Generative Artificial Intelligence in Digital Game Design and Development: A Bibliometric and Complex Network Analysis

Murilo Mazzotti Silvestrini[1], **Pedro Henrique Salmaze**[2], **Leonardo Tórtoro Pereira**[3]

[1] State University of Campinas (UNICAMP)
Institute of Mathematics, Statistics and Scientific Computing
R. Sérgio Buarque de Holanda, 651 – 13083-859 – Campinas – São Paulo – Brazil

m261854@dac.unicamp.br

[2]University of São Paulo (USP)
Institute of Mathematical and Computer Sciences
Av. Trabalhador Sancarlense, 400 – 13566-590 – São Carlos – SP – Brazil

pedrohsalmaze@usp.br

[3]São Paulo State University (UNESP)
Institute of Geosciences and Exact Sciences
Rua 24 A – 13506-900 – Rio Claro – São Paulo – Brazil

leonardo.t.pereira@unesp.br

Abstract. Introduction: *Generative AI (GenAI) has reshaped digital game design since 2017, yet the structural roles of its core concepts remain unmapped.* ***Objective:*** *This study maps GenAI in digital games (2018–2025) using bibliometric and keyword co-occurrence network analysis.* ***Methodology:*** *From 111 peer-reviewed articles (PRISMA-adapted pipeline, LLM-based semantic screening + human validation, Cohen's $\kappa = 0.60$), we applied four centrality measures and Louvain community detection.* ***Results:*** *The network (29 nodes, 103 edges) shows atypical topology — density 0.254, clustering 0.675, modularity 0.136, 62.1% inter-community weight — with all centralities converging on five hubs: game, generative model, PCG, training, and LLM. Two anchors emerge: game (domain anchor, highest degree/eigenvector) and generative model (technology anchor, betweenness 2.8× higher). Both findings are robust across nine threshold configurations and 100 Louvain runs. Lexical analysis shows a paradigm shift: LLMs (+3.47 pp) and diffusion models (+1.28 pp) ascend, while GANs and generic deep-learning vocabulary dilute. The field is highly integrated, anchored by an umbrella technological concept, in early modular specialization. Search strings, screening prompt, per-record decisions, and reproduction scripts are publicly available.*
Keywords *Generative Artificial Intelligence, Digital Games, Bibliometric Analysis, Complex Networks, Procedural Content Generation.*

1. Introduction

Generative Artificial Intelligence (GenAI) has emerged as a disruptive force in computer science, producing original content (text, images, audio, 3D models, code) from learned patterns [Wu et al. 2025, Xia 2025]. Since content creation accounts for a substantial share of production costs in digital games [Ternar et al. 2026], GenAI is transformative for design and development workflows.

The field spans a heterogeneous spectrum: Large Language Models (LLMs) enable dynamic narratives and NPC behaviors [Yang et al. 2025, Sweetser 2024]; GANs and diffusion models accelerate visual asset production [NVIDIA 2024, Chen et al. 2023]; and deep-learning-based Procedural Content Generation (PCG) automates levels and balancing [Sudhakaran et al. 2023, Nie et al. 2025]. The 2017 Transformer [Vaswani et al. 2017] was the anchor event; from 2020 onward, after GPT-3, DALL-E, and diffusion models [Ho et al. 2020], output accelerated rapidly [Lao 2025, Korsah 2025].

Despite this growth, structural mapping remains limited. Recent narrative and scoping reviews [Korsah 2025, Ternar et al. 2026] offer thematic panoramas but do not reveal clusters, bridges, or dynamics; quantitative bibliometric studies [Xia 2025, Čišić et al. 2025] map clusters with off-the-shelf science-mapping tools but do not differentiate the structural roles of the concepts inside them. No study has yet combined community detection with several concurrent centrality measures to recover those roles.

