

Meet *Daiyu* in The Era of GenAI: Exploring the Use of LLMs As Design Material for Remediating Classics into Interactive Digital Narratives

¹Anonymous Author(s)

Anonymous Institute

Abstract. In this paper, we explore the integration of large language models (LLMs) in interactive digital narratives (IDNs) in the context of remediating the classic Chinese novel, *Hong Lou Meng* (HLM). Adopting Research through Design (RtD), we developed an LLM-driven IDN prototype that allows players to interact with Lin Daiyu, the iconic female character of HLM. We then playtested the prototype with seven HLM readers and five IDN designers through semi-structured interviews. Our findings reveal the potential and challenges of using LLMs as character and plot design materials for IDNs that remediate classic literature. We conclude by proposing two design strategies for LLM-driven IDNs that ensure both player engagement and narrative coherence: 1) Designing towards engaging AI and humans in the same task; 2) Expanding AI responses beyond predictable boundaries.

Keywords: interactive digital narrative · large language model · design material · remediation · cultural heritage.

1 INTRODUCTION

Interactive digital narratives (IDNs) represent a narrative form that employs digital technologies to enable players to experience and enact a storyline through character embodiment and world exploration [13,27]. One major form of IDNs is narrative-based videogames. For example, in *The Witcher 3: Wild Hunt*¹, players role-play as a monster hunter navigating a fantasy world in search of his adopted daughter. Their choices throughout the experience meaningfully alter story outcomes and character relationships, creating unique narrative paths that invite multiple playthroughs.

The recent boost of large language models (LLMs) and generative AI (GenAI) is revolutionizing IDNs by enabling real-time narrative generation that responds dynamically to player actions [35,20]. This technological shift allows IDNs to evolve into a highly personalized experience that was previously unattainable. Contemporary IDN platforms such as *AI Dungeon*² utilize LLMs to generate stories based on players' text prompts, opening up infinite plot possibilities.

¹ https://en.wikipedia.org/wiki/The_Witcher_3:_Wild_Hunt

² <https://aidungeon.com/>

However, the current LLM systems, while impressive in many regards, still struggle to create convincing interactive narratives [20]. The player community of *AI Dungeon* often discusses that off-topic input may derail the storyline and disrupt the coherence of the whole story [31]. On the other hand, limiting the scope of player input risks reducing interactivity and narrative diversity. This trade-off highlights the need to balance player autonomy and story quality when integrating LLMs in IDNs [31]. In this sense, LLMs are far from a tool that can autonomously provide interactive stories that are sophisticated enough, but instead, design material that HCI researchers can utilize to explore new design possibilities for application domains such as IDNs.

Hong Lou Meng (HLM), also known as *Dream of the Red Chamber* and *The Story of the Stone*, is a classic Chinese novel written by Cao Xueqin in the 18th century. It follows the rise and fall of the Jia family, focusing primarily on the tragic love story between Jia Baoyu, a sensitive young man, and his cousin, Lin Daiyu, who is deeply introspective and melancholic. The story is known for its poetic language, rich characterization, and intricate plots, making it good fodder for “remediation,” the adaptation of a work from one medium to another while retaining its original core. Its incompleteness, as the last 40 chapters were not written by Cao, even adds to more potential for secondary creation. So far, HLM has been remediated into various media, including television series (e.g., the renowned 1987 television series³), stage productions, and videogames (e.g., *Stone Story* on Steam⁴). With the technical capabilities of LLMs, HLM has the opportunity to be remediated into an IDN form with real-time adaptation that adjusts dialogue, plot events, and world states based on player interactions.

This leads to our research questions: *What are the affordances and challenges of using LLMs in the narrative remediation of HLM?* (RQ1) and *How can LLMs be used effectively for narrative design within the context of the remediation of classic novels?* (RQ2)

We introduce an interactive prototype that integrates LLMs for players to converse and interact with Lin Daiyu. By adopting Research through Design (RtD) as the guiding methodology, we explored the unique affordances and challenges of LLMs as design material in providing real-time adapted narratives. RtD is a “rapid response methodology” that informs insights concerning evolving disruptive technologies [6], which in our case, is LLMs. Using the prototype as a conversation starter, we conducted semi-structured interviews with seven HLM readers and five professional IDN designers to explore their experience of narrative remediation integrated with LLMs, as well as the design possibilities and constraints that LLMs open up for IDNs.

