play framework, which was relevant to the current study's focus on game development and software engineering. More recently, [5], [14], and [11] have also explored the use of AI and machine learning in game design, demonstrating the potential for the current paper's results to have implications beyond the realm of D&D campaigns.

3 Methodology

As noted in the introduction, the application architecture consists of 3 key components. In this section, we will understand the methodology using which the components function and interact with each other.

3.1 ADM

The methodology used in the development of ADM involves building upon the GPT-3 architecture [1]. This cutting-edge technology is specifically designed for natural language processing, which makes it the ideal choice for an AI-driven dungeon master that can effectively communicate with human players. The GPT-3 architecture can produce human-like responses that are contextually appropriate and can easily handle complex queries. The use of this technology ensures that the dungeon master can seamlessly interact with the players and provide an immersive gaming experience.

ADM has several key features that make it a unique and dynamic addition to any Dungeons and Dragons campaign:

- **Character analysis**: ADM analyzes the players' race and class and keeps them in mind throughout the campaign. ADM generates more favorable or challenging situations for the players based on their character traits. For example, let's say that the players have a half-elf rogue and a dwarf cleric in their party. ADM has analyzed the characters' race, class, and history

to understand their strengths, weaknesses, and personality traits. Based on this analysis, ADM decides to create a situation where the players must navigate a political conflict between a human and an elven kingdom. The half-elf rogue's elven heritage makes them a valuable asset in negotiating with the elven kingdom, while the dwarf cleric's distrust of elves makes them more suspicious of the elven kingdom's motives.
As the players delve deeper into the conflict, they discover that the human kingdom is being controlled by a powerful wizard who seeks to exploit the elven kingdom's resources for their own gain. The players must then use their unique abilities and perspectives to rally support against the wizard and prevent a devastating war between the two kingdoms.
In this example, the players' race and class have been used to influence the plot points by creating a conflict that is rooted in their characters' backgrounds and personalities. The half-elf rogue's elven heritage and the dwarf cleric's distrust of elves add depth and complexity to the situation, making it more engaging and immersive for the players.
ADM can also incorporate the players' backstories into the game to create a more immersive experience.

- **Quest and story creation**: ADM generates interesting and unique storylines that keep the players engaged and motivated. ADM improvises scenes based on player feedback to make the scenes dynamic. The quests created are personalized to the players' preferences, character traits, and backstories, creating a unique gaming experience every time.

- **NPC management**: ADM also manages NPCs (non-player characters) throughout the campaign. ADM creates and controls NPC actions and interactions with the players. Each NPC is given unique personalities, traits, and motivations that affect their behavior and decision-making, adding depth and complexity to the campaign.

- **Rule enforcement**: The role of the DM in enforcing the rules is crucial in D&D. The DM is responsible for interpreting the rules and ensuring that the players follow them. This involves making judgments about how specific rules apply in different situations and sometimes making on-the-spot rulings to keep the game moving smoothly.
  In addition to enforcing the rules, the DM can also provide feedback and guidance to the players. This can include correcting rule violations, offering suggestions for how to improve gameplay, and helping players understand how to use the rules to their advantage.
  ADM enforces the game's rules, ensuring that the players adhere to the procedure of the game. ADM also provides real-time feedback to players, correcting rule violations and offering suggestions for how to improve their gameplay.

- **Randomness**: ADM incorporates randomness into the game, introducing unpredictable events that keep the players on their toes. Randomness can affect the outcome of battles, alter the course of the storyline, and add excitement to the game.

- **Data Tracking**: ADM tracks player progress, including character development, quest completion, and overall game performance. This data is used to provide real-time feedback to the players and adjust the game's difficulty level accordingly. Another aspect of the character data that is tracked is the players' HP (Hit points). This is especially crucial during battle scenes when a player needs to be notified if their character's health is at a critical level.

- **Personalization**: ADM can be personalized to cater to the preferences of each player. ADM can also be given different personalities to match the players' preferences. For example, the system can be programmed to adopt a friendly and helpful persona for players who prefer a more supportive gameplay experience. Alternatively, it can adopt a more challenging and competitive persona for players who prefer a more intense and engaging experience. By offering a range of customizable personality options, ADM can provide a truly personalized gaming experience that adapts to the needs and preferences of each individual player. This level of customization ensures that players can fully immerse themselves in the game and enjoy a unique and engaging experience every time they play.

