

AI in the Game Industry: Latest Research Advances and Future Direction

Sibo Song

*Beijing 21st Century School, Beijing, China
jacksong090121@gmail.com*

Abstract. The paper is a summary of the current developments of AI in the gaming sector between 2023 and 2026, including automated generation of content, intelligent game agents (NPCs) and large language models (LLMs), game testing and quality assurance by means of AI, modeling of players and dynamic difficulty management, as well as reorganization of the industry and the new concept of the so-called AI-native games. This study reveals that the use of AI will not be used to assist in the creation of contents but rather become part of the core elements of game mechanics. Real-time diffusion models can model game engines (GameNGen), LLM agents are already implemented in commercial QA systems (TITAN framework), reinforcement learning and player modeling can be applied to adjust the difficulty of the game dynamically based on the data. At the same time, there exists a discussion in the field of efficiency of the process, the creative aspect of it, and how much restructuring of the organization should be done. The authors also suggest some of the criteria which can be used to determine whether a game is an AI-native game (counterfactual testing) and indicate the open issues like automatic playability tests, user acceptance, safety of the contents and ownership of the copyright.

Keywords: Artificial Intelligence (AI), Large Language Model (LLM), Game Agents, AI-native games

1. Introduction

The gaming sector is now in the process of AI transformation. The AI was previously a means to automate the production of content that could be used by developers to assist them during the game creation, and this has changed as it is anticipated to form one of the elements of the core gameplay mechanics and be incorporated within the gameplay itself. The four technical fields are demonstrated to have advanced with the help of research between 2023 and 2026. These consist of player experience, intelligent characters, quality assurance and content generation. Diffusion models can be used to create a real-time engine for games. It is possible to make the NPC dialogue and behavior more human-like in principle with the aid of large language models and at the same time LLM agents will become part of the commercial QA pipelines, which should enhance the efficiency of testing and decrease the number of manual activities performed. There is also the reinforcement learning and player modeling approach that has shifted the dynamic difficulty control to data-based

approach instead of using heuristics. This paper may demonstrate the efficiency gain at the industry level but on the other hand there will be some controversy of creativity this paper cannot ignore.

The newest frontier consensus is the notion of "AI-native games". Indeed, generative AI is not just accelerating production in the existing pipeline anymore, but it is starting to redefine the very nature of games in a fundamental way. This study makes the following main judgements in this review:

This paper see the technical trajectory. The role of AI may range between automated PCG content generation and LLM-driven intelligent NPCs, and then to AI-native games which are based on generative AI as the core mechanism. Also, the game development is going from hand-coded production to Development 3.0. It should be mentioned that human developers can change their roles into supervision and high-level decision-making [1, 2].

Deep generation can be used as a game engine: GameNGen is able to simulate DOOM in real time at 20 frames per second with the help of diffusion model, and frame PSNR can reach 29.4 which is comparable to lossy JPEG. It should be mentioned that human raters cannot easily differentiate between simulated frames and the actual game [3]. In addition, playable levels of games can be generated by large language models in data-driven manner. The performance may increase with dataset size [4].

NPC dialogue can be enhanced, and behaviors can become more realistic using LLMs. Indeed, in small-town simulation, generative agents planned parties and distributed invitations without scripting, and this demonstrates plausible social actions [5]. Nevertheless, the cost of deployment and real-time responsiveness remain bottlenecks, this study employs learned-policy substitution to decrease the cost, and human-like memory mechanisms are used to address the issue of memory problem [6, 7].

This paper believe that AI game testing is feasible in real production environment. As a matter of fact, TITAN framework will reach 95 percent task-completion rate on LLM agents on a large commercial MMORPG and has been implemented across eight QA pipelines already so the pressure of testing due to high-frequency updates under games-as-a-service model will be mitigated [8].

Player modeling has transformed DDA into a data-driven approach as opposed to heuristic rules. Also, this study applies reinforcement learning to train adaptive opponents that are able to match the skill of players [9], and A/B tests in real mobile games show that difficulty-aware interventions can improve retention and engagement [10].

