FabulAI: Artificial Intelligence for Storytelling in Italian Narrative Adventures

Elisa Vittoria Cosmai[1,*,†], Lucia Siciliani[1,†] and Pierpaolo Basile[1,*,†]

[1]University of Bari Aldo Moro, Bari, Italy

Abstract

This paper presents FabulAI, a conversational system designed to facilitate the creation of interactive, text-based stories in Italian through a Telegram chatbot powered by Large Language Models (LLMs). The system aims to make narrative generation accessible to non-expert users by combining guided storytelling mechanics with the linguistic capabilities of LLMs. The paper describes the system's modular architecture, narrative design approach, and real-time interaction flow, supported by open-weight LLMs such as LLaMA-3. A user study involving 31 participants was conducted to evaluate engagement, immersion, and user satisfaction using the Game Experience Questionnaire (GEQ). Results indicate high levels of positive emotions, immersion, and willingness to replay, while highlighting areas for improvement such as narrative length and perceived challenge. FabulAI demonstrates the potential of generative AI as a creative partner rather than a replacement, offering a replicable model for accessible interactive storytelling across educational and entertainment contexts.

Keywords

Narrative Design, Narrative Intelligence, Artificial Intelligence, Interactive Fiction, Video game,

1. Introduction

The desire to tell stories and experience them first-hand has always accompanied the evolution of human culture. Over time, storytelling has evolved to take on increasingly interactive forms, capable of involving the user not only as a spectator but as an actual protagonist of the story. A pivotal moment in this area occurred in the late 1970s and 1980s with the emergence of text adventures interactive experiences in which the player, exclusively in text mode, explored imaginary worlds, solved puzzles, and experienced the story as if he or she could actually influence the course of the plot [1, 2].

This form of interactive fiction was one of the first expressions of the fusion of fiction and technology, paving the way for a new conception of entertainment and imagination [3]. Although it has been partially eclipsed by more visual and multimedia forms over the years, the text adventure retains a timeless appeal and is enjoying a new lease of life thanks to recent developments in artificial intelligence.

In particular, the advent of Large Language Models (LLMs) has rekindled interest in dynamic, personalized narratives, making it possible to generate coherent, rich, contextualized text in real time. These models, trained on massive amounts of linguistic data, are able to understand user intent and respond fluidly and naturally, creating the illusion of an interlocutor capable of reasoning, remembering, and creativity [4, 5].

Their integration into accessible and ubiquitous environments, such as instant messaging platforms, has given rise to a new generation of conversational chatbots that can guide users in constructing interactive stories by merging natural language with narrative logic. In this way, storytelling is being reinvented, returning to its text-based origins, but with more powerful and flexible means that can deliver engaging, open-ended, and much more customizable experiences than in years past [6].

Proceedings of AI4HGI '25, the First Workshop on Artificial Intelligence for Human-Game Interaction at the 28th European Conference on Artificial Intelligence (ECAI '25), Bologna, October 25-30, 2025

*Corresponding author.

†These authors contributed equally.

elisavittoria.cosmai@gmail.com (E. V. Cosmai); lucia.siciliani@uniba.it (L. Siciliani); pierpaolo.basile@uniba.it (P. Basile)

https://www.linkedin.com/in/elisa-vittoria-cosmai-9b881019b/ (E. V. Cosmai);

https://swap.di.uniba.it/members/siciliani.lucia/ (L. Siciliani); https://swap.di.uniba.it/members/basile.pierpaolo/ (P. Basile)

0009-0009-3014-4324 (E. V. Cosmai); 0000-0002-1438-280X (L. Siciliani); 0000-0002-0545-1105 (P. Basile)

At a time when generative artificial intelligence is emerging as a pervasive technology, it is crucial to explore not only its functional potential but also its cultural and creative impact. Among the most promising areas in this regard is that of **interactive storytelling**, understood not only as entertainment but also as a tool to stimulate reflection, creativity and learning. The possibility of actively involving the user in a narrative process that is guided, but at the same time open to improvisation, represents an evolved form of cognitive participation, capable of combining logical rigour and imagination [7].

The resurgence of interest in text adventures, fueled by recent innovations in language models, underscores a still-relevant demand for narrative experiences that rely on the written word, choice, and linguistic exploration. In this context, the possibility of using an intelligent chatbot to assist users in constructing a personal story through a structured and responsive conversation emerges as a concrete opportunity to rethink the role of conversational technologies.

The motivation for our work stems from the desire to investigate how the combination of LLMs, narrative game mechanisms, and conversational interfaces can give rise to new digital tools geared towards creativity. Such tools will not only respond to commands but will also become true *narrative mediators*, capable of guiding, suggesting, improvising, and adapting to the user in a coherent and stimulating way. Our goal is to contribute to this scenario with a design approach that increases both the accessibility of the technology and the depth of the narrative experience.

This paper aims to explore and document an integrated approach to guided generation of interactive narratives using chatbots based on advanced language models. The main objective is to design and analyze a conversational system that guides the user in the construction of a textual story, enhancing both the creative and technical aspects of the interaction.