The use of the GPT-3 architecture allows the dungeon master to generate complex narratives that are tailored to the players' preferences and actions. The system is designed to learn from the players' decisions and adapt the story accordingly. This personalized storytelling helps to create a more dynamic and enjoyable gaming experience for the players. Overall, the methodology used in developing the AI-driven dungeon master is geared towards creating a sophisticated and intelligent system that can enhance the gaming experience for players.

3.2 Summarizer

After ADM generates the campaign, the next step is to generate a summary of the campaign. This is a crucial step as it enables the user to quickly understand the key elements of the campaign without having to go through the entire campaign in detail. To perform the summarization task, three different transformer models are used. These models are based on the idea of self-attention mechanisms, which enable the models to focus on different parts of the input sequence during processing. By doing so, the transformer models [13] are able to capture complex relationships between words and generate more accurate and context-aware outputs. This is achieved through multiple layers of self-attention and feedforward neural networks in the transformer models, which enable them to process the input data and generate a summary that accurately captures the key elements of the campaign. The use of transformer models has enabled significant improvements in the performance of summarization tasks and has become a popular technique in the field of natural language processing (NLP).

The three summarization models used in this architecture are:

- **BART-large-CNN-SAMSum**: A state of the art transformer based model for abstractive text summarization. Abstractive summarization involves generating a summary that captures the essence of the original text rather than simply selecting and rephrasing parts of the original text. The model was specifically designed for summarizing news articles. It was trained on a combination of the CNN/Daily Mail dataset and the SAMSum dataset, which consists of social media conversations and summaries. The model is based on the BART architecture, which is a pre-trained transformer-based model for natural language generation tasks. BART-large-CNN-SAMSum is a variant of the BART-large model, which was fine-tuned on the two datasets to improve its performance on summarization tasks.

- **Distilbert-cnn**: A transformer-based model for abstractive text summarization that was introduced in 2020. It is a modification of the original DistilBERT architecture, which is a smaller and faster version of the popular BERT model. The DistilBERT-CNN model is designed to summarize news articles and other types of long-form text. The model combines the DistilBERT architecture with a convolutional neural network (CNN), which is a type of neural network commonly used for image-processing tasks. The CNN is used to process the input text and identify important features and patterns, which are then combined with the learned representation from the DistilBERT model to generate the summary.

- **Google T5**: A transformer-based language model that was introduced in 2019. It is a large-scale, pre-trained model that has been trained on a wide range of natural language processing tasks, including text summarization. The T5 model is unique in that it is designed to perform multiple NLP tasks using a single model architecture. It uses a "text-to-text" approach, where each task is framed as a "translation" problem, where the input text is "translated" into the output text. This enables the model to handle a wide range of tasks, including text summarization.

3.3 Viz

The next step in the process is to use the DALL-E API [8] to create stunning visuals of the D&D campaign. DALL-E is an image-generation AI that can create images from textual descriptions. The concise summaries generated by the summarization models are fed line by line into the DALL-E API to create a series of images that truly describes the events and settings of the campaign. The API works by using a neural network to translate textual input into visual output. The neural network has been trained on a large dataset of images and textual descriptions. The DALL-E API generates images by combining visual elements in response to the input textual descriptions. The images can range from simple depictions of objects to complex scenes with multiple elements. Once the image is generated on running the script, they are stored as JSON responses. These JSON responses contain data about the generated images, such as their size and pixel data. The responses are then decoded to produce the actual images.

The series of images generated by DALL-E are then displayed on a webpage using the Flask framework, which is a popular Python web framework for building web applications. Flask provides a simple and efficient way to create web applications that can interact with other Python libraries and APIs. The images generated by the DALL-E API are used to create a photo album of the D&D campaign. The photo album serves as a record of the campaign and allows the players to revisit the events and settings of the game in a visual way. It adds a new dimension to the game and helps to make it more immersive and engaging. The series of stunning visuals allow the players to relive their adventure, adding an extra layer of excitement to their experience. This allows the players to recall the most significant moments of their game and serves as a memento of their campaign. It improves the players' gaming experience and enhances the game's storytelling aspect.

4 Results

4.1 Campaign Results

In this section, we present the results of the campaigns generated by ADM. We ran 10 campaigns and evaluated the stories generated by ADM using three metrics: originality, diversity, and creativity.

The originality score was calculated using cosine similarity, which measures the similarity between two documents.

A higher score indicates a greater degree of difference from existing stories. The diversity score reflects the variety of themes, settings, and characters present in the story. The creativity score measures the level of novelty and imagination displayed in the story.