This paper addresses this gap by investigating (2018–2025): the quantitative production profile, thematic communities and their interconnections, emerging/declining concepts across sub-periods (2017–2019, 2020–2022, 2023–2025), structural hubs, and research gaps. Specifically: **RQ1**: What is the quantitative production profile (temporal, source, venue) of GenAI-in-games research? **RQ2**: Which thematic communities structure the field, and how do they interconnect? **RQ3**: Which concepts emerge or decline across sub-periods? **RQ4**: Which terms act as structural hubs, and what roles do they play? **RQ5**: What research gaps are revealed by the network topology?

The main contributions are threefold: (i) a reproducible pipeline combining PRISMA retrieval, LLM-based semantic screening with human validation, thesaurus-harmonized keyword normalization, and co-occurrence network analysis with Louvain community detection and multi-measure centrality; (ii) a structural characterization of the GenAI-in-games literature across three technologically meaningful sub-periods; and (iii) the identification of two distinct structural anchors — a domain anchor and a technology anchor — whose coexistence explains the integrated topology of the field.

2. Related Work

GenAI game research has developed along three technological strands—language, image, content generation—reviewed here in turn, along with reviews and bibliometric studies mapping them.[1]

The first strand applies Transformer-based LLMs to narrative and NPC behavior [Sweetser 2024, Yang et al. 2024]. A scoping review of the GPT family in games and its subsequent update [Yang et al. 2024, Yang et al. 2025] converge on five dominant areas—PCG, mixed-initiative design, mixed-initiative gameplay, AI playing games, and user research—delimiting the design space empirical work has

[1]Works cited here as arXiv preprints/theses are contextual framing references only; they are not part of the analyzed corpus, which comprises exclusively peer-reviewed, database-indexed publications (Section 3).

since populated. Within that space, GPT-family and Llama-3 systems have produced adaptive NPCs, branching dialogue, and full RPG narratives [Buongiorno et al. 2024, Kumaran et al. 2023, Callison-Burch et al. 2022]. These deployments expose limits: coherence decays over extended interaction [Trukhin et al. 2025], characters drift toward caricature via "Flanderization" [Buakhaw et al. 2025], and cloud inference imposes latency cost local deployment only partially mitigates [Li 2025].

The second strand addresses visual asset production. Following GANs' application to textures, terrains, and 2D assets [Gabani et al. 2023, Singh et al. 2024a], diffusion models have displaced them for most tasks: Stable Diffusion and ControlNet drive level/asset generation [Meira-Rodríguez 2025, Wu 2025], while DALL-E/Midjourney evaluations document control-versus-fidelity trade-offs designers must negotiate [Singh et al. 2024b]. The same displacement is under way in 3D, where Edify-3D [NVIDIA 2024] and text-driven texture synthesis [Chen et al. 2023] generate PBR-textured geometry; the open frontier is integration into commercial engines [Liu et al. 2025].

The third strand, PCG, has undergone a comparable transition from rule-based to learned generators, documented by a survey covering terrain, levels, sprites, and storylines [Liu et al. 2021]. MarioGPT [Sudhakaran et al. 2023] fine-tunes GPT-2 for controllable level generation, and Generative Playing Networks [Bontrager e Togelius 2021] co-train agents and generators, establishing the two templates—sequence modeling and co-evolution—that subsequent work extends via interactive GAN latent exploration [Schrum et al. 2020], PPO-trained dual agents [Özkan 2025], and distillation into steerable models [Nie et al. 2025].

Attempts to map these strands remain predominantly qualitative. Systematic and narrative reviews [Wu et al. 2025, Korsah 2025, Ternar et al. 2026] identify recurring concerns—democratization of production, VR/AR, originality—while adjacent work maps interaction-design space for generative creativity tools [Hu et al. 2025]. Quantitative mapping is more incipient: Xia [Xia 2025] proposes a bibliometric workflow, Čišić et al. [Čišić et al. 2025] map 149 AI-HCI papers, and Lao [Lao 2025] documents post-2020 acceleration in output.

Both groups stop short of the same thing. Qualitative reviews treat concepts in isolation rather than modeling relations; quantitative ones map co-occurrence clusters but report membership without differentiating roles within the hub core—a distinction requiring several centrality measures computed concurrently over the same network and read against one another. Co-word analysis with community detection over keyword co-occurrence networks addresses precisely this, and is well established in bibliometrics—applied to nanotechnology and environmental risk [Radhakrishnan et al. 2017], technology and innovation management [Ni et al. 2019], information retrieval [Ding et al. 2001], and consumer behavior [Muñoz-Leiva et al. 2012]—but has not yet been brought to bear on GenAI in digital games. The present study does so.

