Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD28116

Generative AI in Immersive Games and Player Agency: A Comprehensive Investigation

Sitong Jiang

College of Communication, Boston University, Boston, USA
jstong@bu.edu

Abstract. User experience has become central to the value of contemporary games, but traditional scripted responses and linear storylines still constrain immersion and agency. This review examines how generative Artificial Intelligence (AI) addresses those limits by enabling narrative participation and embodied role construction. It synthesizes recent academic studies and credible industry reports, and organizes applications across three aspects: character art pipelines that improve consistency and speed, dialogue-driven characters that influence story progression, and adaptive worlds that regulate pacing and opportunity in real time. The literature indicates gains in responsiveness, personalization, and production efficiency, but also persistent risks. Models often struggle with reasoning and human nuance, data practices raise concerns about bias, privacy, and transparency, and platform policies and regulation are evolving. To balance opportunity and risk, the review offers practical directions that keep human intent in charge while using AI for scale. These include hybrid authoring with clear constraints, transparent opt in personalization, built in provenance and credit, and evaluation focused on immersive and agency. This review provides a comprehensive overview of current progress and identifies priorities for the responsible adoption of AI. It aims to help readers and industry partitioner understand where generative techniques have made a difference, where evidence remains limited, and how future research can advance immersive play without sacrificing trust.

Keywords: Generative AI, Immersive Experience, Player Agency

1. Introduction

In recent years, the global game industry has expanded rapidly, becoming one of the most profitable and technologically innovative sectors within entertainment. The 2024 Newzoo Global Games Market Report projects that the industry will generate approximately $187.7 billion in annual revenue, with billions of players worldwide. This scale underscores both the cultural significance and economic weight of gaming [1]. As platforms and engines gains greater technical capable, player expectations have shifted toward richer emotional engagement, deeper interactivity, and greater personalization. In this context, user experience has become a primary factor in shaping a game's reputation and commercial success.

To overcome limitations of traditional approaches and meet rising expectations, researchers and practitioners are exploring new technical approaches. Artificial Intelligence (AI), especially machine

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD28116

learning and generative models, has become a powerful tool in the way games are designed, experienced and studied. Advances in deep learning over the past decade and the more recently generative models have broadened the potential for AI to enhance user experience. Large language models and multimodal systems now support story framing, branching dialogue, and design assistance. Image and 3D generators accelerate asset creation and prototyping. These capabilities are gradually moving from experimental settings into industrial practice. Ubisoft's NEO NPC prototype, which is developed with Inworld AI, explores generative character that preserve role integrity during freeform conversations [2]. NVIDIA's ACE initiative bundles speech, animation and conversational intelligence into a toolkit for deployable digital humans, both signaling the shift toward real time agents [3]. At the workflow level, a 2025 Google Cloud survey reports that roughly 87% of game developers use AI agents somewhere in development [4].

Despite clear gains, challenges still exist. Empirical studies with designers find productivity gains from generative tools alongside worries about originality and creative dependency, and industry surveys report broad adoption tempered by unresolved questions about data ownership. Scholars warn that technical progress brings ethical and practical risks [5]. Oliveria and Rito argue that personal data management and classification can embed social bias, with race and gender stereotypes distorting models [6]. Ratican and Hutson emphasize privacy and consent across design and runtime data, highlighting the need for transparent pipelines when player traces and creator materials inform generation [7]. Creators likewise voice concern about their rights, attribution, and potential displacement. Taken together, these limitations signal the need for continued exploration for an approach that keeps humans in the loop and treats AI as an assistive collaborator.

This review examines how generative AI is reshaping immersive experience by enabling narrative participation and embodied presence across games. Drawing on academic studies and credible industry reports, this review identifies where current techniques clearly strengthen immersion and player agency and where evidence is still tentative. Discussion foregrounds unresolved issues in reasoning and human nuance, bias and privacy, and platform regulation and governance. The remainder of this article proceeds as follows: Section 2 surveys application patterns that link generative AI to immersive design and player agency. Section 3 examines the main challenges and outlines future prospects, and concludes priorities for future research.