The results displayed in Table 1 show that ADM-generated stories showed a wide range of originality, diversity, and creativity scores. The originality scores ranged from 0.630 to 0.738, with an average score of 0.690, indicating that the generated stories were generally unique and different from each other. The diversity scores ranged from 0.282 to 0.376, with an average score of 0.330, indicating that the stories covered a broad range of topics and themes. The creativity scores ranged from 0.029 to 0.565, with an average score of 0.200, indicating that some stories were more creative than others.

It is important to note that the originality, diversity, and creativity scores are subjective. A higher originality score indicates a higher degree of dissimilarity between the generated story and existing stories, but it does not necessarily reflect the quality of the story. The diversity score measures the number of unique words used in the story, and the creativity score is based on the uniqueness of the story's plot.

Overall, the results suggest that ADM is capable of generating unique and diverse stories with varying levels of creativity. However, further research is necessary to evaluate the quality of the stories generated by ADM and their suitability for D&D campaigns.

Table 1. Results of ADM-generated Campaigns

Story	Originality	Diversity	Creativity
story_0.txt	0.698	0.376	0.166
story_1.txt	0.703	0.345	0.156
story_2.txt	0.701	0.344	0.233
story_3.txt	0.685	0.339	0.565
story_4.txt	0.630	0.341	0.222
story_5.txt	0.661	0.336	0.090
story_6.txt	0.701	0.313	0.050
story_7.txt	0.676	0.369	0.029
story_8.txt	0.738	0.282	0.081
story_9.txt	0.716	0.281	0.131

4.2 Summarization Results

Our study aimed to evaluate the effectiveness of two other transformer-based models, DistilBERT-CNN and Google T5, for text summarization. We performed the evaluation on 10 summaries generated by ADM. To assess the performance of these models, we used the summaries generated by the BART-large-CNN-SAMSum model as a benchmark for comparison. We chose the BART-large-CNN-SAMSum model as our benchmark because, in our initial experimentation, we found that it produced the highest quality abstractive summaries compared to other models.

To compare the performance of the models, we used ROUGE metrics, which are commonly used to evaluate the quality of text summarization. ROUGE stands for Recall-Oriented Understudy[4] for Gisting Evaluation and measures the overlap between the generated summary and the reference summary (in this case, the summary produced by the BART-large-CNN-SAMSum model).

By comparing the ROUGE scores of the summaries generated by the DistilBERT-CNN and Google T5 models with the benchmark summaries generated by BART-large-CNN-SAMSum, we were able to assess the effectiveness of these models for text summarization. This allowed us to identify which model performed the best and to what extent the model's performance deviated from the benchmark.

Table 2. The ROUGE-1 metrics results for Google T5 model

Summary File	Rouge-1 Recall	Rouge-1 Precision
summary_0.txt	0.8292682927	0.3063063063
summary_1.txt	0.5227272727	0.3538461538
summary_2.txt	0.3877551020	0.2714285714
summary_3.txt	0.5000000000	0.2222222222
summary_4.txt	0.4629629630	0.3521126761
summary_5.txt	0.6097560976	0.2450980392
summary_6.txt	0.5090909090	0.2857142857
summary_7.txt	0.5957446809	0.3043478261
summary_8.txt	0.33333333333	0.2653061224
summary_9.txt	0.5428571428	0.2835820895

Table 3. The ROUGE-2 metrics results for Google T5 model

Summary File	Rouge-2 Recall	Rouge-2 Precision
summary_0.txt	0.560000	0.160000
summary_1.txt	0.240741	0.1313131313
summary_2.txt	0.233333	0.1473684211
summary_3.txt	0.279412	0.1187500000
summary_4.txt	0.242424	0.1600000000
summary_5.txt	0.416667	0.1290322581
summary_6.txt	0.300000	0.1544117647
summary_7.txt	0.300000	0.1451612903
summary_8.txt	0.057692307692	0.045454545454
summary_9.txt	0.25	0.096

We can observe in Table 2 that for the ROUGE 1 metrics, Google T5 model performs best for summary_0 with a Recall value of 0.82, which means it retains all the information from the story, and the performance is not very good for summary_2 with a recall value of 0.38.