3. Methodology

The protocol is quantitative, descriptive-exploratory, and documentary, combining bibliometric [Aria e Cuccurullo 2017] and complex network analyses

[Blondel et al. 2008]. It adapts Zhang *et al.* [Zhang et al. 2012], Fonseca *et al.* [Fonseca et al. 2025], and JBI scoping-review guidelines [Tricco et al. 2018], implemented in Python.

3.1. Temporal Scope

The search window is 2017–2025; no 2017 article survived, so the effective corpus spans 2018–2025. The interval is split into three technologically meaningful sub-periods (Table 1). The first contains only 4 articles, making comparisons involving it indicative trends.

Table 1. Sub-periods, technological milestones, and impact on digital games. Sub-periods are defined by the search window; the effective corpus begins in 2018, so the first sub-period covers 2018–2019 in practice.

Sub-period	Technological milestone	Impact on games
2017–2019	Transformer (2017); GPT-1/2 (2018–2019)	First GAN-based PCG; deep-network level generation.
2020–2022	GPT-3 (2020); DALL-E (2021); DDPM (2020); Stable Diffusion (2022)	Explosion of creative applications; art, texture, narrative; first integrated pipelines.
2023–2025	ChatGPT; GPT-4; Midjourney v5+; Sora; Genie	LLM-assisted design; 3D world generation; conversational NPCs; engine-integrated tools.

3.2. Search Strategy

Searches apply three-concept AND logic (OR within concepts), $\mathbf{C_1} \wedge \mathbf{C_2} \wedge \mathbf{C_3}$, targeting GenAI techniques (C_1), digital games (C_2), and specific applications (C_3) (Table 2), executed on Scopus, WoS, IEEE Xplore, and ACM DL, filtered by document type (article, review, conference paper), English, and 2017–2025. Query strings and execution dates are in the public repository.[2]

Table 2. Search concepts and term lists (OR-joined within each concept).

Concept	Terms
C_1: GenAI	*"generative AI"; "generative artificial intelligence"; "large language model*"; LLM; "diffusion model*"; "generative adversarial network*"; GAN; "text-to-image"; "transformer model*"; ChatGPT; GPT; "stable diffusion"; DALL-E; Midjourney.*
C_2: Games	*"video game*"; "digital game*"; "computer game*"; "game design"; "game development"; "game AI"; gaming; "serious game*"; "virtual world*".*
C_3: Application	*"procedural content generation"; PCG; "level generation"; "narrative generation"; "NPC behavio*"; "asset generation"; "texture generation"; "content creation"; "player experience"; "game testing"; "game balancing".*

[2] Search strings and run dates, the complete LLM screening prompt with its few-shot examples, per-record classification decisions, the synonym dictionary, the threshold-sensitivity grid, and reproduction scripts are available at `https://github.com/muriloms/sbgames-bibliometric-complex-network-analysis-genAI-game-design`.

3.3. Selection Protocol

Figure 1 outlines the five-stage pipeline. Records are merged, deduplicated, harmonized with a manual thesaurus, and filtered by document type (article, non-empty abstract/keywords). We retain English original articles (2017–2025) applying or discussing GenAI for content generation, NPC behavior, game testing, or player experience; purely algorithmic PCG, non-generative components, non-game studies, and editorials are excluded. As retrieval was limited to four indexing databases, the corpus includes only peer-reviewed, database-indexed publications (no preprints).