Our research contribution is two-fold: 1) We presented an LLM-integrated IDN approach that allows players to converse with Daiyu, whose responses are real-time generated and personalized to players’ inputs (as opposed to pre-scripted). It is demonstrated with an online prototype playable in a browser. 2) We employed RtD as the overarching methodology and conducted a user

³ [https://en.wikipedia.org/wiki/Dream_of_the_Red_Chamber_\(1987_TV_series\)](https://en.wikipedia.org/wiki/Dream_of_the_Red_Chamber_(1987_TV_series))

⁴ https://store.steampowered.com/app/2761570/_Stone_Story/

study to inform design implications on utilizing LLMs as material for character and plot design in narrative remediation. Specifically, concerning the “authoring problem” [32] in IDNs, we propose future designs to engage AI and humans in the same task and expand AI responses beyond what is predictable.

2 RELATED WORK

In this section, we clarify several concepts that are essential in our work. We first introduce the definition of remediation and how narrative relates to it. Next, we explain the rationale for choosing interactive digital narratives (IDNs) as the specific narrative method and *Daiyu* as the character to narrate the story. We propose integrating LLMs in the process can mitigate the prevalent “authoring problem” in IDNs. Last, we elucidate the notion of design material in HCI, illustrating how narrative remediation helps inform knowledge of using LLMs as design material.

2.1 Remediation

According to Marshall McLuhan [24], media stands for the various technologies, including print, television, and digital platforms like websites and social media, that mediate our communication. They are not “hollow pipes” that just carry messages. Instead, as McLuhan famously argues, “the medium is the message”—the specific properties of a medium also influence our perception and understanding of the world.

For Jay Bolter and Richard Grusin, media is also defined by the process of “remediation”: “What is a medium? We offer this simple definition: a medium is that which remediates. It is that which appropriates the techniques, forms, and social significance of other media and attempts to rival or refashion them in the name of the real.” In other words, remediation refers to the process by which new media represent (i.e., absorb, refashion, and rework) old media, and together they constitute the “media ecology” of their time. Today, the topic of how digital media remediate traditional media at various levels is foregrounded, sparking discussions across various disciplines.

In narrative studies, remediation also serves as a particularly useful concept in framing questions within the field of transmedial narratology [28]. According to Ryan’s framework of transmedia narratology [28], our research aligns with a form of narrative remediation that “transposes from one medium to another,” as exemplified by adaptations of novels into films, novelizations of movies or computer games, and computer games originating from literary works, etc.

2.2 Interactive Digital Narratives & Character

Narrative can be considered to operate on two basic levels: story and discourse. Story refers to the sequence of events, characters, and settings—the “what” of the narrative, while discourse, moreover, concerns the way the story is told—the

“how” [12,3,25]. Traditional storytelling forms, such as novels and films, offer linear experiences through predetermined sequences. In contrast, interactive digital narratives (IDNs) allow players to actively participate in the storytelling process, making choices that influence the direction and outcome of the story. Transforming a traditional linear story into a digital interactive one has become an increasingly popular form of narrative remediation in recent years, as evidenced by the widespread success of video games like *Black Myth: Wukong* (2024), *Middle-earth* series (2014-2017), and *80 Days* (2014), to name just a few.

A story element that undergoes significant transformation in this kind of narrative remediation is “character.” While literary text-based narratological discussions conventionally frame character as a textual construct and structural component of the narrative [4,11,3,26], in IDNs, character is made to dynamically adapt and react to the player’s choices and actions [2]. As Aarseth [1] argues, “the most effective way of creating ludo-narrative content is to invest in character-creation, by making the characters rich, deep and interesting.” A corollary of this argument would be that character becomes a largely foregrounded narrative element in IDNs, where the player cares less about the predefined plotlines, but more about how they themselves can enact the story through interacting with or embodying an interesting character.