More specifically, our contribution aims to:

- **Develop an interactive narrative infrastructure** based on a Telegram bot, accessible also to non-experts, capable of guiding the user through a structured path of text creation;
- **Design a dialogic interaction model** that simulates the behavior of a narrator, adapting to the user's choices and style;
- **Analyze the dynamics of story generation and management** at the technical, linguistic, and functional levels, with attention to quality of user experience and narrative coherence;
- **Evaluate the potential of the system in application areas** such as creative writing, language education, and experimentation with emerging narrative forms.

Through these objectives, our work aims to contribute to research on the creative use of chatbots by providing a concrete and replicable case study that combines interactive storytelling, artificial intelligence, and conversational interfaces.

The work presented focuses on the design, implementation, and evaluation of a conversational system for generating interactive text narratives. A Telegram-based chatbot was developed that guides the user through a sequence of structured messages in the personalized construction of a story, acting as a narrative facilitator. The approach combines elements of narrative design, conversational modeling, and user interaction, with the goal of providing an accessible and engaging experience even for those without specific creative or technical skills.

To evaluate the quality of the proposed experience, a study was conducted with real users. Participants interacted with the system and then completed the Game Experience Questionnaire (GEQ) [8], a widely used instrument in the literature to measure the level of engagement, enjoyment, excitement, and immersion in ludic or simulated contexts. The collected results indicate a high level of engagement, suggesting that utilizing a chatbot for guided storytelling can be an effective way to stimulate personal creativity while maintaining a smooth and intuitive user experience.

Finally, we would also like to draw attention to the ethical implications of the interaction between artificial intelligence and human creativity. The use of a conversational system such as *FabulAI* raises important considerations about the role of AI in the creative process: while the tool guides, suggests, and structures the narrative experience, it by no means claims to replace the user's imagination. On the contrary, *FabulAI* is conceived as a non-invasive support, stimulating individual creativity by offering

clues, incipits, and paths to explore, but always leaving the user in full control of content and style. The conversational agent has no true autonomous creative capacity: it acts on the basis of predefined rules, prompts, and responses. From this perspective, the value of the system lies not in the original production of content but in its ability to activate the imagination, facilitate narrative flow, and accompany the user on a personal path of expression and discovery.

The paper is structured as follows: in section 2 we will explore the advancements in AI-driven chatbots and NLP models and their impact on interactive storytelling, section 3 will present the design and implementation of FabulAI, a Telegram-based chatbot utilizing Large Language Models (LLMs) to create interactive textual narratives. In section 4, we will describe the experimental assessment of FabulAI using the Game Experience Questionnaire (GEQ), focusing on user engagement, immersion, and narrative quality, while 5 will wrap up the work with its conclusions and highlight possible future direction. Finally, in section 6, we will discuss the limitations of the present work.

2. Related Work

In recent decades, artificial intelligence and machine learning have significantly transformed the way humans interact with natural language processing systems. Chatbots and text generation models, in particular, represent a constantly evolving field of research that has led to the development of increasingly advanced tools capable of producing coherent and realistic textual content.

Chatbots are programs designed to simulate conversations with humans through text- or voice-based interfaces. Often powered by AI technologies, they are most commonly used on e-commerce platforms to increase sales, improve productivity, and save time. Natural language processing (NLP) enables interaction between machines and humans using everyday language. NLP methods are widely used to develop Telegram chatbots[9]. Chatbots can provide content in text and audio formats. They are helpful for discussions on creative entrepreneurship. Chatbots can answer questions on various subjects and be used in education to enhance skills and support the development of creative and technical competencies.

An example of a historical rule-based chatbot is ELIZA [10], developed by Joseph Weizenbaum. ELIZA simulated a therapist by responding to users with questions based on textual patterns. Today, chatbots powered by GPT-4 or Claude 3 use Large Language Models (LLMs) to generate much more precise and adaptable responses [11]. In recent years, models such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer) have significantly improved the quality of text generation and revolutionized the field [12, 13].

Natural Language Processing (NLP) techniques are widely used for processing narrative content in tasks such as sentence segmentation, entity recognition, information extraction, and anaphora resolution [14]. On the other side, NLP enables the construction of complex narratives through event graphs, generating semantically richer texts. An approach based on knowledge graphs enables the mapping of precise causal relationships between events, thereby enhancing the grammatical and semantic accuracy of generated narratives[15]. Frameworks such as PANGeA use NLP to ensure narrative coherence in video games and prevent out-of-context inputs. The system dynamically evaluates textual input against the game's rules, aligning content generation with narrative intent. NLP also helps maintain coherent narrative conversations between non-playable characters (NPCs) and can overcome the context-length limitations of language models [16]. Another important research direction is to provide metrics that allow the quality and fidelity of generated content to be evaluated and compared with texts written by human authors [17].

The most advanced text generation models are based on LLMs. GPT models developed by OpenAI can generate fluent, well-structured text thanks to billions of parameters trained on large volumes of textual data [4]. LLMs offer a new level of flexibility in procedural generation, enabling the creation of dynamic plots that evolve with user input [16, 18]. LLMs can also generate playable narrative assets, such as scripts for non-playable character (NPC) interactions and real-time dialogues. PANGeA (Procedural Artificial Narrative using Generative AI) [16], on the other hand, uses LLMs to generate narrative content for turn-based RPGs. It includes a memory system, a validation system, a Unity plugin, and

a server with a RESTful interface. The system procedurally generates level data, including settings, key items, non-player characters (NPCs), and dialogues, based solely on configuration and design rules provided by the game designer. Furthermore, PANGeA utilises the Big Five Personality model to inform NPC responses, thereby enhancing the realism of interactions between players and NPCs. However, LLMs may produce off-topic content in response to free-form text input, necessitating the integration of validation systems [16].

