

AI Use in Game Design: Optimizing NPC Interaction While Preserving Narrative Stability

Hau Tak Zeng

{zenghau@msu.edu}

College of Communication Arts and Sciences, Michigan State University, Michigan, United States

Abstract. Artificial intelligence is becoming more important in game design, especially in NPC interaction. LLM-driven NPCs can respond to players in real time and make dialogue more open and flexible than traditional scripted systems. However, this openness also creates design risks, including character inconsistency, narrative deviation, and reduced control over story structure. This paper examines how AI-driven NPC systems change narrative design in games and focuses on one central issue: how to preserve narrative stability while expanding interaction. By comparing scripted NPCs, *Detroit: Become Human*, PANGeA, and evidence from educational and serious games, this paper argues that hybrid design is more practical than unrestricted generation in narrative-heavy games.

Keywords: Artificial intelligence; Game design; NPCs; Large language models; Narrative stability.

1 Introduction

Artificial intelligence now affects game design in a more direct way than before. In older games, AI usually stayed in the background. Developers used it for enemy behavior, pathfinding, and other support systems. Now AI is moving closer to the center of play. It affects dialogue, story interaction, content generation, and the way players talk to the game world. Research on LLM-based narrative systems shows that one key problem is “maintaining narrative coherence” when players can type freely [1].

Generative AI is a big part of this change. On the production side, Alharthi finds that generative AI can “accelerate ideation,” “enhance the prototyping process,” and “automate repetitive or labor-intensive tasks” [2]. At the same time, the same study says that concerns about the “originality of AI-generated outputs” are still strong [2]. So AI does not only make design faster. AI also changes how control and creative responsibility are shared between human designers and computer systems.

This change is even more important in NPC interaction. Liu argues that AIGC is changing game interaction by improving dynamic environment generation, allowing “Intelligent dialogue systems break through script-based interaction,” and supporting more dynamic story forms [3]. This kind of openness may make interaction feel more alive. Still, games are not open chat

spaces. Games need rules, timing, structure, and consistency. PANGeA points out that “out-of-scope player input” can push generation away from the intended story if the system is not controlled [1].

Traditional scripted NPCs may feel limited, but they are stable. Designers know what an NPC can say before the player reaches that moment. Designers can control when information appears, how a scene develops, and what each character is allowed to know. LLM-driven NPCs work in a different way. They generate responses during play. This design can make interaction feel more natural, but it also creates risk. If the system does not control the model well, the result may be character drift, plot confusion, or broken pacing.

This problem matters even more in narrative-heavy games. In these games, dialogue is not just extra text. Dialogue shapes emotion, guides player understanding, and controls dramatic rhythm. For this reason, this paper argues that narrative stability should be treated as the main issue in AI-driven NPC design. This paper first explains the difference between scripted NPC systems and LLM-driven NPC systems. Then it discusses the value and risk of generative dialogue. After that, it compares two models: the controlled branching structure in *Detroit: Become Human* and the constrained generative system in PANGeA. Finally, it argues that AI works best when it operates inside a clear narrative framework instead of running without limits.

2 Concepts and background

2.1 Traditional scripted NPCs

Conventional NPC systems are constructed on ready made structures. Dialogue trees are written by designers, lines are linked to quest states, and it is up to designers to determine when to show which response. In this model, there is the option of limited choices made by players. Each choice results to a pre-written and pre-marked response by the designer.

This design will be appearing less flexible than the free-text interaction, but it possesses some real advantages. It provides the designers with control of the experience. Designers are able to maintain the NPC. The pacing and information flow as well as the order of key events can also be controlled by designers. This control is significant in narrative based games. A game might have to conceal a twist, hold off a response, or make a character explain reasons behind his or her actions at a slow pace. This is simplified by scripted systems since each division is already a part of the bigger picture.

So programmed NPCs are not only archaic technology. They follow a different goal. They are not meant to install the indefinite openness. They are intended to allow players to make meaningful choices within a stable story.