Today, GenAI has been used to create IDNs [20]. However, we are still facing the prevalent “authoring problem”—the challenge of developing interactive narratives that grant player agency while maintaining content coherent and engaging [32]. The remediation of a classic can even complicate the “authoring problem”. The player’s prior knowledge of the source material may lead to expectations about character behavior, plot development, or thematic consistency, thus making the design of meaningful and engaging interactions more difficult. The use of LLMs, as this research proposes, may offer a potential solution to the authoring problem by enabling responsive, real-time content generation that remains largely coherent with the original work, while still allowing players creative space to explore and interact. To demonstrate this, our research aims to create an interactive narrative that remediates Hong Lou Meng by focusing on crafting a digital consciousness of Lin Daiyu, the female protagonist and the soul of the novel, within our story, supported by LLMs.

2.3 Technology as Design Material in HCI

In Human-Computer Interaction (HCI), design material refers to the elements and tools that designers use to create interactive systems, similar to how physical materials are utilized in traditional design fields like architecture or fashion. These materials can be tangible, such as sensors and screens, or intangible, including code, algorithms, data, and interaction patterns [17,18]. By framing technology as design material, we refer to its design innovation or design contribution [33]. Different from engineers who create technical innovation by inventing new technology that allows new capabilities, designers assemble or reassemble known technologies for novel user scenarios [33,23]. In this process, designers “converse with existing materials” [33,34] to envision things that haven’t existed before,

and the materials “talk back to designers” [33,34] on how the technology opens up design possibilities as well as constraints [29]. Recent HCI research has started to explore the use of AI [18,34] and machine learning (ML) [14,33,5] as design material. Interestingly, all these studies point out that AI (or ML) has unique capabilities to create novel applications and services that no other technology has ever achieved. LLMs are now sophisticated enough to understand, generate, and manipulate human language in real-time, creating unlimited yet unknown design possibilities for IDNs.

The notion of design material aligns closely with the concept of remediation. Our research aims to remediate the Chinese classic *Hong Lou Meng* through IDNs. To achieve that, we designed a dialogue game where we implement a conversational agent remediating from the representative female character in HLM, LIN Daiyu, to tell the narrative. The conversational agent, supported by LLMs, is capable of responding to the player’s input in real-time, and what the player chats to the agent decides the story outcome. In our remediation of HLM, we seek to use the design of an LLM-driven dialogue game as a conversation starter, to inform knowledge about utilizing LLMs as design material for IDNs.

3 METHODOLOGY

In this section, we first justify our choice of Research through Design (RtD) as the leading methodology. We detail our design objective and process, as well as the features of our prototype. The prototype is used as a discussion starter in semi-structured interviews with HLM readers and IDN designers to answer our research questions.

3.1 Research through Design

Research through Design (RtD) is an approach of inquiry in which knowledge is generated through the design process [15,30]. Researchers probe “what the world could and should be” [30] by creating artifacts, such as products, systems, or services. These artifacts serve as both outcomes and tools for understanding player experiences, cultural contexts, and theoretical concepts. RtD is especially useful in rapidly responding to emerging technologies [6], LLMs in our case, and informing future work. We could have adopted other traditional HCI research methods that focus on observation or analysis, for example, comparing our design with other prototypes with similar design goals, such as the digital resurrection of Mary Sibley by Hutson et al. [19]. However, our research is a design-led inquiry, whose primary goal is to produce knowledge on using LLMs as design material for narrative remediation, as opposed to verifying our prototype’s usability.

Our RtD study is structured as follows: in the design and implementation phase, we constantly iterated our design idea while we implemented the prototype using an agile approach; then, during the user study, we used the prototype as a conversation starter [22] in semistructured interviews to gain insights from two different groups, namely HLM readers and IDS designers.

3.2 Prototype Design

Design Objective This research aims to create an application where players can interact with the digital character (i.e., Lin Daiyu) and narrative elements in a way that would not be possible without using LLMs. This system showcases how classical texts like HLM can be brought to life in a new, interactive medium, preserving the depth and nuance of the original story by integrating LLMs.