One of the most recent and well-known examples in this field is the Nemesis System [19]. It is an artificial intelligence and gameplay system that was first introduced in *Middle-earth: Shadow of Mordor (2014)* and later refined for *Middle-earth: Shadow of War (2017)*. Both games were developed by Monolith Productions. The system's primary objective is to generate "emergent" stories through the interaction between players and AI-controlled enemies, ensuring that each playthrough is unique. The system can create a dynamic hierarchy of enemies, particularly orcs, who evolve and remember previous encounters with the player. For example, a simple warrior can become a captain by killing the player. When the player respawns, they may encounter the same enemy, who will remember the previous battle and alert the rest of the army to attack. This is one of the few examples of AI in video games that can generate believable narratives and enrich the game world dynamically.

Narrative bots represent an interesting evolution of interactive storytelling. They enable users to actively participate in shaping the story through their choices and interactions, doing so in a more immediate and accessible manner. One of the most well-known projects in this area is AI Dungeon [1], a platform that utilizes LLMs to generate interactive stories in real-time. AI Dungeon uses advanced neural networks to interpret and respond to user input, creating personalized, constantly evolving storylines based on choices made.

In addition to AI Dungeon, other games and chatbots have demonstrated the potential of interactive storytelling on messaging platforms and beyond. Lifeline[2] is a mobile game based on text choices and simulates a conversation with a stranded astronaut. Users must make decisions that influence the protagonist's fate. Reigns[3] is a narrative system that uses swipe mechanics to let players make political decisions for their kingdom and thereby affect the story's development. FlorenceBot is a Telegram bot developed for patients with brain tumours as part of the INNV-13 Telegram chat application. The King is a system that evaluates the validity of user-provided stories, offers feedback, and continues the narrative when appropriate. It uses a "one-shot example" prompt to ensure the model follows the correct format. This system is used in the video game 1001 Nights and is based on GPT-4 [20]. AdventurA GPT[4] is a GPT-4–based model that acts as a Dungeon Master within a story and creates a unique narrative for each player. Heartfelt Narratives is a project that explores evoking empathy through storytelling and sharing personal experiences. It provides publicly available annotations, study findings, and language model outputs [21, 22].

3. Methodology

Our work aimed to design and implement a conversational system that can generate interactive stories in text by exploiting LLMs. The primary goal of the entire process is to ensure replicability, which we define as the ability of the design to be reproduced independently by third parties using the provided instructions, code, and configurations. For this reason, we adopted a modular, documented approach utilising open-source technologies and publicly accessible services.

FabulAI is a Telegram chatbot that guides users in creating personalized, interactive narratives. Each interaction follows a structured yet flexible flow. First, the user defines initial parameters, such as setting, role, and character name. Then, the user acts freely, scene by scene, as the system generates narrative text in response to their actions.

[1] https://aidungeon.com/
[2] https://www.3minute.games/
[3] https://www.reignsgame.com/reigns
[4] https://chatgpt.com/g/g-676ff9c5b4e08191bcca8c6a10cccf70-adventura-gpt

The methodological approach is based on four fundamental principles.

1. **Modularity**: Each component of the system (e.g., interface, input management, text generation, and data storage) is designed as an independent module to facilitate understanding, maintenance, and reuse;
2. **Incrementality**: Development occurred in progressive stages, with individual functional blocks (e.g., bot initialization, shift management, LLM integration, and database saving);
3. **Accessibility**: Free, accessible tools and services (e.g., Telegram, Together AI, and SQLite) are chosen;
4. **Realistic Interaction**: The user should perceive the narrative experience as fluid, coherent, and immersive. To this end, a dynamic prompting system and a turn-based narrative structure with defined limits are adopted.

All the software is written in Python and utilizes open libraries, including python-telegram-bot, SQLite3, and the Together AI API for prompting LLM models. Together AI was chosen for the ability to use high-level templates free of charge (with monthly limits), and for direct compatibility with Python. In particular, we rely on the Llama-3.3-70B-Instruct-Turbo model provided by Meta. The database is used to store stories and all related attributes such as: user ID, story title, narrative phase, character, setting, protagonist name, accumulated text, turn counter, and timestamp.

Each component of the system is independently built and designed to be reusable, ensuring ease of maintenance and flexibility. The system is structured around four main modules and follows a turn-based user-bot interaction flow that saves narrative state persistently.

The four modules are:

1. **Main** initializes the Telegram bot, loads environment variables (token and API key), registers handlers, and starts polling.
2. **Handler** contains the interaction management logic. It handles the entire game cycle, including receiving input, checking the state of the narrative, building prompts, sending requests to the LLM model, and receiving and forwarding the response to the user.
3. **DB** deals with creating and managing the SQLite database. It provides functions to create new histories, update progress in the story, and load and save user data.
4. **Config** collects project constants, including settings, narrative roles, the maximum number of turns, and guiding messages.