Industry evidence demonstrates that there is a convergence though tensions persist and the majority of people are aware that efficiency gains exist and can be quantified in numerous situations, however, independent and frontline developers may have reservations on creativity of generative AI, quality of content, copyright concerns and also job risks. Actually, the size of gains may rely on the depth of organizational change concerning AI [11-13].

Now this paper is seeing the rise of the new trend that is referred to as AI-native games. In this case, runtime generative AI is required in core gameplay loop. This study also consider a counterfactual: would it be possible to lose the core game without the generative AI? This paper will analyze 53 titles, and the result will reveal an increasing number of games meeting this condition. Nevertheless, there are still some problems left. The playability testing can be done automatically, the players have to accept the content, the ownership of the content should be determined, but these questions are yet to be examined [2, 14].

This study looks into the frontier literature of 2023 to 2026, especially the 2024-2026, on the topic of AI in game industry and this paper review academic articles as well as best conference papers. This study also has some surveys and empirical studies of industries. The sources are both

Chinese and English. There are various angles of the search strategy. The content creation with PCG and generative AI is considered intelligent NPCs and game agents based on LLM, game testing and quality control by RL and LLM testing. Also, this paper analyzes the player modeling and dynamic difficulty (DDA). Then it comes to the controversies and impact of the industry. It will be verified through citation network backtracking. This can be noted that this may guarantee the presence of supporting evidence and the absence of opposing evidence between disciplines.

This study involves 26 papers, and 21 of them are core references and all passed the existence verification (title-author-year-source consistency) and citation-accuracy tests. The evidence remains largely on a summary level. It is also important to note that there are no artificial experimental facts. Industry reports and qualitative interviews are also used by us, which may complement the industry view, but they do not form the only foundation of core technical statements.

2. Research perspectives and structure of this review

This paper will divide this review into five major perspectives as this paper is able to look at the research that has been done on it in these various ways and through which readers can get an insight into what the AI has contributed to the game industry, how far it has gone in terms of its use, and the issues that have not yet been resolved. First, the content-production perspective is concerned with the way AI produces levels and assets of the game. It is important to mention that it is also possible to generate a whole world of games, but in the case of such games, the main problem would be the compromise between the quality, variety, control and other parameters of the product that should be taken into account by the developers in real life. After that, intelligent-character approach looks at the issue of open dialogue provided by LLMs to NPCs. As a matter of fact, it may even produce realistic behavior, so that the player feels like he or she is inside, and the price, memory, time responsiveness and other resources must be adjusted in such a way as to make sure that the system works well without any unnecessary load. Then there is the quality assurance point of view, where one wonders about the automation of testing the game using AI.

It can ease the burden of QA on the part of the high-frequency release, and this is a good thing since the frequency of releasing becomes shorter and shorter and manual testing will not be able to detect all issues in time. The other question that the player-experience perspective poses is the difficulty personalization of AI. This should also make the narrative flexible, such as how it would fit the players with various levels of skills and interests, which could enhance the overall experience of the game and even make people spend more time playing the game in theory. Lastly, there is the industry view on the role of AI in development organizations. There are also creator ecosystems and the final product of games themselves. These five dimensions have a developmental stage towards production, consumption and industry, and they may explain what the AI has done in the game business, how much the current way of making things has been altered by it, and also what are the controversies between the researchers and developers.

3. Thematic sections

3.1. Automation of content generation: from rule-based PCG to generative engines

The history of PCG is in the game industry, which is a field where the AI will be needed and it can be used as a change to the rule and search driven method into data driven so that this study can use the data directly. Liu [15] discussed a number of ways in which deep learning is applied to create levels and textures as well as other content. It is interesting to note that the process may be

independent. Or it may be joined with conventional practices. Following 2023, there are two branches: generation ability and evaluation standards. The authors of Todd [4] prove that big language models are able to create playable Sokoban level and their performance increases with the increase of the size of datasets. GameNGen by Valevski takes the generation to its maximum. There is a single TPU diffusion model capable of generating an entire DOOM environment in real time at 20 frames per second and human raters would not be able to tell the difference between the simulation and the actual game after five minutes of autoregressive generation, and this is what makes it an interactive prototype because it has been marked as a neural network as the game engine [3]. Simultaneously, Khalifa [14] suggests PCG Benchmark. It is a standard of comparison between the quality, diversity and controllability of 12 generation tasks in the range of levels to rule sets, as well as it offers a benchmark for the algorithms. According to Ratican and Hutson [1], this is Game Development 3.0 phase. The neural networks and large language models will write code, content and story on their own whereas human beings will become supervisors and make decisions.

