

Article

Emerging Practices in LLM-integrated Game Writing

DIGDEM SEZEN, University for the Creative Arts, United Kingdom;
email: digdem.sezen@uca.ac.uk

TONGUC SEZEN, University for the Creative Arts, United Kingdom; tonguc.sezen@uca.ac.uk

10.2478/bsmr-2025-0004

ABSTRACT

This article examines emerging practices in large language model (LLM) integration within game writing, focusing on how these technologies reshape narrative design, creative workflows, and professional roles. Drawing on evolving industry experimentation and academic research, it outlines the relationship between traditional game writing and LLM-driven approaches, surveys new forms of interactive storytelling such as conversational NPCs, multi-agent simulations, adaptive commentators, and LLM-based text adventures, and identifies their narrative affordances and constraints. The article analyses core challenges, including hallucination, bias, narrative incoherence, and control, and discusses current strategies to address them, such as fine-tuning, prompt engineering, and new authoring tools that position writers as system-level narrative architects. It argues that LLM integration represents not automation but a reconfiguration of co-creative authorship between writers, machines, and players, and calls for further research into ethical design, player reception, and the evolving responsibilities of narrative professionals in LLM-augmented game development.

KEYWORDS: large language models; game writing; interactive storytelling; AI-assisted creativity; narrative design

INTRODUCTION

Since the release of ChatGPT by OpenAI in late 2022, advances in large language models (LLMs) have sparked widespread interest in their potential impact on game development. This study examines how LLMs are being integrated into game narratives and how this transformation is reshaping game writing practices. Rather than analysing the wider LLM ecosystem or the architectures, computational performance, and generative capabilities of specific models, the study focuses on how their integration influences creative workflows, professional roles, and the evolving craft of interactive storytell-

ing. The study begins with an overview of traditional game writing practices and their relationship to LLMs. It then outlines key types of LLM-integrated game narratives, highlighting emerging trends. Finally, it investigates specific techniques and tools used by game writers and discusses how writers can adapt their skills to LLM-integrated practices.

This is an evolving field, with new tools and approaches emerging at a steady pace. Game developers and researchers are experimenting with novel applications, making it difficult to capture a comprehensive or lasting snapshot. Rapid industry

innovation outpaces academic theoretical development, which requires an intersection of computer science, game studies, narrative theory, and human-computer interaction while maintaining rigorous academic standards and practical industry relevance. In this regard, we acknowledge that some of the insights presented here may soon become outdated. Nevertheless, identifying current shifts in the role of game writers and narrative designers offers valuable perspectives on how narrative craft may evolve over time.

TRADITIONAL GAMES WRITING PRACTICES AND LLMs

Game writing, particularly the crafting of interactive narratives, has long been intertwined with constantly evolving computational methods (Sharples, Pérez y Pérez 2022). Discussions on how algorithmic structures shape narrative experiences by requiring nontrivial effort to exercise agency form part of the foundational scholarship in game studies (Murray 1997; Aarseth 1997). Studies on software built on rule-based systems positioning human authors as general rule-setters without control over the details of their literary output can be traced back to the early 2000s (Manovich 2002; Ensslin 2007). Procedural narrative design frames interactors as initiators of stories emerging in simulations driven by rules and populated by story fragments (McRae 2020; Paterson et al. 2020). According to Rettberg and Rettberg (2024), these diverse approaches to algorithmic narrativity implies a co-construction of meaning between humans and algorithms, combining the human ability to understand experience through narrative with the power of machines to process and generate data. In the context of games, this co-construction is shared among the human author or game writer, the algorithmic system, and the player as interactor.

Neural networks, the underlying technology of LLMs, introduced a shift away from previous regulated algorithmic patterns toward probability-driven mimicry (Ensslin, Nelson 2025). With the right

prompts, they can generate probabilistic outputs in the form of coherent short stories on a given theme, within a particular style, from a particular point of view, thus exemplifying a new form of collaborative writing between humans and machines – a type of cyborg authorship (Rettberg, Rettberg 2024). Investigating the question of authorship further, Hongisto (2025) argues that literary works generated through neural networks exemplify a mode of collective co-creation. According to him, this collective process involves machines, their programmers, the source texts used for training, and their authors. He further contends that curators who select and structure the training data may also be considered co-authors when guided by artistic intent. However, LLMs tend to obscure these collaborative relationships due to their complexity and extensive scale of their training datasets. The economic incentives and closed architectures of commercial models conceal the role of human authors and curators. As a result, their outputs can be problematic in terms of accountability and may be driven by intrinsic cultural biases and should be engaged critically (Freed 2025). To function effectively in domain-specific or creative contexts such as games which involve players as additional creative actors, LLMs typically require further adaptation. In this regard, how LLMs should be linked to narrative elements and functions in games, how effective they can be, and how the responsibility of generating elements of interactive narrative experiences in games can be shared between humans and LLMs are key questions to be answered. Human-computer co-creativity in game writing should address the individual and collaborative creative processes of humans and machines, the results of these processes, and the context of games themselves as an environment (Kantosalo, Takala 2020). However, to evaluate these, we must first address the diverse nature of game narratives, the collaborative practices that bring them to life, and the tools that structurally shape their creation.

There are diverse types of games with diverse approaches to narrative. Most games, except perhaps more abstract puzzle or strategy games, require some level of narrative element which traditionally must be crafted by game writers. The level of dramatic agency given to or perceived by players, and the ways in which this is achieved through technological and design means, are key variables of how a game is written (Smed et al., 2021). A game with a linear main story not influenced by player actions may still require game writers to craft material for cut-scenes, scripted events, or environmental storytelling. In choice-based genres such as adventure games, role playing games (RPGs), and visual novels, some of the core narrative elements game writers are expected to develop may include action and dialogue choices as well as other story related variables traced by the system (i.e. certain positive or negative actions not explicitly categorised as player choices), and the diverging story branches (i.e. non-player character (NPC) reactions to player choices) including multiple endings that the choices and variables may lead to (Finley 2023). Balancing players' narrative agency and the coherence of potential narrative experiences becomes a key responsibility for game writers in this creative process. Emergent narratives, meaning game narratives created mainly through play in simulated game worlds without any directorial control by the designers, also require elements crafted by game writers, such as a library of possible events and other narrative components which can be linked to starting points, environments and objects, conflicts, objectives and so on (Paterson et al. 2020). Writers aim for thematic coherence and multiplicity of narrative potential in these cases. Overall, given the diversity of narrative structures and player experiences, the integration of LLMs into game writing cannot be guided by a singular, universal goal or standard. Such integration demands a nuanced, context-sensitive approach that recognises the differing narrative demands of each game.