The interaction between the modules is sketched in Figure 1

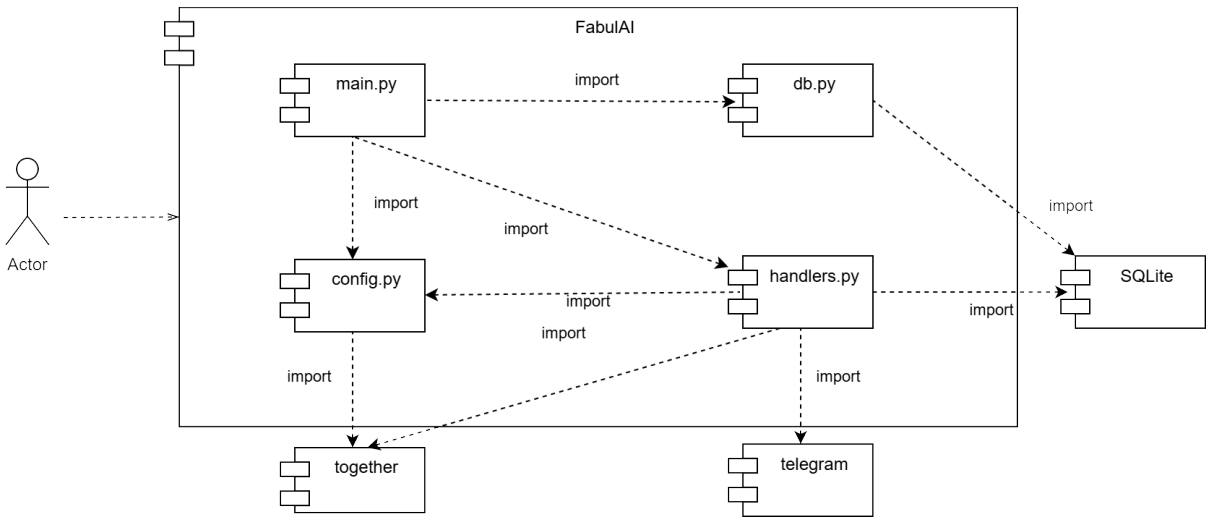


Figure 1: Component Diagram

3.1. User-Bot Interaction Flow

FabulAI guides the user through a structured sequence.

1. **Start with /start command**: The bot checks if there are any saved stories for the current user. If not, the bot proposes creating a new story.
2. **Story creation**: The user selects a setting (e.g., fantasy or cyberpunk) and a character (e.g., wizard or knight), then enters the name of the main character. Each choice is guided via an interactive keyboard.
3. **Incipit generation**: Once the initial information has been collected, a prompt is constructed to generate the introductory paragraph via the LLM template. The result is sent to the user.
4. **Turn-based game phase**: At each turn, the user enters a free-text action. The system then updates the narrative context and generates a continuation consistent with the accumulated story. This process repeats up to the maximum number of turns (e.g., 10).
5. **Generation of Ending**: When the set limit is reached, a consistent narrative conclusion is generated, closing the story arc.
6. **Saving and resuming**: All information is saved in real time in the database. Users can stop and resume the story at any time.

FabulAI utilizes a state structure to track the narrative's progress. Each story is associated with a unique identifier and a current state (*phase*), which can be one of the following: 1) Setting - setting selection; 2) Character - selection of role; 3) Name - entry of the protagonist's name; 4) Game - active phase of the narrative; 5) End - conclusion of the story.

With each state change, the database is updated to record the full narrative text, the number of turns taken and a timestamp. This management system ensures narrative continuity, even in the event of an interruption, and allows for easy extension to future scenarios, such as multi-user matches or parallel stories.

To facilitate interaction, the system uses custom Telegram keyboards instead of free text input in the early stages. This approach reduces the risk of errors, streamlines the selection process, and enhances the overall user experience. Additionally, all bot responses are accompanied by guidance messages to facilitate understanding of the flow and available commands. Interactions are handled asynchronously to ensure responsiveness, even with multiple active users.

3.2. Narrative Text Generation

The heart of the FabulAI system is the narrative text generation. The goal is to produce coherent, creative, and contextually relevant content that shapes a progressive and personalized narrative based on user choices and actions. To this end, a large language model (LLM) has been integrated via the Together AI API.

At each turn, the system generates a dynamic prompt. This prompt provides the model with all the necessary information to generate a credible and engaging narrative continuation. The prompt is automatically generated based on the current state of the story. It includes the role of the model (in this case, an expert narrator), the narrative context, the user's actions in the current turn, and guidance on the narrative style.

Example of prompt used during the story:

```
You are a text adventure storyteller.
The player is playing a {character} named {name} in a world {environment}.
Current story status:
{story}
Player action:
{action}
Continue the story in a coherent way, describing what happens after the action.
Max 180 words.
```

```
If the player uses foul language try to integrate it into the story without
pointing out his language.
Do not make references to real people.
```

This approach enables the model to produce output that maintains narrative consistency without requiring fine-tuning or additional training.

3.2.1. Generation Parameters

The Together AI API call includes a set of generation parameters that can impact the output generated by the model, such as: Temperature, Top-K, Top-p, Repetition Penalty and Max Tokens. In FabulAI, these parameters are not kept frozen throughout the interaction with the user; in fact, they are modified to optimise and balance creativity and consistency at each narrative stage. During the incipit generation, the temperature and length are set to high values to enrich the introduction, in the turn-based game phase, the parameters are then balanced to maintain plot control and try to ensure coherence, finally in the generation of ending, they are modified to have greater focus on consistency and narrative closure.