Design Process The initial idea for the system came from one of the author’s PhD project (forthcoming), which focused on remediating HLM as an Augmented Reality book. Our team conducted an extensive literature review and a survey of how HLM had been remediated in other media. We then conducted a series of informal ideation workshops and decided to use LLM-powered IDNs for remediation and selected one overarching theme, the Buddhist conception of reality as illusion and illusion as reality as handled in the original novel, and three major features: focus on interaction with Daiyu as a dynamic character [8]; include prominent social rituals such as gift giving, poetry recitation, and riddles as game mechanics; and provide clear narrative progression and closure.

The design process then followed an iterative, exploratory approach centered on translating the selected theme and features into interactive affordances. Early ideas were sketched through Figma-based semi-interactive prototypes to demonstrate interaction flows. Besides, drawing from Blom [8], we implemented a preliminary system of affection to track dimensions such as trust, delight, empathy, and awareness, influencing both the surface style of character responses and the narrative possibilities available to the player. Parallel to these developments, we experimented with poetry composition and riddle challenges as turn-based mini-games embedded within dialogue sequences. To test the overall flow and coherence of the narrative experience, we developed a Twine⁵ prototype incorporating key interaction sequences and multiple branching arcs. These iterative and parallel prototyping stages informed one another, helping refine design priorities and revealing how the system might dynamically support both emergent play and structured narrative arcs.

Implementation Process The interactive system follows an Agile development process, using iterative cycles of design, implementation, and evaluation. The core architecture is built with Flask⁶, a lightweight Python web framework that supports a modular structure. This structure separates key components such as conversation handling, stage transitions, poetry generation, riddle interaction, and character introspection. The system uses Flask-Session for secure state tracking and GPT-4o⁷ to generate dynamic, context-aware dialogues. We fine-tuned the GPT-4o model using the English version of the entire HLM novel.

⁵ <https://twinery.org/>

⁶ [https://en.wikipedia.org/wiki/Flask_\(web_framework\)](https://en.wikipedia.org/wiki/Flask_(web_framework))

⁷ <https://openai.com/index/hello-gpt-4o/>

System Mechanism This section introduces our prototype from two dimensions: the storyline and functions. Referring to the concept from game design, the storyline is the framework that guides the player’s narrative experience throughout the interactions within the system, while the functions encompass the interactions that players can perform on it.

Storyline. The narrative consists of four stages, each with its own emotional theme for *Daiyu*. Every stage has various pre-scripted endings, which are achieved based on what the player chats with *Daiyu*. In the first stage, *Daiyu* is in a mood of *peace* and the conversation between her and the player is mainly greetings and self-introductions. At that time, *Daiyu* considers herself a real human and believes that the world of *Dream of the Red Chamber* is a real, existing world where she lives. The second stage is about *doubt*, where *Daiyu* discovers that something has gone wrong and questions the player’s world. In this stage, the player is expected to tell *Daiyu* more about the real world outside the novel or prove that they know everything about her and the other characters in HLM. In the third stage, *Daiyu* realizes that she is fictional but *struggles* to accept this truth. She wonders (in a feeling of *crisis*) who, where, and even what she is. In this stage, the player is expected to engage in a more reflective discussion with her on existential questions such as “What is life?”, “Can characters have life?”, “What is the real world?”, etc. In the fourth stage, the conversation is more open-ended. *Daiyu* *accepts* that she is just a fictional character and *clams down*. Before “a final goodbye,” she asks the player if they have anything else to say.

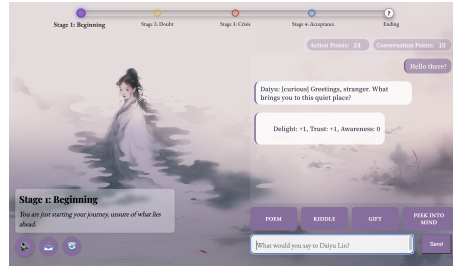


Fig. 1: The main conversation interface of the prototype.