Traditionally, creating the diverse narrative content required in games demands an equally diverse range of writing practices and expertise. Depending on the size of the project, the responsibilities of game writers may be distributed across multiple roles, allowing for specialisation in a collaborative development. As one of the main professional bodies representing game writers, The Writers' Guild of Great Britain (2021) identifies a series of core roles in game writing, each with distinct but interconnected responsibilities unique to games: overseeing the overarching structure of the story, narrative directors are responsible for shaping the main narrative arc and designing overall branching pathways that give players meaningful choices. They ensure that the broader vision of the story aligns with the interactive nature of games, setting the tone and thematic depth that guide the entire narrative process. Tasked with crafting the details, including dialogue, textual artefacts, and contextual backstories, game writers ensure that game worlds feel alive and deep, and characters, quests, and environments feel authentic and immersive. Game writers create the narrative threads that players engage with directly and are expected to pay close attention to language and tone (Lasssheikki 2022). Playing a critical bridging role, narrative designers are positioned at the intersection of storytelling and game design (Trygg et al. 2024). Their primary focus is achieving ludonarrative harmony, a state in which a game's mechanics and narrative structures are aligned to reinforce the same themes and emotional experiences, ensuring that what players do in the game meaningfully supports what the story conveys. This involves ensuring that player actions feel meaningful within the context of the story and that the story evolves in ways that enhances gameplay (Hocking 2009). Narrative designers collaborate closely with both writers and game designers, orchestrating a balance between interactivity and narrative coherence (Lasssheikki 2022). Beyond these core roles, there may be additional writing tasks, including

creating tutorials, technical documentation, or marketing content, all of which contribute to the broader narrative experience. The integration of LLMs into the collaborative process of game writing should consider the specific requirements of each role. It is important to explore how the strengths and generative potential of LLMs can support and complement different forms of creative expertise, while also acknowledging their limitations as tools.

By its nature, game writing demands a specialised and evolving array of technological tools and creative approaches to balance dramatic agency and authorial control to provide a coherent and meaningful narrative experience to players (Smed et al. 2021). In broader terms, systems offering interactive narrative experiences consist of procedural components such as choice systems and world rules, narrative design elements including potential events or possible dialogues and mechanics to link them, and narrative vectors or major story directions determined by developers, as well as other narrative assets and user interfaces shaping the overall experience (Koenitz 2023). Despite their shared characteristics, different game development and writing tools may be more suitable for different storytelling requirements. The spatial exploration, navigation, and collision logics inherent in mainstream game engines such as Unity (2025) should be expected to impact the narrative and/or dramatic form and intent of the crafted story, whereas specialised tools like Twine (2025) and Articy *Draft* (2025) allow authors to craft stories with multi-sequential narratives following the principles of hypertextuality (Ensslin, Nelson 2025). Each tool provides game writers various options, including story mapping and dialog systems for crafting non-linear branching stories, and character development tools with attribute variables and trackers linking narrative structures to game actions and states. Overall, the specific affordances of game authoring tools play a key role in shaping what is written and how it is written. In this regard, it is reasonable to expect that

the integration of LLMs will both challenge and transform contemporary game writing practices, introducing novel approaches to narrative creation and collaboration.

EMERGING TRENDS IN LLM-INTEGRATED GAME NARRATIVES

The integration of LLMs into games is a growing field with multiple experimental studies investigating various approaches and techniques (Gallotta et al. 2024, Yang et al. 2025). Generation of new narrative content, their functioning as conversational agents, and the combination of these capabilities in the form of interactive stories are amongst the most actively explored areas of LLM integration into game stories (Yu et al. 2025). The shared core promise of these visions is the expectation to offer high levels of dramatic agency to players (Farokhmanesh 2024; Warley 2024). While studies show that players introduce interesting and unpredicted directions to stories while interacting with LLMs, systems are not always capable of linking them to actionable solutions, leading to player confusion (Peng et al. 2024). There is no consensus on how LLM-driven coherent interactive narratives may be achieved, and it is unknown how LLMs will impact the narrative quality of games (Dring 2024). Therefore, to discuss the possible directions for LLM-integrated game writing practices, we must first give an overview of emerging trends and experimental uses of LLMs in game narratives.

While earlier research primarily investigated the use of LLMs as assistive tools for generating quests, characters, or dialogue to be used in a traditional fashion in games development, both digital and tabletop, recent studies increasingly explore the ways of direct integration of LLMs into game systems, enabling real-time, player-facing interactions (Yang et al. 2025). This shift reflects a growing interest in how LLMs can function not only as creative support but as active components of interactive narrative experiences. The creation of NPCs that players can engage with in completely open-ended conversa-

tions is one of the most prominent visions for LLM integration (Cox, Ooi 2023). A survey by the gaming AI company InWorld (2023), whose authoring platform will be discussed later in the paper, reports that 99% of gamers believe that such NPCs would positively impact gameplay. Early public implementations of LLM-driven fictional characters built on emerging AI systems can be found in a variety of applications.

Utilising the AI framework of StarPal, The Lizzy Bennet interactive AI Avatar at Jane Austen's House Museum exemplifies how LLMs trained with specialised knowledge bases can act as characters from literary classics while functioning as topical experts (GAIN, 2024). Similarly, a seasonal AI-driven version of the fictional character Darth Vader in the online action video game *Fortnite* serves as both a combat companion and a conversational character. Using Google's Gemini, 2.0 Flash model, this version can engage in short dialogues with players, answering questions or responding to remarks while maintaining a diegetically coherent persona grounded in *Star Wars* lore where the character originates, while also being informed by extradiegetic, real-world knowledge on issues like gaming culture (Fortnite 2025). Built on LLMs' widely recognised ability to role play (Shanahan et al. 2023), both applications appear to be able to mimic the established personalities of these fictional characters, while being positioned in relatively static roles designed for repeating interactions within specific contexts.

Proof-of-concept demos by InWorld explore the feasibility of LLM driven characters within specific scenarios derived from cyberpunk and detective fiction. In the now-delisted demo *InWorld Origins* (2023), players are expected to interact with characters in a crime scene to collect clues and solve the case through open-ended dialogues. Steam reviews of the demo suggest that despite experiencing technical limitations, players responded positively to the concept. Multiple reviewers noted the enjoyment they found in testing the system's conversational boundaries and discovering

unexpected responses, including those breaking the narrative (Ibid). Presented as a complex LLM-driven interactive narrative experience, the digital installation project *YELL* by Huang et al. (2025) also uses InWorld's framework. Engaging in conversations with a man across three different periods of his life, visitors are invited to help him process his emotions and memories. The story of personal transformation at the centre of the experience is achieved through three distinct LLM-driven characters, each representing a different version of the protagonist and each holding a personal secret that visitors must uncover. At its core, this structure resembles the interrogation of multiple witnesses, each offering a fragment of the full picture in *InWorld Origins*, but by reframing the technology and narrative architecture *YELL* achieves a layered narrative effect about one complex character. Overall, these mainstream examples illustrate not only the current limitations of LLM-driven characters, but also the importance of the context of their use, and how these factors can inspire innovative narrative design solutions.