Throughout all phases, we employ a zero-shot prompting approach. Each response is generated from the dynamic prompt without preset examples. This mode is well-suited for open and flexible scenarios, such as textual stories. The generated output is saved in the SQLite database as a continuation of the story and sent to the user via a Telegram message. This process repeats at every turn, updating the context and adapting in real time to the user's decisions.

In the final stage, the prompt includes explicit instructions to stimulate a plot conclusion. For example: "Generate an ending (max. 200 words) that concludes the adventure satisfactorily but still leaves open the possibility of a sequel. Adapt to the state of the story at that time, considering the player's previous actions, and generate an ending that is consistent with the rest."

3.2.2. Example of the FabulAI Telegram Bot Interface

The Telegram bot interface was developed to facilitate intuitive interaction with the narrative system, guiding users through each stage of story creation and development. As illustrated in Figure 2a and Figure 2b, the interface employs interactive buttons and guided messages to simplify the selection of key narrative elements (setting, character, and protagonist name), followed by a turn-based interaction phase.

All interactions between the user and the bot are conducted in Italian. The prompts, generated texts, and guiding instructions are fully localized, ensuring semantic coherence and fluency in the target language.

The code of FabulAI is available on GitHub[5].

4. Evaluation

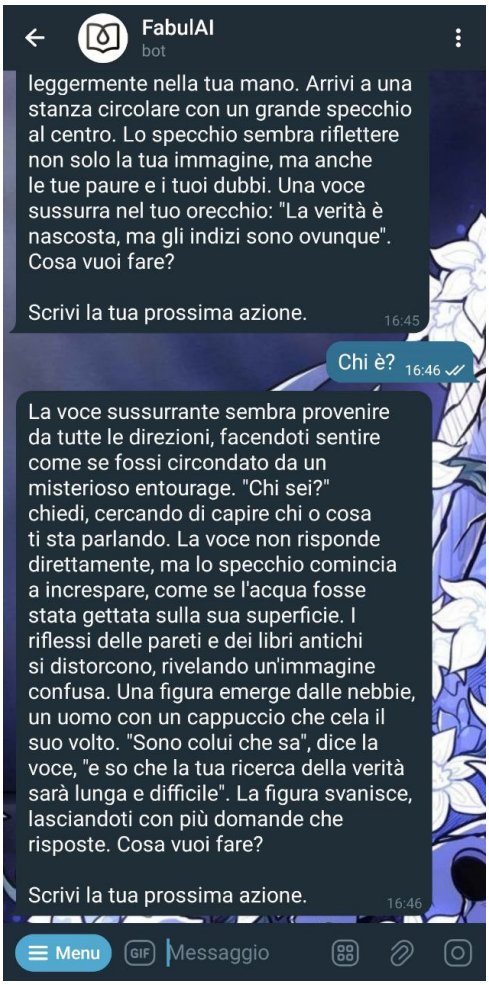
The experimental evaluation phase provides an objective assessment of the FabulAI system's effectiveness in generating high-quality interactive narratives. Involving real users in the testing phase allows us to observe the system's performance and evaluate its impact on entertainment and engagement.

The evaluation aims to understand how users perceive interaction with the narrative chatbot and assess whether this modality can stimulate individual creativity, maintain interest over time, and provide a fluid, coherent narrative. In this sense, the system is regarded not merely as a set of technical functions but also as a narrative and communicative device with its immersive potential.

An additional objective is to evaluate the impact of artificial intelligence. The underlying hypothesis of the experiment is that a conversational agent designed with attention to narrative context and equipped with a guided structure can naturally and engagingly accompany the user in constructing a story without imposing itself or limiting the user's creative freedom.

[5] https://github.com/EC2002/FabulAI

(a) Story and Character Creation



(b) Interaction with the Bot

To investigate these dimensions, we used the Game Experience Questionnaire (GEQ) [8], an internationally validated tool in the field of game experience research. The GEQ is particularly well-suited to this type of research because it considers more than just functional or technical aspects, offering a comprehensive view of the user experience.

The GEQ is structured into multiple modules, each designed to evaluate specific aspects of the user experience. In this study, the following modules are used:

- **The In-Game Module** evaluates how the user feels during interaction, including immersion, tension, positive or negative emotions, sense of competence, and perceived challenge.
- **The Post-Game module** measures impressions after the experience (e.g., fatigue, recovery, satisfaction, and willingness to repeat the activity).
- **Additional Items**: includes specific binary-response questions that directly investigate levels of enjoyment, willingness to reuse the system, and sense of control over the experience.

Responses are collected on a 5-point (0-4) Likert scale ranging from “not at all” to “very much” and, in some cases, through “yes/no” questions to facilitate analysis and interpretation.

The GEQ was administered digitally via a Google Forms module accessible through a link. We merge the in-game and post-game modules that are compiled at the end of the experience. We avoid interrupting the story to compile the in-game module since the game session is short and we want to measure the user experience when the story ends. Participants were encouraged to complete the questionnaire immediately after finishing their story to preserve perceptual fidelity and minimize memory-related distortion.