2. Application of generative AI in immersion and agency

Generative AI is increasingly integrated into both game production and runtime systems, influencing the fundamental components of immersion and agency. In this section, this review groups recent work into three overlapping aspects: character art and asset pipelines that improve consistency and responsiveness, conversational characters and narrative systems that let players shape narrative through language, and adaptive worlds that respond promptly to circumstances.

2.1. Character art design

2.1.1. Sprite sheet diffusion for consistent 2D character motion

Hsieh et al. propose Sprite Sheet Diffusion, a pipeline that adapts a video generation domain, Animate Anyone, with a reference image and pose guider to produce consistent multi-frame sprite animations from a single character image. The authors report "faithful, consistent, and high-quality" action sequences across common moves like jump, run, and attack. They also provide ablations on

DOI: 10.54254/2755-2721/2025.LD28116

appearance consistency and temporal stability [8]. For 2D games, this reduces the hand-drawn frame burden while preserving style, which supports visual continuity

2.1.2. Text-to-3D characters

For 3D assets, Poole et al.'s DreamFusion demonstrated that pretrained 2D diffusion model can guide Neural Radiance Field, turning text prompts into view-consistent objects without paired 3D data [9]. Lin et al. advanced this approach with Magic3D, making it more practical for production. Their coarse-to-fine optimization produces a textured mesh suitable for rigging and level-of-detail adjustments. In user studies, participants showed a clear preference for the higher resolution mesh output and faster turnaround, as this allows artists refine silhouette, costume readability, and scale directly within the engine rather than through repeated repaints [10]. For immersive experiences, the shorter loop improves visual continuity for avatars and none playable characters (NPCs), and helps sustain embodied presence across perspectives and sessions.

2.1.3. Controllable diffusion for identity-preserving character art

To keep identity and pose stable while still allowing variety, Zhang et al. add spatial controls to large diffusion models through ControlNet, conditioning on edges, depth, segmentation, or human pose. In character workflows, these controls anchor face shape, stance, and garment boundaries while generation fills in materials and detail. By mitigating style breaks between concept art, in-game models, and promotional stills, this approach makes it feasible to explore costume or expression variations without losing recognition, thereby supporting both immersion and a clearer sense of agency in dialogue and narrative scenes [11].

2.2. Conversational characters and narration participation

Generative AI is changing how players converse with worlds and how stories advance in response. Recent work on language-based interaction has shifted conversation from decorative flavor to a primary mechanism for advancing play. Cox and Ooi analyze player feedback from a commercially released title that enabled Large Language Model (LLM) driven replies, deriving guidelines for persona definition, memory, retrieval from world lore, and safety filters. Players reported higher responsiveness and variety, but also noted drift, verbosity, and occasional breaks in tone when memory and guardrails were weak [12]. Complementary research shows that adding structured memory and reflection contributes to more believable multi-day behavior. In the Generative Agents study by Park et al., characters that store observations, retrieve relevant experiences, and periodically synthesize them into higher-level plans produce coherent social routines. Users judge these routines to be more lifelike, which is essential for conversational scenes that drive the plot [13]. These observations link directly to agency. When characters retain memory, remain consistent in their portrayal, and reference shared facts, players report a greater sense of being recognized. This recognition often increases their willingness to engage in conversational play rather than replying on menu-based dialogue selection.

Industry prototypes illustrate how these ideas are being implemented in practice. Ubisoft's NEO NPCs integrate unscripted dialogue with memory, contextual awareness, real time emotion and animation [2]. By doing so, they position conversation as a system that can unlock story elements, adjust quests, and respond to situational cues, rather than a predetermined set of lines. Their goal is not only to enrich dialogue, but also design interactions that carry consequences. Similarly,

DOI: 10.54254/2755-2721/2025.LD28116

NVIDIA's ACE for Games packages speech recognition, text-to-speech, conversational intelligence, and real time facial animation into deployable models for game engines [3]. These stacks enhance immersion by reducing latency between player input and character response. They also provide teams standardized, repeatable methods for achieving character performance.