2.2 LLM-Driven NPCs

This pattern is altered by NPCs that are driven by LLM. These non-player characters do not just choose among predetermined alternatives. They produce reactions when the game takes place depending on prompts and the situation and the player. Examples include PANGeA which may create settings, items, events and personality-biased NPCs that accompany free-form dialogue with the player [1]. Another argument presented by Liu is that AIGC is capable of helping to

contribute to the new forms of interaction by providing dynamic content, intelligent dialogue, and more adaptive storytelling [3].

There are two strengths in this design. To begin with, it increases the rigidity of conventional dialogue trees. Second, it has the potential to relieve certain writing pressure as the system can generate numerous potential responses even though a designer is not obliged to write each line manually. Also, Nagarkar mentions that AI-introduced NPC systems should mean building dynamic environments, influencing NPC behaviours, and assisting in the development of stories that respond to player action [4].

Nonetheless, this openness makes the design difficult. There is nothing banal as games in a conversation. Games have rules of the world, roles of characters, tone and priorities of the story. When an NPC is able to say virtually anything, then the designer has to make decisions as well about what the NPC should never say. In the absence of such a restriction, the free-form interaction may harm the form of the game. The model can give out information prematurely. The model can overlook the existing stage of story. The model can also provide an answer that cannot fit the character or the scene yet sounds smooth. When it talks about free-form text it produces content which does not fit within the intended narrative scope, it makes sense [1] as explained by PANGeA.

Then the question of the quality of language is not so simple. It is possible to sound natural in response and to damage the game anyway. Quality in narrative design entails limits. In case AI makes the dialogue more open, design should also provide more powerful systems in such a way that the dialogue is contained in a useful scope.

2.3 Narrative stability: a problem of design

Narrative stability in this paper refers to the ability to maintain the story as it is being interacted upon. The characters must remain the same. The world of the game must have its logicity. The right time should have taken into consideration important information. There must be something meaningful that results of player choices are supposed to produce.

PANGeA is advantageous in this case, since it does not just address this issue as a prompt one, but a system problem. It consists of a framework comprising of a custom memory system and a novel validation system [1]. The cognitive disposition archives the previous events and the circumstances. The validation system compares the player input and output produced with the rules of the gameplay, and the storytelling limits [1]. The validation system also increased performance in the evaluation of the paper. In the reported test cases, both Llama-3 (8B) and GPT-4 increased up to 98 and 99 percent with the addition of a validation, respectively [1].

This point matters. Easy generation will not suffice on its own. And, to make generative NPCs functional in narrative games, the developers must have systems that are supportive of the model. Not only does it rely on the model as far as narrative consistency is concerned. Narrative consistency is also subject to the control structure created around it.

2.4 Generative artificial intelligence and design procedure

The designers also have changed their workflow using AI. The impact of its effect goes even farther than player-facing dialogue. Alharthi demonstrates that AI tools that are based on generative AI, can facilitate ideation, prototyping and repetitive production tasks [2]. This is

one of the reasons why it is helpful to many developers. In the same study, designers are also concerned that originality, over reliance and undermining the human creative work remain issues [2].

So this is not a change only beneficial. AI does not just remove labor. AI transforms the type of work being done by designers. Designers might not have to spend a lot of time on typing out each line individually. However, they might use more time in the formulation of prompts, determination of rules, checking the outputs, and error correction.

It is in this change that the role of the designer is complicated. Designers have ceased to be content creators only. They also turn into the managers of the system. This change is both useful, though it is very responsibility creating. When the control system is weak, designers may have lost excessively much control over tone, coherence and intent.

This aspect is also present in the study of workflow. Also, the systematic review of generative AI in game character-creating states that the generative tools can simplify the workflow, decrease technical pressure, and reduce technical barriers, particularly in the case of prototype and asset development [5]. Nevertheless, that review largely dwells on character development process, rather than dialogue or story integrity. It can be used to prove statements about efficiency in production, but it will not do well as evidence of narrative control [5].