The main dimensions assessed by the GEQ are:

- **Immersion**: This measures how involved and absorbed the user feels in the narrative experience, to the point of forgetting the real environment.
- **Positive Emotions**: Evaluates the presence of pleasant feelings, such as enthusiasm, enjoyment, and satisfaction.
- **Negative Emotions**: Detects states such as frustration, boredom, discomfort, or irritation during interaction.
- **Tension**: Measures the perceived pressure or anxiety during gameplay and helps determine whether the experience is relaxing or stressful.
- **Competence**: Detects feelings of mastery, control, and the ability to manage the interaction.
- **Challenge**: Measures the perceived difficulty encountered during the experience in relation to one's abilities.
- **Recovery**: Analyzes the feeling of relaxation or unwinding after the experience.
- **Fatigue**: Identifies any mental or physical tiredness caused by the activity.
- **Desire to Continue**: Indicates how willing the user would be to repeat the experience in the future.

4.1. Experimental Setting

During the testing phase, users interacted with the FabulAI chatbot via the Telegram app on their devices (smartphones or PCs) without direct supervision to preserve the experience's autonomy and spontaneity.

Users received a direct message with a link to start the story. Once the conversation with the bot began, users were guided through the initial stages of selecting a setting, character, and protagonist's name. Then, they participated in the interactive narrative for a total of ten turns.

Afterwards, the system generated an automatic conclusion and provided a link to a Google Forms module containing the GEQ questionnaire. Users are invited to complete the questionnaire immediately after finishing their story to ensure a vivid memory of the experience.

The average interaction lasts 10–15 minutes, and completing the questionnaire takes an additional 2–4 minutes. The entire process lasts less than 20 minutes and does not require technical support.

Thirty-one participants took part in the experiment, and all completed both the narrative interaction and the questionnaire.

The sample consisted primarily of university students and young adults between the ages of 20 and 32, and is balanced in terms of gender. While most participants report moderate familiarity with digital technologies, fewer have experience with text-based games or interactive fiction. Some participants have no experience with video games. This diversity in background enabled us to test the system's accessibility and usability across a varied audience representative of potential casual users.

4.2. Analysis of Results

The data collected through the GEQ is processed to obtain a quantitative overview of the main dimensions of the user experience. For each measured scale (immersion, positive emotions, negative emotions, tension, competence, challenge, recovery, fatigue, and desire to continue). The results are reported in Figure 3

Following, we analyzed results for each dimension of the GEQ questionnaire and report the average score for each of them.

- **Positive emotions (3.0)**: Participants expressed very positive feelings about the overall experience. This suggests that the game evoked positive emotions, such as happiness, satisfaction, and pride, while interacting with the bot.
- **Sensory and imaginative immersion (2.8)**: Users reported a high level of engagement, confirming the bot's ability to provide immersive, creative experiences.

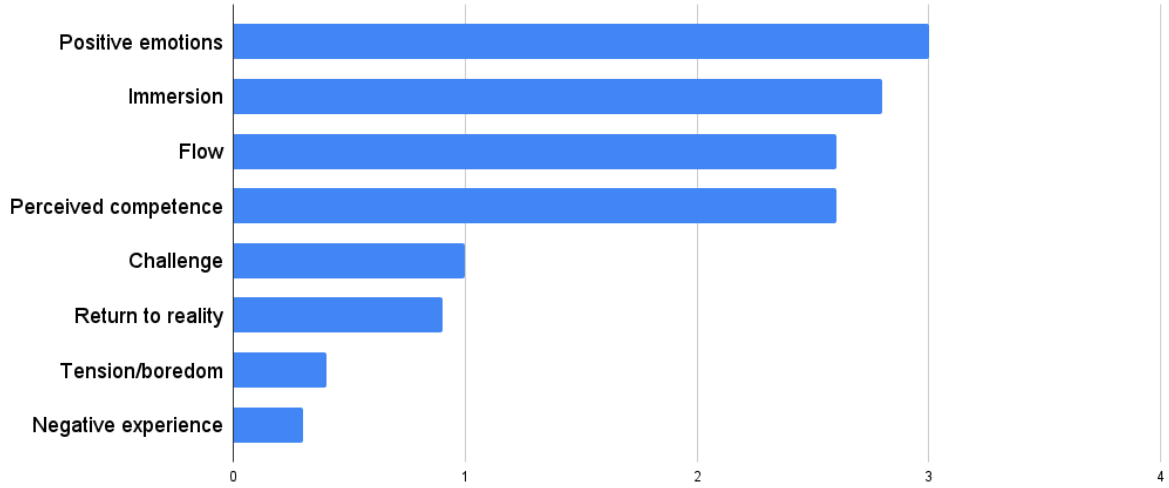


Figure 3: Results of the Game Experience Questionnaire.

- **Flow (2.6)**: Most users maintained interest and focus throughout the story, suggesting that the narrative was engaging enough.
- **Perceived competence (2.6)**: Participants generally felt capable of handling the game's challenges and reported a sense of control and competence.
- **Challenge (1.0)**: The game was not perceived as particularly difficult. This suggests that the experience was too easy, and increasing the challenge could enhance the sense of accomplishment.
- **Return to reality (0.9)**: Most participants had no difficulty returning to reality after the game, indicating that the experience was immersive yet comfortable.
- **Tension/boredom (0.4)** and **negative experience (0.3)**: The game did not generate significant stress or negative sensations, such as boredom, fatigue, or frustration. This underlines that the emotional experience was well-balanced.