2.3. Adaptive world and pacing

Adaptive worlds are designed to keep challenge, tone, and opportunities aligned with the player in real time. The concept is not new. Techniques like dynamic difficulty adjustment can adjust encounters based on player performance and need, while procedural generation shapes content based on player experience and selection. Immersion deepens when the world responds to what the player does, and agency grows when the system recognizes intent and offer possibilities rather than guiding to the fixed script.

Pacing systems bring this view to fruition. The AI Director in Left 4 Dead monitors play and adjusts spawn rates, resources, and rhythms to maintain tension and relief. Valve's developer documentation describes how director scripts shape intensity curves and coordinate set-pieces while preserving authored beats [14]. The method treats drama as a controllable variable rather than a fixed curve, which often results in gameplay that feel authored yet surprising.

Recent large language models extend adaptation from encounter tuning to embodied planning in open worlds. Wang et al.'s research on Voyager presents an open-ended autonomous agent in Minecraft that writes and refines its own skills, learns from the environment, and progressively explores the world. In controlled evaluations, it collects more unique items, reaches tech-tree milestones faster, and transfers learned skills to fresh worlds better than prior methods. Even though it runs outside a commercial game, the architecture points toward embodied play companions and director systems that adapt goals and tools to the player's evolving context [15].

These techniques have since evolved to include more complex sensing and response pipelines. Contemporary implementations blend three functional layers. First is sensing: telemetry, speech, and interaction traces reveal player intent and pace. Second is memory: persistent data such as session history and character state ensure continuity across time. And third is policy: decision layers manage how and when the system should react, inject resource, or prompt an ally to offer help. The adaptive systems complement traditional scripting with a broader design that designers can author flexible templates while allowing runtime engines to tailor the experience. When sensing, memory, and policy are properly aligned, players feel recognized and influential.

3. Discussion

While generative AI has enabled new forms of interactivity and immersion in games, these advances also bring several unresolved challenges. The following section explores the key limitations that remain across reasoning, bias and privacy, and regulatory compliance.

3.1. Challenges

3.1.1. Huaman nuance, reasoning, and contextual understanding

Generative AI models often generate responses that sound coherent but lack deeper reasoning or human nuance. Immersion in games depends not only on credible dialogue or reasonable plot pacing, but also on the system's ability to respond with emotional intelligence and social cognition. These qualities remain difficult for current models to achieve. Players might express hesitation,

DOI: 10.54254/2755-2721/2025.LD28116

curiosity, or moral conflict, but the system responds as if these were simple commands rather than reflective cues. This disconnect is particularly evident in multiplayer or narrative-driven games, where dialogues shapes character development and consequence. Cultural slang, memes, and online trends, common elements in real time player communication, are rarely reflected meaningfully in the model responses. AI also struggles to recognize the emotional context behind voice and action, often responds with neutral or mismatched content. Even when the output sounds convincing, the limitations can disrupt the immersive experience.

There is also a growing concern that systems are simplifying player behavior into statistical patterns rather than individual expression. Predictive models trained on large datasets may overlook personal nuance in favor of the most common or efficient scenario. For example, a game might adjust difficulty based on similarities with other player profiles, unconsciously overriding the actual intentions or emotional states of players. Meanwhile, in multiplayer or competitive contexts, adaptive systems may raise questions about fairness among players. If an AI system grants subtle advantages to maintain player engagement, some may feel that outcomes are manipulated, especially when the system cannot explain its logic or criteria. These opaque decisions will erode trust and make it harder for players to understand or influence the game on their own terms.