2.5 Educational and serious game context evidence

In the context of educational and serious games research, the usefulness of AI-driven NPCs as well as their limits are demonstrated. Xiao, Li and Guo discovered that ChatGPT-driven NPCs enhanced students in their intrinsic motivation, and situational interest, and also the experimental participants demonstrated better performance in writing [6]. These results indicate that AI-oriented NPCs are capable of enhancing interaction and learning in case the task is obvious.

The same outcome is documented in a serious game concerning digital literacy by Chernbumroong et al. According to their research, the system must retain the pedagogical coherence, as well as the flexibility of AI generation [7]. It is also noted in the study that adaptive AI support might have a trade-off with certain technical problems like system latency [7]. The authors present an argument that generative AI can be utilized as a dynamic interface layer in serious games with one condition being that the design must contain structured guidance and a clear understanding of the role of the NPCs [7].

A wider literature review by Mitsea, Drigas and Skianis also indicates that AI-based serious games may be used to provide adaptive, personalized and interactive training environments. Simultaneously, the review indicates that there are still issues in assessment, technical incorporation, and protracted validation of such systems [8].

So such studies are useful, but their meaning is restricted. They demonstrate that AI-based NPCs may be effective into the situation when the task is well-defined and the surrounding is controlled. These do not show that free AI conversation is prepared to purpose all game stories.

3 Case analysis

3.1 Detroit: become human as a controllable narrative standard

The held comparison example is a useful case of Detroit: Become Human. This game does not represent an instance of NPC generation using a large language model. Nonetheless, it is useful to this paper as it demonstrates how a very interactive narrative can remain consistent with a very powerful author control [9].

The game is structured on the choice of the player and the endings. Decisions made by the player in the story can influence the scenes, relationships and outcomes. The design of this makes the player feel that they have a lot of agency. Meanwhile, that agency remains within an already created framework. The conversation, scene action and potential implications is also scripted out. Thus the player does have power, but the story does not open in an unlimited manner.

That is the reason why the game is significant here. It demonstrates that high interactivity does not require free-form AI generation to be significant. The game establishes interaction with predetermined branching and not via open dialogue. The designer still dictates information timing, character arcs and dramatic results. That is, the game provides the player with space to play, and the game does not disrupt the narrative sequence.

3.2 Detroit: become human needs to be regarded as a contrast case

Simultaneously, Detroit: Become Human cannot be regarded as an evidence that the NPC systems based on language-models prove their effectiveness in the narrative-focused games already. This game is effective because of a well-crafted branching stories as opposed to open-ended generation [9].

This distinction is relevant in that the issue of narrative stability is the concern of the present paper. In a branching game such as Detroit: Become Human, the designer is able to see the paths that the player follows. The designer is able to manage the way information is presented, how emotional scenes are presented and how consequences are presented using time. In an unconstrained LLM system this control is weaker unless further structure is added in the design. Automatic dialogue that sounds natural is not enough to ensure a consistent characterization, proper pace, and consistency of the world [1].

Accordingly, the merit of Detroit: Become Human does not lie in the fact that it demonstrates that AI is capable of dealing with the tricky narrative design. The worthiness of it lies in the fact that it offers a reference point. It demonstrates the accomplishments of a solid narrative design when it is well planned and structurally restricted.

3.3 PANGeA as a limiting generative model

PANGeA takes another ride. It is not entirely dependent on the text that the designers rewrote in branches. Rather than that, it discusses what large language models can do to create narrative content in turn-based role-playing games without compromising, yet maintaining structure [1].

PANGeA has a value in utilizing AI, but in utilizing it in its specific form. The system does not allow the model to be an unrestricted speaker. The system fits the model into an ordered narrative process. The memory system assists in storage of previous background information.

The generated content is compared with the game rules and maintained in accordance with the current narrative state by the validation system [1].

It is also demonstrated in the paper that such support systems are important. PANGeA suggests that game designs that do not allow derailment of the game by input of out-of-scope players require frameworks that allow designers maintain the notion of narrative coherence in the game, although it is desirable to involve active players within the game [1]. All risk is not eliminated in this design. Nonetheless it presents a far more justifiable course of approach than free generation.