While mainstream examples focus on a single or a limited number of LLM-integrated NPCs with equal importance, discussing future directions for LLM-integrated NPCs, Gallotta et al. (2024) differentiate between foreground and background NPCs. According to them, foreground NPCs are characters that are expected to play a key role in the narratives envisioned by game writers, such as a companion character who constantly follows the player, and thus are developed accordingly. As such, LLM-generated dialogue for these characters should take into account not only the narrative context, but also the player's past actions and choices. In contrast, background NPCs are characters primarily considered to be decorative, such as pedestrians in a virtual street who have no other role in the story other than populating that street. Thus, their dialogue generation is expected to only serve the believability of the game world and reflect their general social role within it. This distinction

echoes established conventions regarding the hierarchical complexity of fictional characters and how it is reflected in their contributions to story structures (Chatman 1978). However, as analyses of live-action role-playing games (Brind 2020) and narrative retellings of tabletop role-playing campaigns (Skorkowski 2024) demonstrate, high levels of player agency supported by the flexible narrative facilitation can blur or even collapse these distinctions. In such player-driven contexts, background NPCs may evolve into central supporting characters. Recent fan-created game modifications that integrate LLMs into open-world digital RPGs suggest that this dynamic may also be possible in LLM-integrated game narratives. For example, the mod *Mantella*, which enhances the digital RPG *Skyrim*, links every NPC in the game to an LLM, allowing for open-ended dialogue with each character (2025). According to its developers, this system enables players to convince any NPC to become a companion, and even create quests for them, regardless of their initial role in the game (Ibid). In this regard, LLM integration may lead to a re-evaluation of narrative conventions in games in terms of background character design and could direct game designers and writers to pay equal attention to each character.

Multi-agent behaviours, where different NPCs interact with one another, is another aspect of LLM integration into games (Gallotta et al. 2024). In *Mantella*, this phenomenon manifests in spontaneous conversations between LLM-driven characters. Reflecting each NPC's assigned role within the game world, these exchanges function as a complex and unscripted version of the ambient background chatter common in open-world RPGs, revealing not only worldbuilding details but also micro-stories. While their research focuses on the automatic generation of cartoon episodes based on established franchises using LLMs in conjunction with image-generation models, Maas et al. (2023) demonstrated that multi-agent interactions can be used to generate a complete story from a basic synopsis. Their

findings suggest that multiple agents acting in coordination towards a narrative goal could enable new options for LLM-integrated interactive storytelling involving complex group dynamics.

Park et al. (2023) on the other hand experimented with multiple LLM agents populating a sandbox game world inspired by *The Sims* series. These agents produced believable individual and social behaviours, including forming new acquaintances, going on dates, and collaborating toward specific goals. Traditionally, sandbox games as simulations encourage players to perceive emergent stories (Adams, 2019). The tendency of players to see connections and patterns that may or may not exist plays a major role in their experience of these stories (Fijak, Stokalski 2019). Studies on the retellings of *The Sims* has shown that the game can have profound emotional impact thanks to its social mechanics, even though the characters communicate in gibberish (Eladhari 2018). While natural language interactions with LLM-driven NPCs in sandbox environments can add an additional level of complexity to players' experience, it should be noted that they may also reduce the overall emotional impact by reducing the desired ambiguity. This potential challenge should be addressed and evaluated while designing future sandbox games with multi-agent NPCs driven by LLMs.

A third type of LLM-driven NPC identified by Gallotta et al. (2024) is the commentator or reteller. These characters track the player's actions and narrate past events or discuss current behaviours within the game. This type of NPC plays a unique narrative role, shaping player perception through dynamic contextualisation of actions. The most direct example is a sportscaster in a sports game who provides real-time commentary on player performances (Ibid). Expanding on this concept, Mieschke and Radicke (2023) propose using LLM-based retellers as adaptive in-game radio systems, dynamically generating news reports, interviews, and advertisements based on the player's actions and the broader narrative context. They

describe a scenario in which a reporter interviews a civilian about the aftermath of a mission carried out by a player-controlled vigilante, both summarising the player's actions and contextualising them within the game's sociopolitical environment. Similar reactive radio systems based on traditional branching dialogue structures can be found in earlier titles as well.

In the *Fallout* series, for example, radio hosts respond to players' moral choices, shifting their tone and commentary based on in-game behaviour. This tonal shift is not merely cosmetic; according to Iversen (2012), the system acts as a central narrative device, creating a sense of relational and personal continuity that extends beyond mechanical ways of progression such as levelling up. These retellings highlight the broader consequences of player actions and enhance immersion by reinforcing the game's worldbuilding. LLM-driven retellers, designed with similar narrative goals, may replicate and even amplify this function, as they can in theory respond dynamically to any possible game state. Rist (2024) proposes another use of this NPC type focusing on the commentator aspect. Inspired by games like *Firewatch* and *Stanley Parable*, his iVoice system acts as an inner voice which reacts to a player's movements in a virtual space. His prototype positions the player as an investigator exploring a crime scene. The LLM-driven inner voice generates either instructions for the player or provides sarcastic comments as the mind of a hard-boiled detective. As a proof of concept, the system lacks the narrative complexities of the games which inspired it, but it demonstrates that LLMs can voice different narrative styles while reacting to player actions in real time. Further research and experimentation can lead to more mature ways of using internal narrative voice in games through LLMs.

Besides the different approaches to NPCs, the emergence of LLMs has also led to the reinvigoration of text-based narrative games, resembling traditional text-based adventure games in which players use text commands to control characters

and influence their environments. Traditionally, text adventures are built on two main components: the world model, which runs a simulated world and its inhabitants according to a series of rules; and a parser, which translates textual player inputs into commands the world model can process and then translates back outputs from the world model into textual narratives (Montfort 2003). In other words, the set-up of the game-world and the vocabulary of the parser shape the limits of traditional text adventures. Players are expected not only to overcome challenges in the game world but also to recognise possible actions they can take through its narrative descriptions (Ingold 2011). Overtaking the role of the parser and replacing the rule-based world model with probabilistic generation, LLMs add to the genre not only an enhanced natural language perception capability but also a more flexible game world which players can explore and influence more freely. However, the key challenge and desired characteristic of LLM-integrated text adventures is not just the unique and personalised adventures they generate at each session, but also these being narratively coherent.