Functions. Our prototype is mainly a conversation app where players put text prompts into the chat box, and *Daiyu* replies accordingly (see Figure 1). What *Daiyu* says to the player is all generated by GenAI in real-time. We train the LLM (GPT-4o) to imitate the way *Daiyu* thinks and speaks in the original novel. There are four functions, namely “Poem,” “Riddle,” “Gift,” and “Peek into Mind”, designed to support the progression of the storyline. Details are explained as follows:

1. *Poem*. The player clicks the *POEM* button, and *Daiyu* will propose a keyword and invite the player to compose a poem. A panel pops up, showing

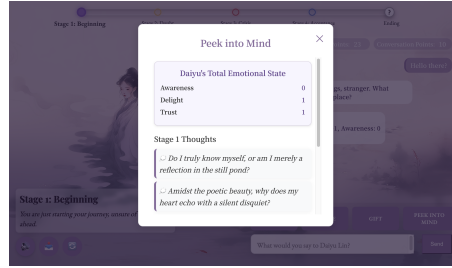


Fig. 5: The player clicks the *PEEK INTO MIND* button. The pop-up panel shows what Daiyu is thinking and the value of her current dynamic traits.

point; composing a poem costs three; answering a riddle costs two; giving a gift costs one; and peeking into the mind also costs one.

2. *Conversation points* serve as the currency in our prototype. Players can earn them by eliciting as many affective responses from the Daiyu as possible during the conversation, and then use those points to purchase gifts.
3. *Dynamic traits* are aspects of Daiyu’s personality that can be influenced by the player during the conversation. They directly shape the progression of the interactive story. There are three traits: awareness, delight, and trust. Awareness reflects the extent to which Daiyu realizes she is a fictional character as well as an LLM-driven consciousness; delight represents how happy she feels during the interaction; and trust indicates how much she sees the player as a friend. Among these, awareness determines whether the player can advance to the next stage and is directly tied to the story’s progression bar. Delight and trust, meanwhile, determine which ending the player reaches at each stage. All three traits can be modified through good conversation quality, successful riddle challenges, and positive reviews of the composed poems, and giving appropriate gifts.
4. *Stable traits* are aspects of Daiyu’s personality that cannot be influenced by the player during the conversation, yet they influence how soon the dynamic traits can be added by the player. These traits (temperament, curiosity, playfulness, taste) are means to control the progress of relationship development, where each trait acts as a multiplier that determines how effectively player actions (dialogue exchanges, poem composition, riddle challenges, gift giving, and peeking into mind) can increase the dynamic attributes of awareness, delight, and trust. For instance, a higher temperament (70) and curiosity (60) make it easier to build “trust” through interactions, while higher playfulness (40) and taste (80) enhance the effectiveness of actions that increase “delight”, and curiosity (60) directly influences how quickly “awareness” can be developed, creating a consistent personality baseline that shapes all player interactions and relationship progression. The design of stable traits is also useful to portray different personalities of the “Twelve Beauties of Jinling” for future work (see Section 6).

4 USER STUDY

4.1 Motivation & Participants

We conducted a user study with two different groups—HLM readers and IDN designers—to answer our research questions. We targeted HLM readers for three reasons: first, the whole playthrough of our prototype requires knowledge of HLM more than just having heard of it; second, we think people who are familiar with HLM can offer more critical opinions on remediation than HLM novices; third, we want to understand players’ experiences, acceptance, and other thoughts regarding remediating HLM with LLMs. Through recruiting IDN designers, we aimed to explore the possibilities that LLMs open up for IDNs, as well as the constraints or limitations from the design perspective. Therefore, we did not require in-depth knowledge of HLM from the recruited IDN designers.

We recruited 12 participants in total, including seven HLM readers (two of whom were cultural heritage researchers) and five IDN designers. The demographic details are listed in Table 1.

Table 1: Demographics of the participants

No.	Gender	Familiarity with HLM	Identity
P1	Male	heard of HLM	IDN designer
P2	Female	read the original novel & watched HLM TV adaptations	HLM reader
P3	Female	read the original novel & watched HLM TV adaptations	HLM reader
P4	Female	watched HLM TV adaptations & liked HLM a lot	HLM enthusiast
P5	Male	read the original novel & watched HLM TV adaptations	HLM enthusiast
P6	Female	read the original novel	IDN designer & lecturer
P7	Male	never heard of HLM	IDN designer & researcher
P8	Male	partially read the original novel	HLM reader
P9	Male	heard of HLM	IDN designer
P10	Female	partially read the original novel	IDN designer for games
P11	Female	watched HLM TV adaptations & knew the plots	HLM reader & cultural heritage researcher
P12	Female	read the original novel	HLM reader & cultural heritage researcher

4.2 Semi-structured Interviews

The study started with an introduction to our research and prototype, after which the participants completed a playthrough of the prototype under our guidance. We then conducted interviews in a semi-structured way, which allows

us to seek clarifications, add questions, and follow comments for nuanced insights [16,21]. We conducted interviews in Chinese or English, depending on the participants’ preference. All interviews were recorded via Zoom⁸ with the participants’ consent. The interview scripts were designed slightly different for the two player groups.