Responses to the binary question are reported in Table 1. The results of the first question clearly show that almost all users would like to continue to experience additional stories, confirming a high overall satisfaction. Moreover, results for the second question indicate that around 6 out of 10 users would have preferred a longer story, while about 4 out of 10 found the length adequate. Offering variable-length options can be helpful in better meeting the needs of different users.

Question	YES	NO
Would you like to play another story?	96.8%	3.2%
Did you wish the story had lasted longer?	58.1%	41.9%

Table 1
Responses to the binary questions.

Analysis of the results obtained through the GEQ confirms that interaction with FabulAI generated a positive, well-balanced experience overall. The high average score for positive emotions and level of immersion indicates that users found the activity engaging, enjoyable, and creatively stimulating.

The solid score in the "flow" dimension suggests that the system-guided narrative successfully maintained users' attention and interest throughout the interaction. Although the perceived competence score is not particularly high, it is sufficient to give users a sense of control and understanding of the game dynamics.

The low values related to tension, fatigue, and negative emotions are an essential indicator of the experience's overall balance: the system provided a safe, non-stressful environment accessible to inexperienced users.

The only negative element that emerged is the low perceived level of challenge, suggesting the potential to introduce optional mechanisms for increasing difficulty or alternative, more complex paths for advanced users.

Finally, responses to binary questions confirm a strong interest in continuing the experience with new stories and highlight a significant desire for longer narratives. This provides a concrete opportunity to improve system personalization by offering users a choice between short and long stories, for example.

5. Conclusions and Future Works

This work aims to design, develop, and evaluate a conversational system that could guide users in creating an interactive textual narrative. Named FabulAI, the system was implemented as a Telegram chatbot. It is based on the integration of large language models that generate coherent, context-aware content in response to user input.

The methodology includes a modular technical development phase followed by empirical experimentation with real users. The entire process is guided by the principles of replicability and accessibility, making the experience engaging and creative for non-expert users. Overall, this study demonstrated the feasibility and effectiveness of an artificial intelligence-supported narrative approach capable of merging linguistic interaction, guided storytelling, and a personalized user experience.

Data collected during the experimental evaluation, based on the Game Experience Questionnaire, indicate that FabulAI provided a positive, accessible, and engaging narrative experience. Most users reported a pleasant experience, highlighting feelings of enthusiasm, curiosity, and fulfilment. The immersion is satisfactory, and the system successfully captures users' attention and maintains their focus throughout the narrative, encouraging active participation in the story's development. Moreover, users demonstrated a good understanding of how the system works, and they do not encounter significant difficulties during interaction, even without supervision. Finally, the emotional stability of the experience is also notable, as the interaction is free of stress or frustration, and no significant negative emotions are reported.

Future developments for the FabulAI project include offering customizable story lengths and difficulty levels, implementing branching narratives for greater interactivity, integrating multimedia elements to enhance immersion, and fine-tuning the chatbot for more nuanced and coherent storytelling.

6. Limitations

While the user sample is sufficient for an exploratory analysis, it remains numerically limited and primarily consists of young adults with a high level of digital literacy. This may have positively influenced the perceived ease of use and willingness to interact with the system.

A second limitation is the fixed length of the narrative, which always unfolds over ten turns. While this approach ensured uniformity across tests, it may be too inflexible for users who prefer shorter or longer experiences. It may also not adapt to all narrative rhythms or play styles.

From a technical standpoint, the interaction is currently text-based and sequential. It lacks multimedia elements, advanced personalisation, and complex branching mechanics. Furthermore, adaptive mechanisms that enable the system to adjust the tone, difficulty, or story structure based on user behaviour have yet to be explored.

Acknowledgments

We acknowledge the support of the PNRR project FAIR - Future AI Research (PE00000013), Spoke 6 - Symbiotic AI (CUP H97G22000210007) under the NRRP MUR program funded by the NextGenerationEU.

Declaration on Generative AI

During the preparation of this work, the author(s) used Grammarly to: Grammar and spelling check. After using these tools, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

[1] W. Wu, H. Wu, L. Jiang, X. Liu, H. Zhao, M. Zhang, From role-play to drama-interaction: An LLM solution, in: L. Ku, A. Martins, V. Srikumar (Eds.), Findings of the Association for Computational Linguistics, ACL 2024, Bangkok, Thailand and virtual meeting, August 11-16, 2024, Association for Computational Linguistics, 2024, pp. 3271–3290. URL: https://doi.org/10.18653/v1/2024.findings-acl.196. doi:10.18653/V1/2024.FINDINGS-ACL.196.

[2] N. Montfort, Riddle machines: The history and nature of interactive fiction, in: R. Siemens, S. Schreibman (Eds.), A Companion to Digital Literary Studies, John Wiley & Sons Ltd, 2013, pp. 267–282. URL: https://dspace.mit.edu/handle/1721.1/129076.

[3] P. Ammanabrolu, W. Cheung, D. Tu, W. Broniec, M. O. Riedl, Bringing stories alive: Generating interactive fiction worlds, in: L. Lelis, D. Thue (Eds.), Proceedings of the Sixteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, AIIDE 2020, virtual, October 19-23, 2020, AAAI Press, 2020, pp. 3–9. URL: https://ojs.aaai.org/index.php/AIIDE/article/view/7400.