3.1.2. Bias, privacy, and transparency

Generative AI systems trained on large datasets will unavoidably absorb the biases embedded in those datasets. When these systems are used in games, the impact becomes highly visible in characters, world-building, and adaptive systems. Scholars such as Oliveria and Rito caution that classification systems used for emotion recognition or personalization may reinforce racial, gender, or cultural stereotypes [6]. A character's behavior may unintentionally reflect skewed assumptions about identity, social roles, or emotional norms, thereby alienating or misrepresenting players.

Privacy is also a central concern. Adaptive models often draw on real-time player data such as movement patterns, interaction traces, voice input, or narrative choices. While this can enhance personalization, without clear boundaries, it can be used in ways that players do not expect or approve of. Once applied, these systems must respond to complex or ambiguous input. If the system misinterprets intent or fails to moderate harmful content, trust is quickly eroded. The good news is that platforms have begun to respond to these concerns. Valve's Steam policy requires developers to clearly disclose when AI is used, explain whether content is generated in advance or during gameplay, and describe how they prevent systems from producing harmful or illegal material [16]. These requirements reinforce the idea that player safety and informed consent must be treated as core elements of design. Immersion and personalization can be powerful, but they must not come at the cost of transparency or respect for user data.

3.1.3. Regulation and platform governance

Generative features now operate at the intersection of content regulation, data protection and security requirements. Development teams must navigate privacy laws that govern personal data, including voice, biometrics, and behavioral telemetry. They are also subject to content rating standards that restrict violence or sexuality, as well as platform policies that require disclosure of AI use and implementation of safeguards for live generation. These overlapping requirements create a shifting regulatory baseline. What is acceptable in one region or store may require different disclosures, data handling practices, or moderation measures in another. For immersive systems that learn from player traces or generate dialogue in real time, these obligations are substantive. They

DOI: 10.54254/2755-2721/2025.LD28116

shape what can be stored, how long memory persists, and whether certain kinds of adaption are permitted.

Regulations introduce more than just legal obstacles; they also bring additional technical work. Teams are expected to provide clear consent processes that give players control over memory and personalization, as well as audit logs that document when content is created, screened, or rejected. Such practices allow players to understand how and why the system responds. To further support transparency and accountability, provenance tracking should be embedded within the asset pipeline. This makes it possible to reconstruct training sources, prompts, and edits if a dispute occurs. At the same time, real time features like generative dialogue or reactive environments must go through safety systems without noticeable delay. This is especially difficult in virtual reality or fast-paced action games where even minor delays or interruptions can disrupt immersion. If safeguards are too slow or opaque, they undermine pacing and presence. But if they are too permissive, they create enforcement challenges and reputational risks.

3.2. Future prospects

As generative AI continues to evolve, future development in games must prioritize the balance between automation and human direction. Generative systems are well suited to scaling variations and accelerating production, while human creators remain responsible for establishing tone, ethics, and intent. A practical approach is hybrid authoring, in which designers define structures, constraints, and style, and models operate within those boundaries. Extracting information from carefully crafted stories and conversational memories will anchor facts and sounds, making responses feel coherent and emotional. When the uncertainty is high, systems can display a brief notice and fall back to authored lines or safe defaults rather than improvise.

Personalization works only when players understand and control it. Systems should pair consent with clear setting, offering opt in memory systems for continuity and easy reset options for privacy. Let it default to minimal data collection and expand only with explicitly opt ins. Games should also allow players view, edit, and delete the memory that drives personalization, and have session data expire unless they choose to keep it. Moreover, developers should provide short and plain language notices, such as why a quest branched or why assistance appeared, to make adaptation transparent and to give players the context they need to adjust or revoke permissions.

Evaluations and tests should concentrate on factors that shape to immersion and agency. Designers can combine play telemetry with measures of engagement and satisfaction, and test whether adaptive pacing reduces boredom without raising frustration. By preserving manual oversight of pacing, voices, and visual standards, creative work can maintain its integrity and vibrancy. Under these circumstances, generative AI can extend what designers express, while keeping players' privacy, safety, and authorship at the center of immersive play.