We can observe in Table 3 that for the ROUGE 2 metrics, Google T5 model performs best for summary_0 with a Recall value of 0.56, which means it retains all the information

Table 4. The ROUGE-1 metrics results for distilBert model

Summary File	Rouge-1 Recall	Rouge-1 Precision
summary_0.txt	0.7882	0.6036
summary_1.txt	0.6389	0.7077
summary_2.txt	0.6552	0.5429
summary_3.txt	0.8657	0.5370
summary_4.txt	0.3780	0.4366
summary_5.txt	0.6111	0.5392
summary_6.txt	0.4842	0.4694
summary_7.txt	0.5584	0.4674
summary_8.txt	0.3553	0.5510
summary_9.txt	0.8172	0.4533

Table 5. The ROUGE-2 metrics results for distilBert model

Summary File	Rouge-2 Recall	Rouge-2 Precision
summary_0.txt	0.7765	0.5946
summary_1.txt	0.6111	0.6769
summary_2.txt	0.6552	0.5429
summary_3.txt	0.8507	0.5278
summary_4.txt	0.3537	0.4085
summary_5.txt	0.6111	0.5392
summary_6.txt	0.4737	0.4592
summary_7.txt	0.5065	0.4239
summary_8.txt	0.3553	0.5510
summary_9.txt	0.7950	0.4400

from the story, and the performance is not very good for summary_8 with a recall value of 0.057.

We can observe in Table 4 that for the ROUGE 1 metrics, distilbert model performs best for summary_3 with a Recall value of 0.86, which means it retains all the information from the story, and the performance is not very good for summary_8 with a recall value of 0.35.

We can observe in Table 5 that for the ROUGE 1 metrics, distilbert model performs best for summary_3 with a Recall value of 0.85, which means it retains all the information from the story, and the performance is not very good for summary_4 with a recall value of 0.35.

4.3 Visualization Results

The proposed model was able to generate appropriate visualizations from the D&D campaign effectively. Some of the lines from different summaries and the respective images are shown below.

"Greg the wizard, Ralph the Warlock, and Mandy the Bard had been traveling together for several months, each seeking their own purpose."

Figure 3 features three characters, each with distinct clothing and features, traveling together on a path. The image

Figure 3. Image generated by DALL-E: Greg the wizard, Ralph the Warlock, and Mandy the Bard

also conveys a sense of camaraderie and adventure, with the three characters appearing to be united in their quest. The image is able to capture the essence of the textual description, highlighting the characters' sense of adventure and purpose while also conveying a sense of magic and fantasy.

Figure 4. Image generated by DALL-E: Hidden chamber

"They discovered hidden chambers filled with secrets and treasures, which shed new light on the true nature of the artifact they were seeking."

Figure 4 features a dark and mysterious room with various objects and artifacts scattered around. The image evokes a sense of mystery and discovery, with the viewer being drawn in to explore the hidden depths of the chamber.

"A powerful necromancer is controlling the goblins, using them as his spies."

Figure 5 features a dark and eerie scene, with the necromancer standing at the center of a group of goblins. The

Figure 5. Image generated by DALL-E: Necromancer and the goblins

necromancer is depicted as a sinister figure with dark robes and glowing eyes, while the goblins are shown as small and mischievous creatures. The overall atmosphere of the image is one of danger and foreboding, with the viewer sensing that the necromancer and his goblin minions are up to no good. The image also suggests a sense of power, with the necromancer exerting control over the goblins through his magic.

Figure 6. Image generated by DALL-E: Strange sightings in a cave

"The villagers had reported strange sightings and eerie sounds coming from a nearby cave."

Figure 6 features a dark and foreboding cave entrance, with strange lights and shadows emanating from within. The image suggests a sense of danger and mystery, with the viewer wondering what strange creatures or supernatural forces might lurk within the cave. The image also features a villager near the entrance to the cave.

Figure 7 shows the photo album created on the webpage for a D&D campaign. It describes the events generated using GPT-3 and makes it helpful for campaigns that stretch out over a long period of time or that have many different locations and characters. This helps the players visualize the world more clearly, increasing their immersion in the game. This can help players feel more connected to the story and their characters. This can help build a sense of community around the campaign and allow others to get involved in the story.

5 Conclusion

Our research into automated systems for Dungeons and Dragons campaigns has shown promising results. Our experiments have demonstrated the effectiveness of three main modules: the automated dungeon master (ADM), the Summarizer, and Viz. Our experiments showed that ADM was able to generate a range of unique and diverse stories with varying levels of creativity. Viz generated appropriate visualizations of key plot points, providing players with breathtaking artwork.