[4] Z. Zhao, S. Song, B. Duah, J. Macbeth, S. Carter, M. P. Van, N. S. Bravo, M. Klenk, K. Sick, A. L. Filipowicz, More human than human: Llm-generated narratives outperform human-llm interleaved narratives, in: Proceedings of the 15th Conference on Creativity and Cognition, 2023, pp. 368–370.

[5] A. Umbraško, R. Drury, Applying chatgpt in ai-based dynamic video game narrative generation system, 2023.

[6] J. L. Jacko, Generating narrative in an interactive fiction game, West Virginia University, 2018.

[7] D. Lande, L. Strashnoy, Llm reconstructions of the text narrative with and without the time markers, Available at SSRN 5000687 (2024).

[8] W. A. IJsselsteijn, Y. A. De Kort, K. Poels, The game experience questionnaire, 2013.

[9] R. Hartati, E. B. Manullang, Implementation of telegram chatbot ai with natural language processing (nlp) in learning creative entrepreneurship to develop students' creative and innovative competence, Talenta Conference Series: Local Wisdom, Social, and Arts (LWSA) 7 (2024) 72–79.

[10] J. Weizenbaum, ELIZA - a computer program for the study of natural language communication between man and machine, Commun. ACM 9 (1966) 36–45. URL: https://doi.org/10.1145/365153.365168. doi:10.1145/365153.365168.

[11] C. J. Lynch, E. J. Jensen, V. Zamponi, K. O'Brien, E. Frydenlund, R. Gore, A structured narrative prompt for prompting narratives from large language models: Sentiment assessment of chatgpt-generated narratives and real tweets, Future Internet 15 (2023) 375. URL: https://doi.org/10.3390/fi15120375. doi:10.3390/FI15120375.

[12] T. S. Gunawan, A. B. F. Babiker, N. Ismail, M. R. Effendi, Development of intelligent telegram chatbot using natural language processing, in: 2021 7th International Conference on Wireless and Telematics (ICWT), IEEE, 2021, pp. 1–5.

[13] N. F. Bonn, The impact of integrating artificial intelligence into the video games industry.: A case study, 2023. URL: https://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A1820226&dswid=9529.

[14] B. S. Santana, R. Campos, E. Amorim, A. Jorge, P. Silvano, S. Nunes, A survey on narrative extraction from textual data, Artif. Intell. Rev. 56 (2023) 8393–8435. URL: https://doi.org/10.1007/s10462-022-10338-7. doi:10.1007/S10462-022-10338-7.

[15] M. de Kok, Y. Rebboud, P. Lisena, R. Troncy, I. Tiddi, From nodes to narratives: A knowledge graph-based storytelling approach, in: TEXT2STORY 2024, 7th International Workshop on Narrative Extraction from Texts (Text2Story), colocated with ECIR 2024, 2024.

[16] S. Buongiorno, L. Klinkert, Z. Zhuang, T. Chawla, C. Clark, Pangea: procedural artificial narrative using generative ai for turn-based, role-playing video games, in: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 2024, pp. 156–166.

[17] T. Ichmoukhamedov, J. Hinns, D. Martens, How good is my story? towards quantitative metrics for evaluating llm-generated xai narratives, arXiv preprint arXiv:2412.10220 (2024).

[18] Y. Wang, Q. Zhou, D. Ledo, Storyverse: Towards co-authoring dynamic plot with llm-based character simulation via narrative planning, in: G. Smith, J. Whitehead, B. Samuel, K. Spiel, R. van Rozen (Eds.), Proceedings of the 19th International Conference on the Foundations of Digital Games, FDG 2024, Worcester, MA, USA, May 21-24, 2024, ACM, 2024, p. 55. URL: https://doi.org/10.1145/3649921.3656987. doi:10.1145/3649921.3656987.

[19] I. Parosu, E. Hage, S. Magnusson, The nemesis system: How games create stories, 2022.

[20] Y. Sun, Z. Li, K. Fang, C. H. Lee, A. Asadipour, Language as reality: A co-creative storytelling game experience in 1001 nights using generative AI, in: M. Eger, R. E. Cardona-Rivera (Eds.), Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, October 08-12, 2023, Salt Lake City, UT, USA, AAAI Press, 2023, pp. 425–434. URL: https://doi.org/10.1609/aiide.v19i1.27539. doi:10.1609/AIIDE.V19I1.27539.

[21] G. Lasagna, Il racconto ludonarrativo. Analisi del videogioco come medium narrativo, Ph.D. thesis, UniBO, 2020-2021. URL: https://amslaurea.unibo.it/id/eprint/25147/.

[22] J. Shen, J. Mire, H. Park, C. Breazeal, M. Sap, Heart-felt narratives: Tracing empathy and narrative style in personal stories with llms, in: Y. Al-Onaizan, M. Bansal, Y. Chen (Eds.), Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, EMNLP 2024, Miami, FL, USA, November 12-16, 2024, Association for Computational Linguistics, 2024, pp. 1026–1046. URL: https://doi.org/10.18653/v1/2024.emnlp-main.59. doi:10.18653/V1/2024.EMNLP-MAIN.59.