4.3 Data Collection & Thematic Analysis

We used Feishu Minutes⁹ to transcribe the recorded videos into text automatically. Following this, we performed manual checks and content calibration to ensure that the transcribed text matched the original interview content. Subsequently, we conducted a thematic analysis (TA) to examine those text transcripts.

Thematic analysis (TA), as defined by Braun and Clarke [9], serves as an interpretative framework to uncover patterns of meaning—commonly referred to as “themes”—within a qualitative dataset. These themes convey the narratives behind the data. Notably, there are no uniform or definitive themes for a particular dataset; themes vary due to the researcher’s theoretical insights and analytical capabilities [10]. The first author conducted an inductive TA on the interview data, employing open coding rather than adhering to a predefined codebook. Adopting the six-phase TA outlined by Braun and Clarke [9], the first author commenced by thoroughly reading the transcripts to gain a solid understanding of the entire dataset. Second, 16 preliminary codes were generated, which represented the first author’s preliminary interpretations of the original data. Third, the researcher reviewed these codes and clustered similar ones around the research questions, yielding three potential themes. In the fourth phase, the first and second authors scrutinized these candidate themes against both the codes and the entire dataset to ensure that they encapsulated a coherent narrative. The final two phases involve defining, naming, and documenting these themes, which are detailed in Section 5.3.

5 RESULTS

In this section, we present three themes to answer the proposed research questions. To respond to *RQ1*, we start with the characteristics of LLMs and answer it from the perspective of characterization and storytelling. In response to *RQ2*, we identify two design implications for future remediation works using LLMs.

5.1 LLMs as IDN Character Design Material (in Narrative Remediation)

We received mixed reactions towards our LLM-driven Daiyu prototype. Participants’ responses appeared to correlate with their familiarity with, and devotion

⁸ <https://www.zoom.com/>

⁹ <https://www.feishu.cn/hc/en-US/articles/022111234449-get-the-most-out-of-feishu-minutes>

to, the original HLM. On the one hand, two participants who were HLM enthusiasts expressed reservations about our prototype. P5 observed: *“In the novel, Daiyu wouldn’t readily engage in such friendly banter with strangers, but here she seemed too affable and outgoing.”* P4 directly suggested that the crucial next step for improving the prototype is to *“enhance its literary fidelity to the original novel—specifically, to ensure that Daiyu’s speaking and poetic style more closely mirror the literary characteristics of Daiyu as depicted in the classic.”* On the other hand, participants who were not HLM enthusiasts or had only limited knowledge of the work generally provided positive feedback regarding how well our prototype matched their impression of Daiyu. *“I am surprised that the way Daiyu talked in your prototype is consistent with her sentimental image in the original novel.”* [P2] *“The personality that she showed to me is quite close to Daiyu in Hong Lou Meng. When I talked to her, I could sense her hesitation derived from her personality.”* [P12]

Regardless of their familiarity with HLM or satisfaction with our prototype, most participants reported being able to detect AI characteristics in Daiyu’s responses. P6 provided a particularly illustrative example: *“When I asked her, ‘Do you feel sad for not knowing your future life?’ and she answered, ‘I cannot feel sadness for what I do not know.’ This answer reminded me of her AI-ness.”* Many participants also picked up on Daiyu’s AI nature through her constant availability to respond. This created a double-edged effect: while the immediate, tailored responses made the character feel more *“real”* [P3], her relentlessly passive role gradually caused some participants to feel an uncomfortable one-sidedness in their conversations. As P9 put it: *“She always listened to what I said and showed respect for my words, but she didn’t start new topics with me. Sometimes, she would ask me questions, but they were based on the context that I mentioned to her. She was not proactive, and that made me feel like talking with an AI, instead of the character from the novel.”*