The use of automated systems in D&D campaigns has the potential to revolutionize how the game is played and make it more accessible to newcomers. DMing can be a challenging task that requires creativity, quick thinking, and the ability to adapt to unpredictable situations. Automating this process could reduce the workload on the DM and make the game more accessible to newcomers.

However, creating an AI that can effectively run a D&D campaign is daunting. D&D is a game that relies heavily on storytelling, improvisation, and the ability to adapt to unforeseen circumstances. These are all tasks that are notoriously difficult for AI systems to accomplish. Furthermore, D&D campaigns can be highly subjective, with different DMs and players having different interpretations of the game's rules and lore.

Despite these challenges, our research has shown that automated systems can effectively assist in running and summarizing D&D campaigns.

6 Future Work

In future research, we aim to evaluate the quality of the stories generated by ADM and their suitability for D&D campaigns. We also plan to investigate the potential of incorporating player feedback into the automated system to improve the overall experience. Ultimately, our research aims to push the boundaries of what is possible with automated systems in D&D campaigns and provide new insights into the intersection of AI and tabletop gaming.

Additionally, we plan to continue to enhance our automated system for D&D campaigns with more exciting features. One such feature is in-game Text-to-speech voiceover, which will enable players to hear descriptions and dialogues

Figure 7. Photo album of a D&D campaign

in the game read aloud by the AI. Another exciting feature is Speech-to-text input, which will allow players to speak their actions and dialogues and have them automatically translated into text and incorporated into the game.

We also plan to incorporate the input as character sheet feature, which will enable players to input their characters' information and abilities into the system, making it easier for the AI to generate tailored scenarios for each player.

Moreover, we aim to implement Dynamic music generation, which will enable the AI to generate music in real time that matches the mood and tone of the game. This feature will provide players with an immersive and dynamic gaming experience.

Finally, we also plan to integrate Stable Diffusion in order to generate artwork from the campaign and draw comparisons between it and the images generated by DALL-E.

By continuing to improve our AI system with new and exciting features, we hope to revolutionize the way D&D campaigns are played and make them more accessible and enjoyable for everyone.

References

[1] Tom B. Brown et al. *Language Models are Few-Shot Learners.* 2020. arXiv: 2005.14165 [cs.CL].

[2] Yun-Gyung Cheong and Byung-Chull Bae. "Procedural content generation for games". In: *Communications of the Korean Institute of Information Scientists and Engineers* 31.7 (2013), pp. 16–25.

[3] Matthew Guzdial and Mark Riedl. "Automated game design via conceptual expansion". In: *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment.* Vol. 14. 1. 2018.

[4] Chin-Yew Lin. "Rouge: A package for automatic evaluation of summaries". In: *Text summarization branches out.* 2004, pp. 74–81.

[5] Xiangyang Liu et al. "Late Prompt Tuning: A Late Prompt Could Be Better Than Many Prompts". In: *arXiv preprint arXiv:2210.11292* (2022).

[6] Mark J. Nelson and Michael Mateas. "Towards automated game design". In: *AI*IA 2007: Artificial Intelligence and Human-Oriented Computing: 10th Congress of the Italian Association for Artificial Intelligence, Rome, Italy, September 10-13, 2007. Proceedings.* Vol. 10. Springer. 2007, pp. 40–51.

[7] Konstantinos Ntokos. "The blackthorn manor: a case study in teaching software engineering for computer games courses using codeplay framework". In: *The Computer Games Journal* 9.1 (2020), pp. 61–74.

[8] Aditya Ramesh et al. "Zero-shot text-to-image generation". In: *International Conference on Machine Learning.* PMLR. 2021, pp. 8821–8831.

[9] Noor Shaker, Julian Togelius, and Mark J. Nelson. "Procedural content generation in games". In: (2016).

[10] Tanya Short and Tarn Adams. *Procedural Generation in Game Design.* CRC Press, 2017.

[11] Disha Shrivastava, Hugo Larochelle, and Daniel Tarlow. "Repository-level prompt generation for large language models of code". In: *arXiv preprint arXiv:2206.12839* (2022).

[12] Roland Van Der Linden, Ricardo Lopes, and Rafael Bidarra. "Procedural generation of dungeons". In: *IEEE Transactions on Computational Intelligence and AI in Games* 6.1 (2013), pp. 78–89.

[13] Ashish Vaswani et al. "Attention is all you need". In: *Advances in neural information processing systems* 30 (2017).

[14] Zhuofeng Wu et al. "IDPG: An instance-dependent prompt generation method". In: *arXiv preprint arXiv:2204.04497* (2022).