Each candidate is screened by an LLM semantic classifier (*gpt-4o-mini*, temperature 0.1, fixed seed, three few-shot examples) evaluating three criteria: technological adequacy (GenAI), domain adequacy (digital games), and application adequacy (empirical/prototype). Records satisfying at least two criteria are kept (2-of-3 rule); the full prompt, examples, and per-record decisions are in the repository (footnote 2). A stratified random sample of 50 decisions (25 accepted, 25 rejected) was independently re-examined by a human annotator, yielding 80.0% observed agreement and Cohen's $\kappa = 0.60$ (substantial). The 10 disagreements comprised 7 false positives (domain-adjacent work lacking a generative component) and 3 false negatives (missed reuse of generative assets); all were resolved by human adjudication, which prevails in the final corpus.

1. Identification — queries in Scopus, WoS, IEEE Xplore, ACM DL

↓

2. Merging, deduplication, thesaurus harmonization $\rightarrow$ 451 records

↓

3. Documentary filters — article type, abstract, keywords $\rightarrow$ 227 records

↓

4. LLM semantic screening + human validation (3 criteria, $\geq 2/3$ rule) $\rightarrow$ 111 records

↓

5. Final analytical corpus $\rightarrow$ 111 records

Figure 1. PRISMA-adapted five-stage selection and preparation pipeline.

3.4. Analytical Procedures

Four analytical layers are used. The descriptive bibliometric layer computes production indicators (annual time series, source-database distribution, top venues, author counts, raw keyword frequencies). The keyword normalization layer applies a curated synonym dictionary and rule-based lemmatization (English plurals only), avoiding morphological over-collapsing; for abstracts, NLTK's WordNetLemmatizer with POS tags resolves verb–noun ambiguity before canonical mapping. The lexical-dynamics layer cross-tabulates top normalized keywords across the three sub-periods, flagging emerging and declining terms by percentage-point changes in per-period relative share.

The complex-network layer constructs the keyword co-occurrence network: nodes are normalized keywords (minimum frequency ≥ 3), edges denote co-occurrence in the same article (minimum edge weight ≥ 2), and isolated vertices after edge filtering are removed. Structural indicators ($|V|$, $|E|$, density, mean clustering coefficient, component count, modularity Q) characterize the global topology; four centrality measures (weighted degree, betweenness, closeness, eigenvector) are computed concurrently using shortest-path distance, with convergence across measures taken as evidence of structural robustness of the identified hubs. Communities are detected via the Louvain algorithm [Blondel et al. 2008] (fixed seed 42), and the intra- versus inter-community edge-weight ratio quantifies integration or fragmentation of the field.

4. Results

4.1. Corpus Characterization

After merging and deduplication, documentary filters retained 227 of 451 records for semantic screening. The LLM classifier yielded 111 relevant articles (48.9% of the screened set): 60 satisfied all three criteria (26.4%) and 51 satisfied exactly two (22.5%). Exclusions: *Not game-related* (87/116; 75.0%), *Theoretical only* (15), *Non-generative AI* (8), and *Wrong domain* (6) — consistent with the broad semantic reach of the native queries.

The corpus is dominated by ACM DL (79; 71.2%), followed by IEEE Xplore (28; 25.2%) and WoS (4; 3.6%). Scopus exclusive records were out-of-scope types or duplicates already indexed elsewhere; no unique eligible article was lost. Approval rates (ACM 42.7%, IEEE 73.7%, WoS 100%) indicate differential semantic alignment. Top venues: *ACM Trans. Graphics* (32), *IEEE Trans. Games* (20), and *PACM HCI* (18), covering graphics, game-specialized research, and HCI.

Temporal growth is rapid: annual counts are 1, 3, 5, 6, 9, 8, 39, and 40 (2018–2025). Aggregated by sub-period, 2017–2019 (Transformer/GAN) contributes 4 articles (3.6%), 2020–2022 (Diffusion/GPT-3) 20 (18.0%), and 2023–2025 (ChatGPT/Midjourney) 87 (78.4%). The 4.9$\times$ jump between 2023 and 2024 coincides with end-user generative system consolidation in late 2022, marking an inflection point.