The LLM-integrated text adventure platform *AI Dungeon* allows players to generate adventure seeds in different settings which are then shaped by players' commands. In the current version of *AI Dungeon*, players can describe what their characters do and say, but they also have the option of inserting lines into the story or rewinding and erasing earlier generated text (2025). Due to the platform's adaptive capabilities, character actions and the insertion of story lines may lead to inconsistent and out-of-context situations, such as using a laser gun in a mediaeval setting. The rewind and regeneration options give players the capability to keep their stories consistent if they wish to do so. Mitchell (2022) argues that these functions are designed to evade the potential strangeness in the experience and prevent a player's critical stance towards the limitations and characteristics of the system. While *AI Dungeon* experiments with

both different models and training data, as a platform it currently cannot automatically provide coherent interactive stories. However, smaller, specialised projects seem to have come closer to this goal.

HistoryLens simulates a day in the lives of historical figures each with a specific challenge (Breen 2024). The players' actions are shaped not only by the limitations of these characters' circumstances but also by the educational setting of the project, which dictates a certain protocol for interactions. Similarly, the indie game *Katanas and Kimonos* also requires the player, who takes the role of a ronin, to consider the cultural and historical limitations of its setting when entering commands, and if a player ignores them, the game requests alternative inputs (Ross 2023). Both projects achieve this through detailed prompt engineering, which will be discussed below. *Eternagram* invites players to assist a character in recovering their memory while exploring alternate climate scenarios (Zhou et al., 2024). While the game does not direct player input, its narrative progresses through a series of intermediate goals identified by designers and activated by specific player inputs, or narrative triggers. This structure is inspired by the interactive drama *Façade*, where player interactions dynamically change the story's direction, as they influence the order in which the underlying narrative units are presented (Mateas, Stern 2005). While the game's story flows in a specific direction, the details of how its destination is reached is up to the players. Telling a time travel story with the player character tasked to change history, *EcoEcho* features multiple LLM-driven NPCs who the players must convince to achieve their goals alongside an overarching branching story with predetermined endings (Zhang et al. 2025). *EcoEcho* uses this layered system to preserve narrative coherence. Player inputs corresponding with the themes and goals of the story lead to natural and flexible conversations with NPCs. Off-topic inputs, on the other hand, first cause NPCs to try steer the exchange back through emotional

cues or guiding questions, and, if progress stalls, they default to predefined responses to move the story forward. All in all, in their current state LLMs cannot provide coherent interactive narratives consistently without designer interventions. Techniques to balance the free-form conversation and narrative coherence are still being developed and the field is open for further experimentation.

WRITING LLM-INTEGRATED GAMES

On the surface, LLMs' generative functions seem to often overlap with the classical roles of game writers. Creating plots and characters for classic games mirrors the work of narrative directors, while generating NPC dialogue aligns with the responsibilities of game writers. However, the methods and discourse around the integration of LLMs into game writing is dynamic and far from settled. Not only are the models themselves advancing, but there is also growing momentum in the industry to integrate LLMs into game development, including narrative and dialogue systems, despite the challenges regarding high costs, uncertainty about best practices, and balancing automation with creative control (Guarracino 2025). Combined with ongoing academic research utilising open and commercial models, this push is catalysing experimentation and innovation towards multiple new modes of game writing and narrative design. Consequently, any analysis of this rapidly evolving field must be understood as provisional, reflecting a moment within an ongoing transformation while best practices, standards, and theoretical frameworks still emerging. With these limitations in mind, several current challenges related to the utilisation of LLMs in interactive game narratives and evolving approaches to overcoming them can be identified.

The tendency of base LLMs to hallucinate, producing content that is baseless, internally inconsistent, or unverifiable, is a widely recognised issue (Alansari, Luqman 2025). One contributing factor is the inclination of LLMs to follow player cues with-

out resistance, even when these contradict the established setting or narrative logic (Breen 2024; Sun et al. 2025). In such cases, hallucinations can disrupt narrative coherence and limit prolonged, consistent interactions. However, when hallucinated content introduces new but contextually fitting details, it may be seen not as a flaw but as a generative feature that enriches the story world (Gallotta et al. 2024; Wissigkeit et al. 2025). Bias is another well-documented concern. While negative bias can result in problematic or exclusionary content, positive bias may produce overly sanitised or optimistic responses, especially limiting in genres that rely on moral ambiguity or historical nuance (Breen 2024; Buongiorno et al. 2024). Generic or repetitive content generation also remains a challenge, further underscoring the importance of narrative framing by writers and designers (Gallotta et al. 2024). To address these issues, a range of strategies has emerged. Computational methods for interpreting and guiding the internal workings of LLMs are still in their infancy, though they show promise in surfacing specific topics, even when these are out of context for a given exchange (Templeton, Conerly 2024). Various forms of fine-tuning, meaning the further training of general-purpose models on smaller, domain-specific datasets, are widely used to improve performance on targeted narrative tasks (Kelttek 2024). Meanwhile, prompt engineering – that is, shaping LLM outputs through well structured, detailed, and goal-oriented inputs – remains the most accessible and widely adopted technique for influencing LLMs (Oppenlaender 2023). As research on how human-led game writing and LLMs can and need to function together continues, these methods and the tools built around them are becoming integral components of a game writer’s evolving toolbox.

The recognition of base models’ limitations in contextual understanding and the ability to adapt to player-driven interactions in specific game settings has led to the evaluation of fine-tuning as a solution. While most studies have focused on the

implementation and impact of fine-tuning, from a game-writing perspective the source of the data used for the process can be considered an equally important question as well. Researching utilisation of LLMs in TRPGs, Santiago III et al. (2023) suggested leveraging data from hundreds-of-hours-long *Critical Role* campaigns, a popular TRPG actual play web series, for fine-tuning, emphasising the importance of context and player-focused dynamics. Van Stegeren and Myśliwiec (2021) also explored how fine-tuning could impact quest generation by using a detailed data set consisting of thousands of *World of Warcraft* quests, originally collected, organised, and shared by fans on the online Wowhead database. The study revealed that fine-tuning allowed the system to generate quests that felt authentic to the *World of Warcraft* environment but were also constrained within those boundaries. Further focusing on how data should be organised for fine-tuning LLMs for quest generation, Kelttek (2024) also used fan-created content. Aiming to generate quests for *Assassin’s Creed Odyssey*, he first collected quest entries from the *Assassin’s Creed* wiki and organised them into a coherent dataset. He then processed this dataset to generate a synthetic version suitable for fine-tuning. While the initial data allowed him to analyse and represent the game’s complex narrative structure, the second step reformatted it for model training. His analysis revealed that while fine-tuned models showed potential for more specialised content generation, base models often outperformed them across several evaluation criteria. *Mantella* also employed fine-tuning (2025), using a dataset of over 8,800 player-NPC interactions extracted from *Skyrim*. This dataset is used to train LLMs to generate natural dialogue dialogs that reflected the conversational styles and patterns of the vanilla game. From a game writing perspective these examples show how existing extensive data linked to established franchises can be used to fine-tune LLMs to imitate these source materials’ style.

