

# 北京邮电大学 本科毕业设计（论文）任务书

## Project Specification Form

### Part 2 - Student

学院 School	International School	专业 Programme	Electronic Information Engineering		
姓 Family name	Zhang	名 First Name	Ziquan		
BUPT 学号 BUPT number	2022213501	QM 学号 QM number	221170788	班级 Class	2022215117
论文题目 Project Title	A Study on Story Model Architectures for RPGs Using Large Language Models				
论文概述 Project outline  Write about 500-800 words  Please refer to Project Student Handbook section 3.2	<p>This project aims to design and implement an adaptive role-playing game (RPG) story generation system driven by large language models (LLMs). Traditional RPGs depend on fixed dialogue trees and pre-scripted storylines to tell story, which limit player agency and narrative diversity. With the emergence of powerful generative language models, it is now possible to create interactive story systems that dynamically generate coherent narratives and dialogues in response to player input. The project will explore how LLMs combined with Retrieval-Augmented Generation (RAG) and prompt engineering can be used to build a prototype capable of real-time, context-aware storytelling[1].</p> <p><b>1. Problem Analysis and User Requirements</b></p> <p>Current narrative-driven games offer limited flexibility because all story branches must be manually written. Once players have explored the existing dialogue options, the sense of novelty quickly diminishes. The key requirement of modern RPG players is adaptive interaction—they expect NPCs to respond naturally to any input while maintaining logical consistency and emotional realism[2]. Therefore, the system must:</p> <ul style="list-style-type: none"> <li>• Interpret free-form natural language input from players.</li> <li>• Generate coherent, lore-consistent story continuations.</li> <li>• Maintain NPC memory and persona consistency across interactions.</li> <li>• Operate efficiently enough for real-time response.</li> </ul> <p>The proposed solution integrates natural language understanding, retrieval-based grounding, and prompt-controlled text generation to address these needs[4].</p> <p><b>2. Methodologies and Algorithms</b></p> <p>The system is fundamentally built upon a Large Language Model (LLM), which serves as the core narrative generator. All other components act as enhancement layers that improve the model's ability in storytelling in an RPG setting.</p> <ul style="list-style-type: none"> <li>• <b>Large Language Model (LLM)</b> – The LLM forms the foundation of the narrative system. It is responsible for generating story events, dialogues, descriptions, and character behavior based on player input.</li> </ul>				

	<p>API-based models such as GPT-5 or DeepSeek will be used to ensure fluency, contextual coherence, and adaptability.</p> <ul style="list-style-type: none"> <li>• <b>Retrieval-Augmented Generation (RAG)</b> – RAG functions as a support mechanism that supplies the model with relevant game-world information. A vector store will maintain lore fragments, past interactions, and character metadata. Retrieved context is fed into the LLM to reinforce narrative grounding and help mitigate hallucinations, ultimately enhancing the model’s storytelling ability.[1]</li> <li>• <b>Prompt Engineering</b> – Prompt engineering provides structured guidance that shapes how the model generates narrative content. This includes defining NPC personas, tone, narrative roles, and dialogue constraints. Well-designed prompts help the LLM maintain character consistency and produce responses that align with the evolving story.[1]</li> </ul> <p><b>3. System Interaction and Implementation</b></p> <p>The system will operate as an interactive, text-based RPG simulator in which players engage with the game world through natural language input. A user-facing interface will provide a clear and responsive environment for entering player commands, displaying NPC dialogue, and presenting the unfolding story. [5]This interface aims to create an intuitive and immersive interaction experience, allowing players to influence the narrative in real time.</p> <p>Each non-playable character (NPC) will be equipped with a lightweight internal memory model that tracks recent interactions, emotional states, and player-related events[2]. This enables NPCs to react in a contextually appropriate manner, maintain consistent personalities, and reflect changes in the story world over time. Together, these components form a sustainable interaction loop that supports adaptive storytelling throughout the gameplay experience.[3]</p> <p><b>4. Experiments and Evaluation</b></p> <p>To verify the effectiveness of the approach, several experiments will be conducted:</p> <ul style="list-style-type: none"> <li>• <b>Narrative Coherence Test:</b> Human participants will interact with the system and rate the generated story in terms of narrative flow, emotional tone, and overall coherence. This helps assess whether the system can maintain a consistent storyline across multiple interaction turns.[6]</li> <li>• <b>Latency and Efficiency:</b> The system’s response time will be measured under different settings, including variations in model size and context length. Retrieval costs will also be evaluated to ensure that the prototype can support smooth, real-time interaction.</li> <li>• <b>Qualitative Analysis:</b> User feedback will be collected through short questionnaires, focusing on narrative adaptability, emotional engagement, and perceived immersion.[6]</li> </ul> <p><b>5. Implementation Environment</b></p> <ul style="list-style-type: none"> <li>• Language: Python 3.9+</li> <li>• Databases: For vector retrieval or else RAG</li> <li>• Models: GPT-5 (API) or DeepSeek-R1(API)</li> <li>• Hardware: Laptop</li> </ul>
--	---