4.2. Lexical Dynamics and Thematic Evolution

Raw keywords showed substantial redundancy: 83.0% of 506 unique forms occurred in a single article. The synonym dictionary (105 mappings to 44 forms) unified variants under *generative model*, *LLM*, *PCG*, and *GAN*. The terminological core is led by *generative model* (36), *game* (26), *LLM* (19), *PCG* (18), and *training* (13). The temporal heatmap (Figure 2) shows progression from homogeneous vocabulary in 2017–2019 to concentration on *generative model*, *game*, and *LLM* in 2023–2025, with *AI*, *diffusion model*, and *3d* emerging.

Share-change analysis: LLM (+3.47 pp; $0 \to 19$) and diffusion (+1.28 pp; $0 \to 7$) emerge; ML (-2.78), NN (-2.78), and GAN (-2.42) decline in relative share but remain in absolute terms, indicating dilution, not abandonment. As the first sub-period has 4 articles, the changes are indicative, but the direction rests on the 87 articles of 2023–2025.

4.3. Network Structure and Hub Concepts

With min. frequency $\geq$ 3 and edge-weight $\geq$ 2 filters, isolated vertices removed, the co-occurrence network comprises 29 vertices and 103 edges in a single component. Thresholds (freq $\geq$ 3 = $\geq$ 2.7% of articles; edge $\geq$ 2 = support from $\geq$