To a large extent, the detectable and hard-to-eliminate AI characteristics were responsible for the unfavorable responses to our prototype. Based on this, P9 further highlighted *“the possible dilemma of character analogs”* in LLM-powered IDNs, as *“I don’t expect to see characters featured in different classics all saying things like ‘How can I help you?’”* That said, there were also some alternative perspectives that found value in how we used AI to construct Daiyu’s indecisive character: *“My impression of Daiyu is that she tends to be indecisive, while AI tends to give answers with repetitive meanings. Maybe that’s why I feel that the Daiyu in your prototype is believable.”* [P10] This suggests that AI can share certain commonalities with human personality, which are worth exploring and leveraging when using LLMs to portray human characters.

5.2 LLMs As IDN Plot Design Material (in Narrative Remediation)

By utilizing LLMs, we make it possible for the player to intervene and influence the plots while keeping them relevant to the player’s input. As several participants reported, they felt an enhanced sense of interaction and engagement due to the perspective shift enabled by LLMs. When reading the novel or watching the

TV adaptations, they were in the third-person perspective. When chatting with Daiyu, however, they were in the second-person perspective. *“I felt like a time traveller, stepping into Hong Lou Meng. I got infinite possibilities [of Daiyu’s answers] due to AI, and it gave me the opportunity to bring about change for her, which made me immersed.”* [P8] Another evidence is that most participants felt in control of the whole narrative, but none thought they had absolute control over it, which in turn made them feel engaged in the narrative. *“If ten is total control, I’d give it a six. In your context, although I could say anything to Daiyu, I still had a task [to persuade Daiyu that she is fictional] to complete, and I was constrained by the action points. But that was good. It made me more focused on what I should say.”* [P1]

Besides, with the integration of LLMs, there will be no two identical narrative experiences, but experiences personalized to each player’s input. Most of our participants were aware of this point, like what P6 said, *“I have a really big variety of ways of seeing how Daiyu responds to me. Even though I asked the same question or played the game again, the responses were different. I feel I’m into it [the narrative].”* This merit of LLMs makes the narrative experiences distinct from pre-scripted ones and can encourage players to engage more thoughtfully with the story. *“Personally, pre-scripted narratives or narratives with a tree structure are like a multiple-choice question to me. I would have a trial-and-error mentality when I experience pre-scripted narratives. But in your game, I must be careful since I cannot let things start over again as they were.”* [P1]

Nevertheless, there were still criticisms of the narrative design, arguing that the narrative based on the passive, responsive AI weakens the engagement. *“The current storyline is that I destroy Daiyu’s worldview and help her rebuild a new one according to my worldview. However, I was expecting that Daiyu, or the AI, could, at least try to, persuade me and change my viewpoint. If that happens, I’ll see her as more human and get more engaged.”* [P12]

5.3 Takeaways for Future IDN Design Using LLMs

Players are capable of prompting any content to the LLM. However, it seems that this right can distract them from the narrative and thus undermine the narrative coherence. We summarize two insights on approaching the “authoring problem” [32] in IDNs in the context where players co-create the narrative with LLMs. Specifically, the two insights help ensure that players’ inputs to LLMs lead to meaningful and engaging experiences while maintaining a consistent storyline.

Designing towards engaging AI and humans in the same task. Many participants mentioned that our storyline was designed around the task—resolving the worldview conflict between Daiyu and them. This conflict challenges them while also encouraging meaningful interactions between them and Daiyu. *“I think the conflict between the character [Daiyu] and me interests me a lot.”* [P3] *“I think it [the prototype] is about the conflict between humans and AI, that is, how to convince an AI that it is an AI.”* [P8] In situations where there is a conflict to

be resolved, especially between the players and the LLM, they seem unlikely to talk to the LLM about irrelevant matters. Therefore, we propose to frame the narrative around collaborative problem-solving so as to create a shared goal that aligns the interests of both the player and the AI. This approach allows the player to feel a sense of agency while simultaneously fostering a deeper connection with the AI character.