3.2. LLM-driven NPCs and game agents

Traditional NPCs are based on the scripted lines and behavior trees, and this paper considers how to improve them to open LLM-based interaction that can bypass the constraints of conventional approaches. The generative agents developed by Park are the basis: agents record experiences into memory stream using natural language and they have a chance to reflect during retrieval of relevant memories so that believable behaviors can be produced. Actually, 25 agents in small-town simulation held parties and distributed invitations, and this demonstrates plausible social behaviors [5]. The survey by Hu categorizes LLM game agents into such parts as memory and reasoning, input/output modules, etc., and it records their evolution across six genres of games with important challenges being highlighted since real-time performance and context length are difficult to achieve, and out-of-distribution generalization is also anticipated to be hard [16]. There are two primary bottlenecks of the deployment which are cost and memory. Yu replaces learned policies with repeated LLM inference to compress multi-agent interaction costs [6], and Zheng

NPC memory stores are to be designed as human-like forgetting systems, which can be used in the continuation of conversation [7]. The experiments by Peng indicate that the free-form interaction between a player and NPCs based on LLMs may result in the emergence of narrative points not foreseen by the designer as 28 participants created the development of the narrative beyond the script and this is a first evidence of the phenomenon of the player-driven emergence [17].

3.3. AI for game testing and quality assurance

Games-as-a-service requires a regular update of the contents and this is why automated testing is required. The authors suggest that Wang come up with the TITAN model, which allows the LLM agents to test large MMORPG automatically through state abstraction and action optimization of plans. They are also able to repair their failed routes by self-correction. It has 95 percent completion rate in actual commercial games and is used in eight QA pipelines. This is the transition between research prototype and industrial application [8]. Previous attempts in this line of work involved using reinforcement-learning players that were trained at scale on multi-objective reward functions to detect bugs and strategies that are unbalanced however they had high cost of training and low generalization across games, and it was difficult to use them in practice. Testing agents can be endowed with spatial reasoning as well as long-horizon tasks with the help of LLMs. Notably, one of the weaknesses of visual-spatial reasoning is that of non-API graphical games.

3.4. Player modeling, dynamic difficulty, and personalized experience

Dynamic difficulty adjustment (DDA) is an attempt to balance the level of challenge with the level of ability of the players. The approach in this field was not based on heuristic rules but rather on data driven strategies. de Oliveira and Chaimowicz are training their opponents, who are matched by skill in a fighting game, using reinforcement learning, and they can manipulate win rates of their AI opponent to make it challenging without making it too difficult to the player [10]. Li offers a real-time mobile game that has a difficulty-aware structure. And they have survival analysis of how perceived difficulty will influence churn. They conduct A/B experiments of the live game, and the findings indicate that dynamic interventions that take into consideration the level of difficulty would be effective in increasing retention and engagement [9]. These two studies complement each other because one proves at the technical plane that the AI enemies are able to adjust, and the other proves at the business plane that it is possible to transform the adjustments in the difficulty into business measures. However, there is a reliance on small-scale tests which are short-term. It should be mentioned that there is no enough large-scale evidence of long-term motivation and effect of players.

3.5. Industry organizational change, controversies, and AI-native games

Industry level research indicates that there is a disparity between the efficiency narratives and reality of frontline practice. Ternar review ten qualitative studies published between 2020 and 2025, and they conclude that GenAI can transform game production pipelines and value networks, yet the outcomes are ambivalent [18]. Panchanadikar. Also, Freeman looks into 3,091 posts on indie-developer communities, and discovers that generative AI may encourage innovation practices of indie developers. Simultaneously, it has the potential to harm these practices as well [11]. Daneels interview 20 Belgian developers, and the findings indicate that most of them admit that AI has an assistive value in programming tasks, but they criticize that the narrative and visual quality are mediocre [12]. On organizational scale, Ahmed interview 20 studios AAA to indie, and discover that studios that were designed with AI since their inception can shorten development cycles by months to weeks, whereas those who only integrate AI into their existing workflows become stuck at bottlenecks between teams, and the gain size is dependent on how deep the organizational change was [13].