Mantella is potentially a best-case scenario for this approach, since, as a mod, it uses dialogues written for *Skyrim* to enable conversational NPCs in *Skyrim*. Disregarding any exploitative practices on the use of copyrighted content, fine-tuning LLMs for original game narratives will potentially require new approaches. Handcrafting, generating synthetic data, or combining these together are options currently evaluated by the industry (AI Dungeon 2025). Experiments on crowdsourced worldbuilding for LLM fine-tuning have also shown some promise (Urbanek et al. 2019). Building on Hongisto (2025), the curation of genre or franchise text can also be a direction. While it is unclear how game writers will be able to contribute to this process, an understanding of narrative logic and organisation will potentially be required to generate and evaluate data for fine-tuning.

Prompt engineering refers to the design of instructions to steer LLMs towards specific desired responses (Knoth et al, 2024). According to Eager and Brunton (2023) prompts should describe the specific actions to be performed by LLMs and their outcomes, the context of these actions, the subject matter, constraints, and limitations, and finally, how the content aligns with the desired goal. In other words, a good prompt tells the model who it is, what it should do and produce under which conditions, and gives it rules to ensure that the output aligns with the user's goal. There are different approaches in prompt engineering (Harmon, Rutman 2023). The level of prompt specificity influences the degree of control, coherence, and fidelity to a desired tone or structure. In general, more specific, detailed prompts tend to direct the LLM to follow instructions closely, while very brief or open-ended prompts can produce more surprising or creative outputs. In any case, the prompt must be written with the intended goal and context in mind, since even minor changes in wording or detail can significantly affect the response.

In game writing, prompts can establish character identities, goals, and relationships to provide contextual understanding

and directions to LLMs to respond certain conditions in certain ways (Csepregi 2021). Rao et al. (2024) suggest that prompting should introduce distinctive personas, backstories, goals, and reactions for each character, and support the expected flow of events. While a properly engineered single prompt can direct LLMs to run an interactive text adventure, complex prompts can also be part of larger systems. The size of the project and how LLMs are integrated to the narrative experience can change the prompting strategy drastically.

Katanas and Kimonos, for example, is a detailed but relatively short prompt establishing conditions, rules, setting, and characters for a text-based adventure. *HistoryLens*' prompt establishes the rules and objectives of the game as well as the historical setting, while also providing historical texts to the LLM, which supports the generation of a dynamic world for the player to explore. *Eternagram* uses a two-level prompting approach. It first supplies the LLM with a detailed context including world data, NPC role data, and story background. These elements indicate the details of the narrative and manage the behaviour of the NPC during the interactions. The second lists a series of triggers the LLM is instructed to react to if detected in the players' input. Such triggers, like a question about the NPC's origin, drive the narrative progression, acting as milestones from one stage of the story to the next. This two-tier prompt design ensures that the LLM both understands the game world and responds to the player appropriately. *Mantella* (2025), on the other hand, uses a starting prompt containing detailed background descriptions of over 1,000 *Skyrim* NPCs taken from the *Skyrim* wiki to ensure that the LLMs' responses and interactions align closely with the established lore of *Skyrim*. Even though this prompt is data-heavy, it is just one part of the system. *Mantella* pairs the LLM with *Skyrim*'s normal game engine and rules. The mod's architecture links speech-to-text, the LLM, and text-to-speech with *Skyrim*'s original mechanics, so that NPCs react both to dialogues and to

the underlying game state. Overall, prompt engineering in game writing is a versatile tool. It lets writers leverage their skills in worldbuilding, character development, and narrative design when working with LLMs. By carefully designing prompts and multi-prompt systems to encode character goals, histories, and conversational constraints, game writers can create rich, context-sensitive interactions between players and LLM-driven characters. However, the field of prompt engineering is still evolving. Researchers are continually developing new and advanced prompt engineering techniques to improve LLMs' performance (Yang et al. 2025). In this regard, adaptability should be a core priority for game writers.

In recent years, several game authoring tools have been introduced to support writers in developing components for LLM-driven game narratives. These tools integrate prompting capabilities with narrative planning mechanisms that in some cases resemble traditional branching story structures. Their primary advantage lies in offering author-friendly interfaces that foreground the human craft of narrative design, making the development of LLM-integrated storytelling more accessible and creatively manageable for game writers. In other words, they are designed as interfaces between game writers and LLMs.

Convai and the previously mentioned InWorld provide two prominent commercial LLM-integrated authoring platforms which focus on conversational NPC generation. They provide game writers with the tools to shape character behaviour and dialogue without scripting dialogue trees. In both systems, game writers provide background material, such as backstory, personality traits, and knowledge about the game world, to achieve coherent character identity and consistent tone. Compared to Convai, InWorld places a stronger emphasis on categorical structure. Writers are required to enter each narrative element into its designated section in the platform to prevent unpredictable or out-of-character outputs, giving clues on how

the internal mechanisms of the system organises narrative content as prompts. While they differ in specifics, personality modelling based on predetermined system parameters is also a shared function in both authoring environments and shows how they aim to standardise author input to achieve desired LLM behaviour. Beyond shaping characters, Convai also includes a narrative design module centred around high-level goals for NPCs. These narrative objectives shape how they interact with players. While no dialogue exchanges are expected to be authored by game writers, they are still responsible for defining potential events and character reactions in response to player actions. Further afield, the Drama Llama authoring environment, which was developed as a research prototype, also explores how LLMs can be used in narrative management (Sun et al. 2025). In this system, besides descriptions of the game world and characters, game writers are expected to define key potential actions activated under specific conditions. Early experiments have shown that a limited number of well-defined actions may lead to the generation of engaging simulations (Ibid). Overall, contemporary LLM-integrated game writing tools aim to have LLMs function like improv actors directed by prompts. The focus of the writing shifts to the minds, emotions, and internal motivations of characters, rather than storylines and dialogues demonstrating these. These tools thus represent a reorientation of game writers' primarily responsibility as worldbuilders, character developers, and narrative directors.

CONCLUSION AND FUTURE DIRECTIONS

The integration of LLMs into game narratives presents both a technological and conceptual shift in the craft of game writing. A reevaluation of its already hybrid characteristics, rooted both in the art of storytelling and in computational design, is required. Emerging strategies like fine-tuning and prompt engineering as well as upcoming authoring tools redefine the

role and responsibilities of game writers. Designing the conditions under which stories emerge, and the characters portrayed by machines for and with players, appears to be the key focus. At the same time, the potential adoption of LLMs also raises questions about authorship, coherence, control, and ethical design. Hallucination, bias, and repetitiveness remain unresolved technical and creative challenges. The use of proprietary or community-generated data for early fine-tuning experiments complicates the relationship between inspiration, imitation, and intellectual property. Prompt engineering, though powerful, is still a developing expertise. In other words, there are still unknowns when it comes to LLMs and their application in game writing.