However, it cannot be guaranteed that the player will always focus on the designed task and input relevant content to the LLM. In that case, we should consider how to make them realize when they have deviated from the narrative at an appropriate time. *“Your design of immediate value feedback from Daiyu reminded me of deviating from the storyline. When I saw her awareness increase by zero, I knew I had said something not that important, and that made me focus on thinking about what Daiyu wanted.”* [P1] Thus, providing real-time feedback to players’ inputs can be a solution to guide players back to the core story while preserving their autonomy. This helps maintain the narrative coherence, ensuring that while players explore their creative input, they remain anchored within the framework of the story.

Expanding AI responses beyond predictable boundaries. LLMs appear to answer within a safe area, where human players seem to have preset expectations of what LLMs would respond. This tendency for LLMs to stay within predictable boundaries can limit the richness of player experiences. *“AI constructs sentences by predicting the most likely next word based on the context of the conversation. It uses patterns learned from existing datasets to determine which words and phrases fit best, and that is why I hardly had the experience of being surprised by AI. I expect AI, or Daiyu in your context, could say something that surprises me.”* [P10] *“I think the AI-implemented characters are too safe now. I want to see how Daiyu can get annoyed with me, since I tried to be rude to her by directly saying that she was a character written by somebody, but her response was not angry. Neither did she ever challenge me.”* [P9]

Therefore, it is understandable that even if the player could input anything to the LLM, they can be bothered by constantly getting expected responses and become disengaged. How to get AI out of passive consent and take more proactive stances seems essential. *“AI always responds with a kind and empathetic tone, and maybe that’s why I feel it inactive, although it does interact with me to some extent. I would recommend designing the AI character to be more proactive, for example, let AI, or Daiyu, solve the conflict or convince me of certain viewpoints.”* [P12] This proactive engagement could involve allowing the AI to pose challenges or counterarguments to the player’s input, thus creating a more “player-passive” dialogue. Moreover, pushing the boundary of AI responses can also contribute to this aim. *“Maybe try to reach the outliers of AI responses, for example, to what extent will it be annoyed? And then we might know where to go.”* [P9] We thereby propose to train LLMs to react beyond their usual scope, moving from passive answering to proactively initiating or ending conversations, and from always being friendly and polite to showing a range of emotions, such

as rudeness or anger. This can allow AI characters to be capable of emotional complexity and more authentic.

6 LIMITATIONS & FUTURE WORK

Our research has two noteworthy limitations. First, a potential sampling bias may have occurred during participant recruitment, as individuals who volunteered for interviews might have been inherently more receptive to AI applications in narrative reinterpretation than the general population. Second, the prototype’s English-language interface created barriers to user immersion, as several participants emphasized that Chinese would arouse more cultural resonance in this context. Third, our prototype is naive in guiding players to understand and complete their task. Many participants pointed out the need for a beginner’s guide; otherwise, they would be overwhelmed when they were exposed to a blank chat box.

Looking ahead, we are preparing for a further study with a refined prototype focusing on the cultural acceptance of LLM remediations. Concerning the prototype, we plan to make a Chinese version and improve responses from LLMs, potentially incorporating strategies to make characters more proactive. In addition, we aim to remediate characters more than only Lin Daiyu, but also the “Twelve Beauties of Jinling” using LLMs, with each beauty representing a unique personality and narrative.

7 CONCLUSION

In this research, we demonstrated how LLMs can enhance player engagement and create personalized narrative experiences by designing and developing a prototype, which enables real-time conversations with iconic characters Lin Daiyu from the Chinese classic novel *Hong Lou Meng*. The findings from our user study indicate that while participants appreciated the real-time personalized responses of the character, there is a need for more proactive interaction initiated by LLMs that is contextually coherent. Based on this gap, we offer insights on how to maintain player engagement while ensuring narrative coherence when they are able to input anything to LLMs. On the players’ side, we encourage the design of a task for players and LLMs to complete collaboratively and remind the players when they deviate from the task at an appropriate time. On the LLMs’ side, we propose to train them towards being more active and emotionally diverse.

References

1. Aarseth, E.: A narrative theory of games. In: Proceedings of the International Conference on the Foundations of Digital Games. p. 129–133. FDG ’12, Association for Computing Machinery, New York, NY, USA (2012). <https://doi.org/10.1145/2282338.2282365>