The survey of the three-year industry trends by Kultima also proves that generative AI will become the primary shift in the industry, as it is possible to replace crypto gaming [19]. The Chinese research on the industry focuses on the opportunities to reduce costs and increase efficiency of AIGC. And there are other dangers of copyright and data-security that can be found simultaneously [20], however, it remains at the macro level. Xu suggest a theory of the latest type of game: AI-native games; they argue that an artificial intelligence criterion is the difference between AI-native and AI-enhanced games, and 53 such titles are analyzed and it is concluded that the main design problem is how to arrange open semantics into fixed gameplay [2]. This implies that the extent of assistance provided by the AI may rely on the extent of change in the organization and creativity process, no agreement has been reached on the issue among the industries, and the AI-native games can be used as a test case of the paradigm shift.

4. Controversies and challenges

Controversy 1: Technological optimism vs. frontline skepticism. This study can show that the industry reports provide a blueprint of GenAI transformation. However, interviews with 20 Belgian developers indicate that most of them are doubtful about current tools [12], and indie-developer communities demonstrate both promotion and harm [11], which means there is a significant gap between what the technology can do in principle and what developers experience. It should be mentioned that the difference comes from a gap between technical potential and current tool quality, and benefits can be unequal across team sizes, and this makes the situation complex.

Controversy 2 is creativity. It should be mentioned that controlled assessments indicate that generative AI can enhance the understanding of difficulty with increased training data, but creative diversity does not necessarily increase with data scale [21] and this contradicts industry narrative that GenAI can assume creative work.

Controversy 3 is the efficiency versus organizational restructuring. This study can discover that even the most straightforward copy-and-paste adoption will bring limited benefits and studios that are created with AI in mind can shorten cycles to weeks [13]. Then this study will be able to observe degrees of help based on depth of organizational change, not the tool itself.

Controversy 4: AI-native games. The generation and the AI-native are not identical. There is no single answer to this question of which games may be considered as being native to the AI [2] but there is a counterfactual condition that must be applied.

There are still some issues to be solved. There is no generally accepted method of determining the playability of generated content automatically, which can be applied by the majority of researchers in this area [14]. Besides, the issue of acceptance and safety of the contents of the AI-native games was not considered empirically and the issue of copyright ownership was also not touched upon [2]. The studies of LLM-narrative and difficulty-adjustment are mostly small sample and have a short period and so it is impossible to generalize the findings as there is no external validity and this paper cannot use the results in other contexts. It is important to mention that empirical descriptions of the production process in Chinese language are difficult to locate and they are not many.

5. Future research directions

Direction 1: Automatic playability testing of AI-generated game content. It is anticipated to be standardized in PCG Benchmark, quality, diversity, also controllability. However, playability and fun may not have automatic assessment [14]. Actually, according to the open-source games and the public benchmarks, this paper is able to gather the player-behavior data such as completion rates and flow proxies, and this study trained playability predictors with this data which should be calibrated on human evaluation in principle. This study can also identify various types of games.

Direction 2: Long-term interaction and memory stability of LLM-driven NPCs. The difference is that in most cases, memory mechanisms [7] and cost compression [6] are prototype validations and they do not have data on large-scale long-term deployment which can be used to support real application scenarios with thousands of users. It is important to note that current works experiment on small scale only. This paper ought to construct a long-term memory standard based on open-source MMORPGs or dialogue models and this study creates multi-turn dialogue-consistency measures that can assess the memory consistency in long-term communication with complex user actions. Then this paper will be able to contrast various memory architectures in more than 100 hours of interaction as this study regulates context window and inference cost.

Direction 3: The research of the player acceptance and experience in AI-native games. Gap: there is a gap in the definition of AI-native games (only in literature [2]), the acceptance of players to AI-based games, as well as the trust on it are not considered much, and the fun of playing it is also absent. It should be mentioned that there is no empirical study of this field. Methodology: this study conducts experiments comparing players with the prototype of AI-native mechanics with conventional ones, which this paper will do through the flow and autonomy measurement. Besides, immersion in the narrative can be measured at one point and intensity of AI intervention may be regulated by the genre of the game.