As the field matures, further research is needed into how writers can meaningfully shape the narrative behaviours of LLMs within games. Given the rapid evolution of models as well as approaches and tools in LLM-integrated game writing, such research must remain adaptive to ongoing technological change. A structured comparative evaluation of frameworks can assess both the co-creative processes and

the quality, coherence, and creativity of LLM-generated narrative content. Studies exploring player perception and reception of LLM-driven interactive narratives will provide further evidence on their effectiveness. Investigations into how writers, narrative designers, and players engage with LLM-integrated systems in practice and how this relates to the broader transformation of the game industry itself is also required to clarify how authorship, responsibility, and collaboration are being renegotiated in real production environments. The integration of LLMs into game writing is not simply a matter of automation, but of augmentation. In this regard, questions of ownership, consent, and attribution must be addressed to identify and develop best practices for ethically grounded co-creativity between humans and machines. Future studies building on earlier research in creativity support as well as co-creation, not only in game writing but also other forms of computational creativity, would benefit the development of critical and theoretical frameworks to discuss and evaluate the development of LLM-integrated interactive narrative experiences.

ACKNOWLEDGMENTS

This research is part of the AI and Games Narratives Project, which has been supported by the University for the Creative Arts (IRF2024/25/2).

REFERENCES

- Aarseth, Espen J** 1997. *Cybertext. Perspectives on Ergodic Literature*. London.
- Adams, Tarn** 2019. Emergent Narrative in Dwarf Fortress. Tanya X. Short, Tarn Adams (eds.), *Procedural Storytelling in Game Design*, A K Peters/CRC Press, 149–158.
- Alansari, Aisha; Luqman, Hamzah** 2025. 'Large Language Models Hallucination: A Comprehensive Survey', <https://www.arxiv.org/abs/2510.06265> (10 October 2025).
- Articy Homepage** <https://www.articy.com/> (30 May 2025).
- AI Dungeon Homepage** <https://aidungeon.com/> (30 May 2025).
- Buongiorno, Steph; Klinkert, Lawrence, J.; Chawla, Tanishq; Zhuang, Zixin; Clark, Corey** 2024. 'PANGeA: Procedural Artificial Narrative using Generative AI for Turn-Based Video Games', <https://arxiv.org/abs/2404.19721> (30 May 2025).
- Breen, Benjamin** 2024. 'Large Language Models, Experiential Learning, and AI Simulations in the Classroom: The HistoryLens Framework' -Teaching History 49(1), 10–33.
- Brind, Simon** 2020. Learning from NPCs. Eleanor Saitta, Johanna Koljonen, Jukka Särkijärvi, Anne Serup Grove, Pauliina Männistö, & Mia Makkonen (eds.). *What Do We Do When We Play?* Helsinki; Solmukohta, 44–48.
- ChatGPT Homepage**. <https://chatgpt.com/> (3 February 2023).
- Chatman, Seymour** 1978. *Story and Discourse*. London.
- Conv.AI Homepage**. <https://convai.com/>. (30 May 2025).
- Cox, Samuel, R., Ooi, Wei Tsang** 2023. 'Conversational Interactions with NPCs in LLM-Driven Gaming: Guidelines from a Content Analysis of Player Feedback' – Asbjørn Følstad, Theo Araujo, Symeon Papadopoulos. Effie L.-C. Law, Ewa Luger, Morten Goodwin, Sebastian Hobert, Petter Bae Brandtzaeg, – *Chatbot Research and Design: 7th International Workshop, CONVERSATIONS 2023, Oslo, Norway, November 22–23 2023, Revised Selected Papers, Lecture Notes in Computer Science, vol 14524*. Cham, Springer, 167–184.
- Csepregi, Lajos M.** 2024. 'The Effect of Context-aware LLM-based NPC Conversations on Player Engagement in Role-playing Video Games'. https://projekter.aau.dk/projekter/files/536738243/The_Effect_of_Context_aware_LLM_based_NPC_Dialogues_on_Player_Engagement_in_Role_playing_Video_Games.pdf (30 May 2025).
- Dring, Christopher** 2024. 'Ubisoft debuts NEO NPC AI prototypes |GDC 2024, 19 March 2024, <https://www.gamesindustry.biz/ubisoft-debuts-neo-npc-ai-prototypes-at-gdc#:~:text=Dubbed%20NEO%20NPCs%2C%20Ubisoft's%20efforts,alongside%20its%20Production%20Technology%20department> (30 May 2025).
- Eager, Bronwyn, Ryan, Brunton** 2023. Prompting Higher Education Towards AI-Augmented Teaching and Learning Practice, *Journal of University Teaching and Learning Practice*, 20 (5), <https://open-publishing.org/journals/index.php/jutlp/article/view/670> (30 May 2025).
- Eladhari, Mirjam P.** 2018. Re-Tellings: The Fourth Layer of Narrative as an Instrument for Critique – Rebecca Rouse, Hartmut Koenitz, Mads Haahr (eds) *Interactive Storytelling*. ICIDS 2018. *Lecture Notes in Computer Science*, vol 11318. Cham, Springer, 65–78.
- Ensslin, Astrid** 2007. *Canonizing Hypertext: Explorations and Constructions*. London.
- Ensslin, Astrid; Nelson, Jason** 2025. 'Co-creative multimodal authorship as procedural performance with DALL-E' – Will Slocombe, Genevieve Liveley (eds.), *The Routledge Handbook of AI and Literature*, Routledge, New York, 315–331.
- Farokhmanesh, Megan** 2024. 'Is AI the Future of NPCs?', *Wired Magazine*, <https://www.wired.com/story/artificial-intelligence-ai-non-player-characters-npcs-future/#:~:text=Artificial%20intelligence%20is%20poised%20to,might%20be%20AI's%20killer%20app> (30 May 2025).
- Fijak, Marta; Stokalski, Jakub** 2019. Beyond Fun in Frostpunk. Tanya X. Short, Tarn Adams (eds.), *Procedural Storytelling in Game Design*, A K Peters/CRC Press, 241–256.
- Finley, Toiya K.** 2023. *Branching Story, Unlocked Dialogue Designing and Writing Visual Novels*, Routledge. Boca Raton.
- Fortnite** 2025. 'Speak with Darth Vader in Fortnite', <https://www.fortnite.com/disney#darth-vader-faq> (30 May 2025).
- Freed, Joanne L.** 2025. 'Rethinking intentionality in the era of AI' – Will Slocombe, Genevieve Liveley (eds.), *The Routledge Handbook of AI and Literature*, Routledge, New York, 224–232.
- GAIN** 2024. 'Lizzy Bennett: Interactive AI Avatar', <https://gain.uca.ac.uk/blog/jah/> (30 May 2025).
- Gallotta, Roberto, Todd, Graham, Zammit, Marvin, Earle, Sam, Liapi, Antonios, Togelius, Julian, Yannakakis, Georgios, N.** 2024. *Large Language Models and Games: A Survey and Roadmap*. – *IEEE Transactions on Games*, 1–18.
- Guarracino, Juliana**. 'AI is helping developers supercharge video game characters and make immersive worlds feel more realistic', <https://www.businessinsider.com/ai-video-games-development-faster-production-new-improved-gameplay-2025-8> (5 October 2025).
- Harmon, Sarah, Rutman, Sophia** 2023. Prompt Engineering for Narrative Choice Generation. Lisa Holloway-Attaway, John T. Murray (eds) *Interactive Storytelling*. ICIDS 2023. *Lecture Notes in Computer Science*, vol 14383. Cham. Springer, 208–225.
- Hocking, Clint** 2009. 'Ludonarrative dissonance in Bioshock: the problem of what the game is about'. – Drew Davidson (ed.), *Well played 1.0*. Pittsburgh, PA, ETC Press, 114–117.
- Hongisto, Tuuli** 2025. 'A brief history of computer-generated literature: In search of the author'. – Will Slocombe, Genevieve Liveley (eds.), *The Routledge Handbook of AI and Literature*, Routledge, New York, 36–48.
- Huang, Jiayang, Li, Lingjie, Zhang, Kang, Yip, David** 2025. 'The Dream Within Huang Long Cave: AI-Driven Interactive Narrative for Family Storytelling and Emotional Reflection'. <https://arxiv.org/abs/2504.04968> (30 May 2025).
- Ingold, Jon** 2011. 'Thinking Into the Box: On the Use and Deployment of Puzzles'. – Kevin Jackson-Mead, J. Robinson Wheeler (eds) *IF Theory Reader*, Boston, Transcript on Press, 229–247.
- Inworld** 2023. 'Inworld study on the future of NPCs finds 99% of gamers think AI will enhance gameplay', <https://inworld.ai/blog/future-of-npcs-report> (20 May 2025).
- Inworld AI Homepage**, <https://inworld.ai/> (30 May 2025).
- Inworld Origins Website**, https://store.steampowered.com/app/2199920/Inworld_Origins/ (30 May 2025).