	<ul style="list-style-type: none"> <li>• Version Control: Git / GitHub for source and experiment management</li> </ul> <p><b>References</b></p> <p>[1]. Mulyana, A., Wibisono, Y., &amp; Anisyah, A. (2025). Non-Playable Characters Based On Large Language Models For Role Playing Games (RPG). <i>Brilliance: Research of Artificial Intelligence</i>, vol. 5, no. 2, 785–792, 2025.</p> <p>[2]. S. Buongiorno, L. Klinkert, Z. Zhuang, T. Chawla, and C. Clark, “PANGeA: Procedural artificial narrative using generative AI for turn-based, role-playing video games,” <i>Proc. AAAI Conf. Artif. Intell. Interact. Digit. Entertain</i>, vol. 20, no. 1, pp. 156–166, 2024.</p> <p>[3]. E. Rimer, R. Ploug, A. K. S. Petersen, and M. Scirea, “Talking to NPCs: Three LLM-driven approaches to dynamic RPG dialogue,” in <i>Proc. IEEE Conf. Games (CoG)</i>, 2025.</p> <p>[4]. A. van der Torre, “Large language models and narrative storytelling in video games,” <i>ResearchGate</i>, 2025. doi: 10.13140/RG.2.2.17147.02086.</p> <p>[5]. J. Leandro <i>et al.</i>, “GENEVA: GENERating and Visualizing branching narratives using LLMs,” 2024 IEEE Conference on Games (CoG), Milan, Italy, 2024, pp. 1-5.</p> <p>[6]. Maisto, Alessandro. (2025). Collaborative Storytelling and LLM: A Linguistic Analysis of Automatically-Generated Role-Playing Game Sessions. 10.48550/arXiv.2503.20623.</p>
道德规范 <b>Ethics</b>	Please confirm by checking the box:  <input checked="" type="checkbox"/> I confirm that I have discussed ethical issues with my supervisor.

<p><b>Please discuss ethical issues with your supervisor.</b></p> <p><b>Please refer to Project Student Handbook section 4.1 and Project Ethics Guidance</b></p>	<p>Summary of ethical issues:</p> <p>This project involves the development of an AI-based narrative generation system using large language models. The main ethical considerations concern the responsible use of artificial intelligence, data handling, and academic integrity.</p> <p>No personal or sensitive human data will be collected; only voluntary user feedback may be gathered during prototype evaluation, and all participants will be informed about the study's purpose and their right to withdraw at any time.</p> <p>From an AI-ethics perspective, care will be taken to ensure data privacy, transparency of model behavior, and mitigation of potential algorithmic bias in generated narratives. The system will be explicitly designed to avoid producing content that references real individuals, political topics, sexual material, or any other inappropriate or harmful subject matter. Filtering and validation mechanisms will be applied to prevent the generation of offensive or unsafe outputs.</p> <p>All development and writing will be original, with proper citation of existing work, in full compliance with university policies on plagiarism and academic honesty.</p>
<p><b>中期目标</b> <b>Mid-term target.</b></p> <p><b>It must be tangible outcomes, E.g. software, hardware or simulation.</b></p> <p><b>It will be assessed at the mid-term oral.</b></p>	<p><b>Mid-term Target</b></p> <p>By the mid-term assessment, the project will deliver a working prototype of the adaptive RPG storytelling system that demonstrates the core functionality of dynamic narrative generation.</p> <p>Specifically, the mid-term outcomes will include:</p> <ol style="list-style-type: none"> <li>1. A basic interactive simulator where players can input text and receive AI-generated narrative responses.</li> <li>2. An implemented narrative generation pipeline, integrating LLM-based story creation with retrieval and prompt mechanisms for maintaining context and coherence.</li> <li>3. A short technical report and presentation summarizing system design, workflow, and initial findings for the mid-term oral examination.</li> </ol>

## Work Plan (Gantt Chart)

Fill in the sub-tasks and insert a letter X in the cells to show the extent of each task

	Nov 1-15	Nov 16-30	Dec 1-15	Dec 16-31	Jan 1-15	Jan 16-31	Feb 1-15	Feb 16-28	Mar 1-15	Mar 16-31	Apr 1-15	Apr 16-30
<b>Task 1 [Conduct a literature review on narrative generation and RPG storytelling frameworks.]</b>												
Review academic research on LLM-based narrative generation and dynamic storytelling.		X	X									
Identify technical challenges and gaps in adaptive narrative generation.			X									
Draft the literature review and establish research direction for the project			X	X								
<b>Task 2 [Design and implement a prototype story model using a large language model for dynamic narrative generation.]</b>												
Design system architecture for the adaptive RPG storytelling prototype.				X	X							
Implement the narrative generation pipeline with context retrieval and response generation.					X	X	X					
Integrate the model with a simple text-based interaction interface.						X	X	X				
Test the prototype for narrative consistency and basic functionality.							X	X				
<b>Task 3 [Develop prompt engineering strategies to maintain story coherence and incorporate player choices.]</b>												
Design structured prompts to define character roles and tone.					X	X						
Implement context-sensitive prompt templates for adaptive player interaction.					X	X	X					
Conduct iterative refinement to improve coherence and personality consistency.						X	X					
<b>Task 4 [Evaluate the system through qualitative user testing focusing on story quality and player engagement.]</b>												

Design evaluation metrics (story coherence, immersion, engagement).								X	X			
Recruit participants or simulate test sessions for user evaluation.									X	X		
Collect and analyze qualitative and quantitative feedback.									X	X	X	
Summarize results and propose improvements and future work.											X	X