Direction 4: The motivation of the players because of dynamic difficulty adjustment. Gap: DDA has not been studied over time, but only in short-term and small-scale studies [9, 10], which is why the long-term retention and affect have not been found. Strategy: conduct A/B testing in commercial games over the course of several months, use churn forecasting and subjective surveys to differentiate between the degree of difficulty adaptation and the excessive intervention, manage updates and seasonality.

Direction 5: This paper considers the longitudinal effect of generative AI on indie/small-team creativity. It is interesting to mention that the current studies are primarily based on cross-sectional interviews or community-post analysis [11] and it is not able to demonstrate how things change with time. The method follows a group of independent teams over 6-12 months. This study registers the use of AI tools and output quality development. In addition, creative autonomy must be measured. The number of people in the team and type of tool can be regulated.

6. Future directions

This study observe that in the last two years, there is a definite shift in research on AI in game industry. The direction of AI assistance is shifting towards a tool that does not only speeds up production but becomes gameplay itself and this transition suggests that the interaction between AI and game design is evolving in the modern industry. In early works, AI is considered as procedural content generation (PCG) and assistive creation. Ratican and Hutson define Game Development 3.0 as a new era of development in which large language models and neural networks can create code, content, and even narrative [1], after hand-coded production, or Software 1.0, and data-driven AI, or Software 2.0. Xu takes it one step further and suggests AI-native games. They employ counterfactual criterion to draw the distinction, with the test being whether the essence of the gameplay fails without runtime generative AI and they examine 53 publicly available titles to justify such classification [2]. All these works demonstrate present major tension.

Indeed, the problem is not how to generate but rather how this study can arrange AI - generated outputs into playable game mechanics and assess them. The ability to create things is no longer a barrier in the current research work.

On content level, deep learning has transformed the field of PCG. Liu explored deep-learning methods of game content generation in early [15]. They discovered that these techniques are able to create assets such as levels on their own. Also, they can collaborate with conventional approaches. Then the study proceeds along two directions. One is generation capability. The other is evaluation benchmark. Todd proved that large language models can generate playable Sokoban levels using data-driven method [4], and the performance can be improved when dataset size increases. GameNGen by Valevski demonstrates a more radical approach. Indeed, a diffusion model can simulate the entire DOOM game environment in real time at 20 frames per second and it is difficult for human raters to distinguish simulation from reality [3]. It should be noted that evaluation is behind. PCG Benchmark proposed by Khalifa has 12 generation tasks and tries to standardize

quality, diversity, controllability also being taken into consideration [14]. However, automatic playability and fun evaluation are still unsolved.

The topic of intelligent NPCs and game agents is also very popular nowadays. The key difference is that the scripted lines and behavior trees are replaced by open LLM-based interaction, i.e., the NPCs can produce responses dynamically rather than using a script. Work by Park is a valuable source [5]. They document experiences in natural language. The system is capable of periodically reflecting and retrieving memories. Indeed, 25 agents on simulated town gave parties and sent invitations without scripts, which demonstrated plausible social actions. Hu [16] provides an overview, which separates LLM game agents into memory, reasoning, and input/output components as well, and they report how adaptation varies between six types of games. However, there remains a discrepancy between research prototypes and playable NPCs. Yu [6] employs learned policies to avoid repeating LLM inference so that the cost of multi-agent interaction may be compressed whereas Zheng [7] designs NPC memory repositories with human-like forgetting mechanisms to assist in keeping the dialogue going.

It is important to mention that the two directions are real-time responsiveness and long-term memory as the two major constraints in deployment.

The quality-assurance and player-experience is also applied to production in real life by AI. Wang [8] has created TITAN system which can be implemented with LLM agents and perform automatic testing of large MMORPGs and it is capable of achieving 95% completion rate on tasks in commercial game using state abstraction and reflective correction. It has eight QA pipelines. Besides, personalization may be transferred into data rather than heuristics with dynamic difficulty control with modeling of players. Indeed, de Oliveira and Chaimowicz train an opponent that could play against a player who has mastered the fighting game skill [10]. Li [9] proposes a framework of difficulty-awareness of mobile games and A/B experiments indicate that interventions are able to enhance retention as well as engagement. However, there are still the assessments on both sides, and these depend on short term experimentation on small samples, and one cannot forget that long term motivation and affect evidence is not available in great volumes.