3.4 What educational AI-NPC systems have taught us

The same general trend is favored by the educational studies. Xiao et al. demonstrate that NPCs which are based on ChatGPT are capable of enhancing motivation and writing performance in a narrative-based learning environment [6]. Chernbumroong et al. also demonstrate that NPC interaction with AI can generate more engagement and aid learning in a serious game setting [7].

Nevertheless, it is only through well-framed and structured goals that such systems will be successful. These NPCs do not even have to bear the burden of an extended entertainment story. They operate within narrow scope of work with goals that are more easily regulated. Other issues, including latency, redundancy, and disruption of flow, are present even in such environments [7].

All in all, these instances are headed toward the same direction. Games which are highly authored maintain their coherence via branching design. PANGeA maintains consistency using memory and validation. Limited objectives and structured communication help educational AI-NPCs remain practical. Complete freedom is not the common element. The shared factor is control.

4 Discussion and recommendations

4.1 Freedom without structure is weak design

Many discussions of AI assume that more freedom will naturally create a better game experience. This idea sounds attractive, but it is weak from a design point of view. In narrative games, freedom only matters when players can understand the meaning of their actions inside a coherent system.

If an NPC can answer many questions but those answers weaken world logic, pacing, or tone, then the interaction may seem broad but still remain shallow. It may look flexible, but it does not create a strong whole.

This is where many AI-driven NPC ideas still fail. They focus too much on novelty. They show that an NPC can answer many questions, but they do not show that those answers remain useful inside the story. If an NPC forgets earlier events, changes tone for no clear reason, or says things outside the world of the game, then the system may feel impressive for a short time but weaken the overall experience.

So unrestricted openness is not good design by itself. In many cases, it shows that the design has not solved the hardest problem.

4.2 Narrative stability should be a main design goal

This paper argues that narrative stability should be a core requirement in AI-driven game design. Developers should not only ask what an NPC can say. Developers should also ask what the NPC cannot say, which information must stay hidden, and how continuity will be preserved across long interactions.

PANGeA is useful because it takes this issue seriously. Its validation system exists to keep free-form generation consistent with narrative conditions [1]. Chernbumroong et al. make a similar point in another setting by arguing that AI systems need to balance open-ended generation with structured support [7]. Xiao et al. also show that AI-powered NPCs work best when they act inside a bounded scenario instead of an unrestricted conversation space [6].

This logic should not be treated as optional. Any game that wants to use LLM-driven NPCs in major narrative roles needs some system for identity control, world-rule enforcement, and continuity tracking. Without these systems, the game does not really use AI for narrative design. It only attaches a language model to a weak structure.

4.3 AI changes the role of the designer

Generative AI also changes the role of the designer. AI can reduce some repetitive work and help with brainstorming and early prototyping [2]. But AI does not remove the need for writers or narrative designers. Instead, it changes where authorship happens.

StoryVerse makes this point clearly. The paper presents a workflow that connects a writer’s “authorial intent” with the “emergent behaviors” of LLM-driven character simulations through “abstract acts” [10]. This point matters because it shows that even more open narrative systems still need authors to define high-level structure and guide plot movement.

So AI does not remove human judgment. In fact, AI may make that judgment more important. In AI-assisted systems, authors may shape rules and boundaries more often than individual lines, but they still remain responsible for structure, tone, and narrative purpose.

4.4 Recommended architecture for LLM-Driven NPCs

A practical development path is hybrid architecture, not unrestricted generation. Based on the research discussed above, an effective AI-based NPC system should include at least four layers.

First, the system needs character definition. Each NPC should have clear identity limits, knowledge boundaries, tone rules, and behavior rules. This design reduces character drift.

Second, the system needs context management. Long-term interaction needs memory, summarization, or structured retrieval so that NPCs can stay consistent over time [1].

Third, the system needs verification control. The design should check generated responses before they reach the player. The system should compare them with world rules, plot conditions, and safety boundaries [1]. This is one of the clearest lessons from PANGeA’s “novel validation system” [1].

Fourth, the system needs selective use. Not every part of the game should allow the same level of openness. Optional side dialogue can be more flexible. Main plot scenes should remain tightly controlled. Educational AI-NPC research also supports this point because the strongest results appear in settings with bounded goals and structured support [5, 6].