- Iversen, Sara M. 2012. 'In the Double Grip of the Game: Challenge and Fallout 3', *Game Studies The International Journal of Computer Game Research*, volume 12 issue 2, https://gamestudies.org/1202/articles/in_the_double_grip_of_the_game/ (30 May 2025).
- Kantosalo, Anna, Takala, Tapio 2020. 'Five C's for Human-Computer Co-Creativity: An Update on Classical Creativity Perspectives.' – Amilcar Cardoso, Penousal Machado, Tony Veale, Joao Miguel Cunha (eds) *Proceedings of the 11th International Conference on Computational Creativity*, Association for Computational Creativity, 17–24.
- Keltek, Tamer U. 2024. Empowering Narrative Design in Video Games With Large Language Models: A Case Study of Assassin's Creed Odyssey, Bahcesehir University, Unpublished Master Thesis, Istanbul.
- Knoth, Nils, Decker, Marie, Laupichler, Mathias C., Pinski, Marc, Buchholtz, Nils, Bata, Buchholtz, Schultz, Ben 2024. 'Developing a holistic AI literacy assessment matrix – Bridging generic, domain-specific, and ethical competencies', *Computers and Education Open*, vol 6, <https://www.sciencedirect.com/science/article/pii/S2666557324000181> (30 May 2025).
- Koenitz, Hartmut 2023. *Understanding Interactive Digital Narrative Immersive Expressions for a Complex Time*. New York.
- Lasssheikki, Christina 2022. 'Game writers and narrative designers – The evolving role of storytelling professionals in game development.' <https://aaltodoc.aalto.fi/server/api/core/bitstreams/37044d31-528a-4087-9090-17f4bdf65e6f/content> (30 May 2025).
- Maas, Philipp; Carey, Frank; Wheeler, Chris; Saatchi, Edward; Billington, Pete; Shamash, Jessica Y. 2023. 'To Infinity and Beyond: SHOW-1 and Showrunner Agents in Multi-Agent Simulations', <https://fablestudio.github.io/showrunner-agents/> (30 May 2025).
- Manovich, Lev 2002 'Models of authorship in new media': SWITCH 17, no. 3. www.scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1186&context=switch (30 May 2025).
- Mantella, Bring NPCs to Life with AI. <https://www.nexusmods.com/skyrimspeditions/mods/98631> (30 May 2025).
- Mateas, Michael, Stern, Andrew 2005. 'Structuring content in the Façade interactive drama architecture' – R. Michael Young, John Laird (eds) *AIIDE'05: Proceedings of the First AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, AAAI Press, 93–98.
- McRae, Edwin 2020. *Narrative Design for Writers*. Owaka.
- Mieschke, Paul, Radicke, Stefan 2023. 'Adaptive LLM-based Game Radio (ALGR)'. – *Proc. IEEE 4th International Conference on Computers and Artificial Intelligence Technology (CAIT)*, 277–283.
- Mitchell, Alex 2022. 'Repetition and Defamiliarization in AI Dungeon and Project December', <https://electronicbookreview.com/essay/repetition-and-defamiliarization-in-ai-dungeon-and-project-december/> (30 May 2025).
- Montfort, Nick 2003. *Twisty Little Passages: An Approach to Interactive Fiction*. Cambridge.
- Murray, Janet 1997. *Hamlet on the Holodeck. The Future of Narrative in Cyberspace*. Cambridge.
- Oppenlaender, Jonas 2023. 'A taxonomy of prompt modifiers for text-to-image generation', *Behaviour & Information Technology* Volume 43 (15), 3763–3776.
- Park, Joon Sung, O'Brien, Joseph C., Cai, Carrie J., Ringel, Meredith M., Liang, Percy, Bernstein, Michael S. 2023. 'Generative Agents: Interactive Simulacra of Human Behavior', <https://arxiv.org/abs/2304.03442> (30 May 2025).
- Paterson, Eddie, Simpson-Williams, Timothy, Corder, Will 2020. *Once Upon a Pixel: Storytelling and Worldbuilding in Video Games*, Boca Raton.
- Peng, Xiangyu; Quaye, Jessica; Rao, Sudha; Xu, Weijia; Botchway, Portia; Brockett, Chris; Jojic, Nebojsa; DesGarenes, Gabriel; Lobb, Ken; Xu, Michael; Leandro, Jorge; Jin, Claire; Dolan, Bill 2024. 'Player-Driven Emergence in LLM-Driven Game Narrative', <https://arxiv.org/abs/2404.17027> (30 May 2025).
- Rao, Sudha, Xu, Weijia, Xu, Michael, Leandro, Jorge, Lobb, Ken., DesGarenes, Gabriele, Brockett, Chris, Dolan, Bill 2024 'Collaborative Quest Completion with LLM-driven Non-Player Characters in Minecraft', <https://arxiv.org/abs/2407.03460> (30 May 2025).
- Rettberg, Scott, Rettberg, Jill Walter 2024 'Algorithmic narrativity: Literary experiments that drive technology.' – *Dialogues on Digital Society*, 1(1), 37–40.
- Ross, Jakob B. 2023. *Katanas and Kimonos: GM Bot Volume I*, Portland, Oregon.
- Rist, Thomas 2024. 'Using a Large Language Model to turn Explorations of Virtual 3D-Worlds into Interactive Narrative Experiences', <https://ieeexplore.ieee.org/document/10645583> (30 May 2025).
- Santiago III, Jose M., Parayno, Richard L., Deja, Jordan A., Samson, Briane P.V. 2023. 'Rolling the Dice: Imagining Generative AI as a Dungeons & Dragons Storytelling Companion.' <https://arxiv.org/abs/2304.01860> (30 May 2025).
- Skorkowsky, Seth 2024. 'Secrets of the Ancients: Ep 3. The Hunt – RPG Review', https://www.youtube.com/watch?v=8mM_cRjXH8w&list=PL25p5gPY6qKW-7aW5rCx62j7LH_SDgA4_&index=3 (30 May 2025).
- Shanahan, Murray, McDonell, Kyle, Reynolds, Laria 2023 'Role play with large language models.' – *Nature* 623 (7987), 493–498.
- Sharples, Mike, Pérez y Pérez, Rafael 2022. *Story Machines: How Computers Have Become Creative Writers*, London.
- Smed, Jouni, Suovuo, Tomi, Skult, Natasha, Skult, Petter 2021. *Handbook on Interactive Storytelling*, Chichester, West Sussex.
- Sun, Yuqian; Wang, Phoebe, J.; Chung, John, J.Y.; Roemmele, Melissa; Kim, Taewook; Kreminski, Max 2025. 'Drama Llama: An LLM-Powered Storylets Framework for Authorable Responsiveness in Interactive Narrative', <https://arxiv.org/abs/2501.09099> (30 May 2025).
- Templeton, Adly; Conery, Tom 2024. 'Scaling Monosemanticity: Extracting Interpretable Features from Claude 3 Sonnet', <https://transformer-circuits.pub/2024/scaling-monosemanticity/index.html> (30 May 2025).
- The Elder Scrolls V Skyrim Homepage, <https://elderscrolls.bethesda.net/en/skyrim10> (30 May 2025).
- The Writers' Guild of Great Britain 2021. Writing for Videogames: A guide for games writers and those who work with them. <https://writersguild.org.uk/wp-content/uploads/2021/12/WGGB-A4-Videogames-6.12.21.pdf> (30 May 2025).
- Trygg, Natasha Bulatovic., Skult, Petter, Smed, Jouni 2024. Narrative Design. Newton Lee (ed.), *Encyclopedia of Computer Games*, Cham, Switzerland, Springer, 1225–1230.
- Twine Homepage, <https://twinery.org/> (30 May 2025).
- Unity Homepage, <https://unity.com/> (30 May 2025).
- Urbanek, Jack; Fan, Angela; Karamcheti, Siddharth; Jain, Saachi; Humeau, Samuel; Dinan, Emily; Rocktaschel, Tim; Kiela, Douwe; Szlam, Arthur; Weston, Jason 2019. 'Learning to Speak and Act in a Fantasy Text Adventure Game', <https://arxiv.org/pdf/1903.03094> (10 October 2025).