Research at the industry level shows the gap between efficiency narratives and frontline practice. The systematic synthesis of ten qualitative studies (2020–25) by Ternar shows GenAI reshaping pipelines and value networks with mixed effects [18]; Panchanadikar and Freeman's analysis of 3,091 indie-community posts finds generative AI both promoting and harming indie innovation [11]; Daneels' interviews with 20 Belgian developers are blunter—most respondents doubt GenAI's practical usefulness, acknowledging its assistive value in technical tasks while criticizing its mediocre narrative and visual quality [12]. Notably, there is organizational divergence: Ahmed's interviews with 20 AAA-to-indie studios show that studios "designed around AI from day one" compress cycles from months to weeks, while those embedding AI into existing workflows get stuck at cross-team collaboration—suggesting the degree of AI's help depends at least partly on the depth of organizational change [13]. Chinese industry research also highlights the cost efficiency opportunities of AIGC and coexisting copyright and data security risks [20] yet mainly remains at a macro level without empirical accounts of production stages.

The overall opinion is that the help of AI has not only been to do it quicker, but the gameplay itself will be different, with AI-native games being the testbed [2], and Peng experiments have also demonstrated that in case of free-form player-NPC interaction, when LLM-based characters are involved, the designer does not know what nodes the narrative will create, so there are some initial evidence on the emergence of players. The open issues are as obvious: the generated content does not have a universal standard of automatic playability [14]; the empirical studies of the acceptance

by the players, safety of the content and copyright rights of the AI-native games are practically non-existent [2]; and most of the available research were conducted on small samples and over a short period, hence the findings cannot be extrapolated to a long-term process. The debate of literature is not whether or not AI is useful, but rather the conditions under which it is useful, the organizational structure, and the stages of creativity where the assistance is stable.

7. Conclusion

This review examined recent developments in artificial intelligence in the game industry, with a focus on content generation, LLM-driven NPCs, automated game testing, player modeling, and AI-native games. The reviewed studies show that AI is gradually moving from a supporting tool to a more important part of game development and gameplay. Generative models can create game content, while LLM-based agents can improve NPC interaction and support automated testing. Reinforcement learning and player modeling also make it possible to adjust game difficulty according to player behavior. However, several challenges remain. Current systems still face problems related to computational cost, long-term memory, content quality, copyright, safety, and player acceptance. Many existing studies are also based on small samples or short-term experiments. Future research should therefore pay more attention to long-term player experience, automatic playability evaluation, reliable AI agents, and the practical use of AI-native game design. Overall, AI is likely to become an increasingly important part of future game development.