This approach is not as flashy as the idea of fully autonomous narrative. But from a design point of view, it is much stronger. It treats AI as a tool inside narrative design, not as a replacement for narrative design itself.

4.5 Where AI-Driven NPCs fit best

The risk of AI-driven NPCs changes across different kinds of games. In educational games, training simulations, bounded role-playing scenarios, and other structured settings, generative interaction can bring clear benefits while the risks remain easier to manage [5-7]. In these cases, some flexibility can improve engagement without forcing AI to carry the full weight of a tightly planned dramatic plot.

In highly orchestrated narrative games, the risk is higher. The more a scene depends on tone, foreshadowing, rhythm, or moral tension, the less room there is for uncontrolled output. In these games, AI should be used in a selective way. AI may help with optional dialogue, replay variation, or low-risk support scenes. But the main dramatic structure should remain protected.

This view may sound cautious. Still, it is more defensible than pretending that current technology is ready to control every part of narrative design.

5 Conclusion

This paper examines how NPC systems based on large language models are changing the interaction structure of game design. Traditional scripted NPCs limit player freedom, but they preserve narrative order through strong author control. Generative NPCs increase flexibility and responsiveness, but they also bring risks such as inconsistent characters, narrative deviation, and weaker control over story boundaries.

The comparison in this paper shows that the main challenge is not whether AI can generate natural dialogue. The deeper challenge is whether the system can stay stable while handling unpredictable player input. If developers pursue openness without enough control, they may weaken the coherence of the game instead of improving its quality.

So the most promising direction is not unrestricted generation. A stronger path is hybrid design. AI can be a useful creative and interactive tool, but it should work inside a system that protects consistency, structure, and design intent. In narrative-heavy games, the goal should not be full automation of story narration. The goal should be to expand interaction while preserving the integrity of the story.

References

- [1] Buongiorno, S., Klinkert, L., Zhaung, Z., Chawla, T., Clark, C.: PANGeA: Procedural artificial narrative using generative AI for turn-based, role-playing video games. Proceedings of the Twentieth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2024). (2024)
- [2] Alharthi, S.A.: Generative AI in game design: Enhancing creativity or constraining innovation? Journal of Intelligence. 13, art. 60. (2025)
- [3] Liu, S.M.: Research on the innovation and practice of game interaction mode empowered by AIGC. Proceedings of the 1st Youth Academic Forum on Social Innovation and Economic Development. pp. 49–52. (2025)
- [4] Nagarkar, S.B.: Improving realism and interactivity in video games: A study of AI-integrated non-player characters (NPCs). Master's thesis, Universitat Politècnica de Catalunya and Sapienza Università di Roma. (2024)
- [5] Wu, Z., Chen, Z., Zhu, D., Mousas, C., Kao, D.: A systematic review of generative AI on game character creation: Applications, challenges, and future trends. IEEE Transactions on Games. (2025)
- [6] Xiao, Y., Li, D., Guo, K.: Using ChatGPT to bring non-player characters to life: Effects on students' storyline-driven game-based writing learning. Computers & Education. 238, art. 105414. (2025)
- [7] Chernbumroong, S., Intawong, K., Asawimalkit, U., Puritat, K., Julrode, P.: Design and evaluation of a generative AI-enhanced serious game for digital literacy: An AI-driven NPC approach. Informatics. 13(1), art. 16. (2026)
- [8] Mitsea, E., Drigas, A., Skianis, C.: A systematic review of serious games in the era of artificial intelligence, immersive technologies, the metaverse, and neurotechnologies: Transformation through meta-skills training. Electronics. 14, art. 649. (2025)
- [9] Quantic Dream: Detroit: Become Human. Official game page. (n.d.)
- [10] Wang, Y., Zhou, Q., Ledo, D.: StoryVerse: Towards Co-authoring Dynamic Plot with LLM-based Character Simulation via Narrative Planning. Proceedings of the 19th International Conference on the Foundations of Digital Games (FDG 2024). (2024)