4. Conclusion

This review explains how generative AI can deepen immersion and expand player agency in games. It synthesizes applications across creation workflows, language-based interaction, and runtime adaptation. The evidence suggests an improvement in responsiveness, variety, and visual continuity, especially when memory outputs and pacing controllers keep intense, clear, and understandable. However, problems remain, including weak reasoning and limited human perception, bias and privacy risks in the use of data, and changing regulatory requirements. In the future, this review will

DOI: 10.54254/2755-2721/2025.LD28116

continue to track new tools and production practices, improve comparable methods to determine whether they truly enhance immersion and agency in games.

References

[1] Newzoo. (2024). 2024 Global Games Market Report. Newzoo. https://resources.newzoo.com/hubfs/Free%20Reports/Games%20Market%20Report%20and%20Forecasts/2024_Newzoo_Free_Global_Games_Market_Report.pdf

[2] Higham, M. (2024). How do Ubisoft's AI-driven npcs handle dynamic player interactions? Game Developer. https://www.gamedeveloper.com/design/how-do-ubisoft-s-ai-driven-npcs-handle-dynamic-player-interactions-

[3] NVIDIA. (2025). NVIDIA ACE for Games. NVIDIA Developer. Retrieved September 2, 2025, from https://developer.nvidia.com/ace-for-games

[4] Kachwala, Z. (2025). Nearly 90% of videogame developers use AI agents, Google study shows. Reuters. https://www.reuters.com/business/nearly-90-videogame-developers-use-ai-agents-google-study-shows-2025-08-18/

[5] Nasir, M. U., James, S., & Togelius, J. (2024). Word2world: Generating stories and worlds through large language models. arXiv preprint arXiv: 2405.06686.

[6] Oliveira, C. V., & Rito, P. N. (2024). Enhancing Player Experience Through Generative Artificial Intelligence: Custom Interaction in Game Design. In International Conference on Videogame Sciences and Arts (pp. 259-268). Cham: Springer Nature Switzerland.

[7] Ratican, J., & Hutson, J. (2024). Adaptive worlds: Generative AI in game design and future of gaming, and interactive media. ISRG Journal of Arts, Humanities and Social Sciences, 2(5).

[8] Hsieh, C. A., Zhang, J., & Yan, A. (2024). Sprite Sheet Diffusion: Generate Game Character for Animation. arXiv preprint arXiv: 2412.03685.

[9] Poole, B., Jain, A., Barron, J. T., & Mildenhall, B. (2022). Dreamfusion: Text-to-3d using 2d diffusion. arXiv preprint arXiv: 2209.14988.

[10] Lin, C. H., Gao, J., Tang, L., Takikawa, T., Zeng, X., Huang, X., ... & Lin, T. Y. (2023). Magic3d: High-resolution text-to-3d content creation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 300-309).

[11] Zhang, L., Rao, A., & Agrawala, M. (2023). Adding conditional control to text-to-image diffusion models. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 3836-3847).

[12] Cox, S. R., & Ooi, W. T. (2023). Conversational interactions with npcs in llm-driven gaming: Guidelines from a content analysis of player feedback. In International Workshop on Chatbot Research and Design (pp. 167-184). Cham: Springer Nature Switzerland.

[13] Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. In Proceedings of the 36th annual acm symposium on user interface software and technology (pp. 1-22).

[14] Valve Developer Community. (2025). Left 4 Dead 2/Scripting/Director scripts. Retrieved September 5, 2025, from https: //developer.valvesoftware.com/wiki/Left_4_Dead_2/Scripting/Director_Scripts

[15] Wang, G., Xie, Y., Jiang, Y., Mandlekar, A., Xiao, C., Zhu, Y., ... & Anandkumar, A. (2023). Voyager: An open-ended embodied agent with large language models. arXiv preprint arXiv: 2305.16291.

[16] Steamworks Development. (2024). AI content on Steam. Steam. https://store.steampowered.com/news/group/4145017/view/3862463747997849618