References

- [1] Ratican, J., & Hutson, J. (2024). Video game development 3.0: AI-driven collaborative co-creation. *Metaverse*, 5(2), 2904. DOI: 10.54517/m2904.
- [2] Xu, Z., Meng, F., Xu, K., Verbrugge, C., Lucas, S., & Zhao, J. (2026). AI native games: A survey and roadmap. arXiv preprint, arXiv:2607.00527.
- [3] Valevski, D., Leviathan, Y., Arar, M., & Fruchter, S. (2025). Diffusion models are real-time game engines. *Proceedings of the 13th International Conference on Learning Representations (ICLR 2025)*.
- [4] Todd, G., Earle, S., Nasir, M. U., Green, M. C., & Togelius, J. (2023). Level generation through large language models. *Proceedings of the 18th International Conference on the Foundations of Digital Games*, Article 70, 1-8. DOI: 10.1145/3582437.3587211.
- [5] Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, Article 2, 1-22. DOI: 10.1145/3586183.3606763.
- [6] Yu, Y., Zhang, Q., Li, J., Fu, Q., & Ye, D. (2024). Affordable generative agents. *Transactions on Machine Learning Research*.
- [7] Zheng, S., He, K., Yang, L., & Xiong, J. (2024). MemoryRepository for AI NPC. *IEEE Access*, 12, 62581–62596. DOI: 10.1109/ACCESS.2024.3393485.
- [8] Wang, C., Tang, L., Yuan, M., Yu, J., Xie, X., & Bu, J. (2025). Leveraging LLM agents for automated video game testing. arXiv preprint, arXiv:2509.22170.
- [9] Li, J., Lu, H., Wang, C., Ma, W., Zhang, M., Zhao, X., Qi, W., Liu, Y., & Ma, S. (2021). A difficulty-aware framework for churn prediction and intervention in games. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 943–952. DOI: 10.1145/3447548.3467277.
- [10] de Oliveira, T. N., & Chaimowicz, L. (2024). Investigating reinforcement learning for dynamic difficulty adjustment. *Proceedings of the 22nd Brazilian Symposium on Games and Digital Entertainment (SBGames '23)*, 66–75. DOI: 10.1145/3631085.3631229.
- [11] Panchanadikar, R., & Freeman, G. (2024). "I'm a solo developer but AI is my new ill-informed co-worker": Envisioning and designing generative AI to support indie game development. *Proceedings of the ACM on Human-Computer Interaction*, 8(CHI PLAY), Article 317, 1–26. DOI: 10.1145/3677082.
- [12] Daneels, R. (2025). Exploring the opportunities and risks of generative AI for game development: Insights from the Belgian game industry. *Proceedings of the First Workshop on Artificial Intelligence for Human-Game Interaction*

(AI4HGI 2025), *CEUR Workshop Proceedings*, 4097.

- [13] Ahmed, Z. (2026). Beyond copy-and-paste: How game studios are reorganizing around AI. Wharton Generative AI Labs, University of Pennsylvania.
- [14] Khalifa, A., Gallotta, R., Barthet, M., Liapis, A., Togelius, J., & Yannakakis, G. N. (2025). The procedural content generation benchmark: An open-source testbed for generative challenges in games. *Proceedings of the International Conference on the Foundations of Digital Games (FDG '25)*, 1-12. DOI: 10.1145/3723498.3723794.
- [15] Liu, J., Snodgrass, S., Khalifa, A., Risi, S., Yannakakis, G. N., & Togelius, J. (2021). Deep learning for procedural content generation. *Neural Computing and Applications*, 33(1), 19-37. DOI: 10.1007/s00521-020-05383-8.
- [16] Hu, S., Huang, T., Liu, G., Kompella, R. R., Ilhan, F., Tekin, S. F., Xu, Y., Yahn, Z., & Liu, L. (2024). A survey on large language model-based game agents. arXiv preprint, arXiv:2404.02039. DOI: 10.48550/arXiv.2404.02039.
- [17] Peng, X., Quaye, J., Rao, S., Xu, W., Botchway, P., Brockett, C., Jojic, N., DesGarennes, G., Lobb, K., Xu, M., Leandro, J. J. G., Jin, C., & Dolan, B. (2024). Player-driven emergence in LLM-driven game narrative. *2024 IEEE Conference on Games (CoG)*, 1-8. DOI: 10.1109/CoG60054.2024.10645607.
- [18] Ternar, D.-A., Denisova, A., Cunha, J. M., Kultima, A., & Guckelsberger, C. (2026). Generative AI in game development: A qualitative research synthesis. *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*, Article 452, 1-24. DOI: 10.1145/3772318.3791206.
- [19] Kultima, A. (2025). Trends in gaming: Views of industry professionals and game scholars. *Proceedings of the 20th International Conference on the Foundations of Digital Games (FDG '25)*, Article 10, 1-7. DOI: 10.1145/3723498.3723743.
- [20] Xu, X., Tian, K., & Li, W. (2023). Next generation artificial intelligence technology (AIGC): Development evolution, industrial opportunities and prospects. *Review of Industrial Economics*, 57(4), 5-22. DOI: 10.19313/j.cnki.cn10-1223/f.20230705.001.
- [21] Chang, C. (2025). Testing creative and comprehensive capabilities of generative AI through game design evaluation. *The National High School Journal of Science*, 19 January 2025.