- Van Stegeren, Judith, Myśliwiec, Jakub** 2021. 'Fine-tuning GPT-2 on annotated RPG quests for NPC dialogue generation'. -*Proceedings of the 16th International Conference on the Foundations of Digital Games (FDG '21)*. New York, NY, USA, Association for Computing Machinery, Article 2, 1–8.
- Warley, William** 2024, 'AI in Gaming The Future of Gaming: Personalized Experiences with AI-Driven NPCs', 31 January 2024, <https://medium.com/@williamwarley/the-future-of-gaming-personalized-experiences-with-ai-driven-npcs-bd1054738188> (30 May 2025).
- Wissigkeit, Tobias, Lope-Nicholls, Daniela, Stindl, Christian, Iurgel, Ido A.** 2025. Exploring Phantasmagorical Game Design: Integrating Anomalous Behaviour of AI and Player Into Conversational AI-Native Games, Eva Brooks, Emma Edstrand, Anders Kalsgaard Møller, Thomas Bjørner (eds) *Design, Learning, and Innovation 9th EAI International Conference, DLI 2024, Virtual Event, November 7–8, 2024, Proceedings*, 120-135.
- Yang, Daijin; Kleinman, Erica; Hartevelde, Casper** 2025. 'GPT for Games: An Updated Scoping Review (2020–2024).' – *IEEE Transactions on Games*, 04, 1–16.
- Yu, Pengfei; Shen, Dongming; Meng, Silin; Lee, Jaewon; Yin, Weisu; Cui, Andrea Y.; Xu, Zhenlin; Zhu, Yi; Shi, Xingjian; Li, Mu; Smola, Alex** 2025. 'RPGBENCH: Evaluating Large Language Models as Role-Playing Game Engines,' <https://arxiv.org/abs/2502.00595> (30 May 2025).
- Yuan, Ann, Coenen, Andy, Reif, Emily, Ippolito, Daphne** 2022 'Wordcraft: Story Writing With Large Language Models. – *Proceedings of the 27th International Conference on Intelligent User Interfaces (IUI '22)*. New York, NY, USA, Association for Computing Machinery, 841–852.
- Zhang, Qinshi; Wen, Ruoyu; Hendra, Latisha B.; Ding, Zijian; LC, Ray** 2025. 'Can AI Prompt Humans? Multimodal Agents Prompt Players' Game Actions and Show Consequences to Raise Sustainability Awareness,' <https://arxiv.org/abs/2409.08486> (30 May 2025).
- Zhou, Suifang; Hendra, Latisha B.; Zhang, Qinshi; Holopainen, Jussi; LC, Ray** 2024. 'Eternagram: Probing Player Attitudes Towards Climate Change Using a ChatGPT-driven Text-based Adventure,' <https://dl.acm.org/doi/10.1145/3613904.3642850> (30 May 2025).