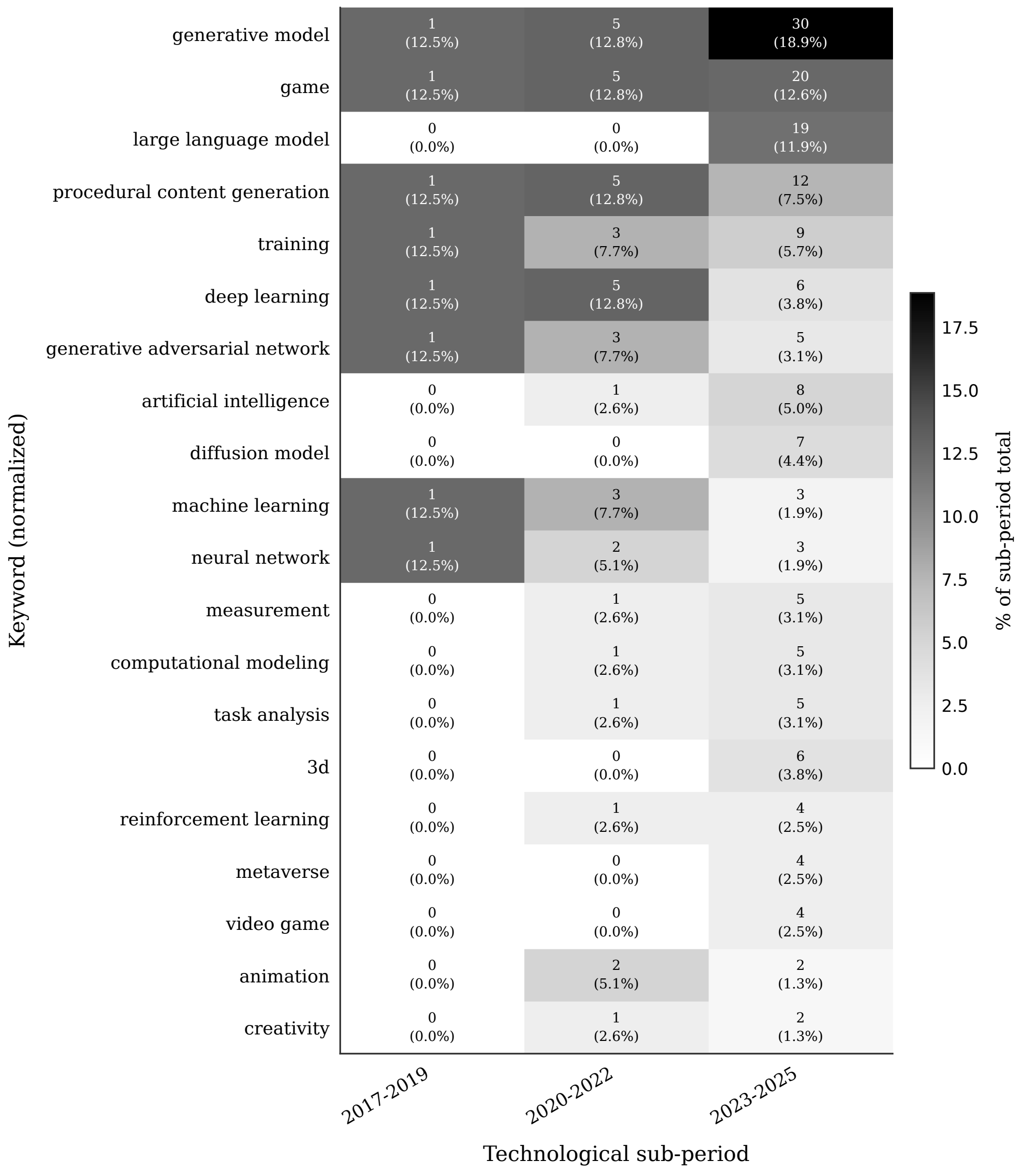

Figure 2. Temporal evolution of the top 20 normalized keywords across the three sub-periods. Values are column-normalized (% of sub-period total).

2 articles) filter idiosyncratic/spurious terms, standard practice [Cobo et al. 2011, van Eck e Waltman 2014].

Structural indicators (Table 3) show atypical topology: density 0.254, clustering 0.675, modularity $Q = 0.136$. Despite three Louvain communities, Q is well below the 0.3 threshold for well-defined structure [Clauset et al. 2004], indicating loose boundaries.

Sensitivity analysis over 9 configurations (freq 2–4, edge 1–3) confirms 4 of 5 hubs in top 5 across runs; Q remains <0.3 (0.063–0.246). The hub core and weak modularity are not artifacts (full grid in repository). Community detection is stable: 100 runs yielded 3 communities ($Q = 0.134 \pm 0.004$, ARI=0.91, NMI=0.93).

The four centrality measures (Table 4) converge on five prominent terms — *game*, *generative model*, *PCG*, *training*, and *LLM* — all in the top 5 of every measure,

Table 3. Structural indicators.

Indicator	Value
$\|V\|$	29
$\|E\|$	103
Density	0.254
Mean clustering coefficient	0.675
Connected components	1
Giant-component coverage (%)	100
Louvain communities	3
Modularity Q	0.136
Inter-community edge weight (%)	62.1

confirming structural robustness.

Table 4. Top 8 keywords by weighted degree and corresponding centrality values.

Keyword	Freq	Deg. (w)	Betw.	Clos.	Eig.
game	26	118	0.070	0.824	0.513
procedural content generation	18	87	0.066	0.737	0.435
generative model	36	83	0.351	0.849	0.372
training	13	59	0.124	0.683	0.316
large language model	19	35	0.100	0.622	0.202
artificial intelligence	9	32	0.034	0.596	0.205
generative adversarial network	9	29	0.043	0.583	0.182
task analysis	6	28	0.004	0.583	0.180

Joint ranking inspection reveals two distinct roles: *game* (degree 118, eigen 0.513) is the domain anchor (dense hub); *generative model* (betweenness 0.351, 2.8× higher than second-ranked *training*, closeness 0.849) is the technology anchor (bridge linking methods to applications). Their coexistence is the structural signature: umbrella technology binding heterogeneity to the game domain.

Figure 3 visualizes communities. C0 (12 vertices, freq 108) centers on *game* and *PCG*, covering game design and AI agents; C1 (9, freq 81) on *generative model*, *GAN*, and *diffusion*, covering generative methods and multimedia; C2 (8, freq 44) groups *training*, *3d*, and *neural network*, covering training and 3D content. Given low modularity, these are weakly differentiated tendencies rather than separate epistemic subcommunities.

Inter-community edge analysis (Figure 4) shows 62.1% inter-community and 37.9% intra-community. Densest exchange: C0-C1 (14.6%), then C0-C2 (11.1%), C1-C2 (5.4%). With $Q = 0.136$, the field is highly integrated, not siloed, consistent with its youth, where articulation between techniques, applications, and training has not yet crystallized.

5. Discussion

Our results answer the research questions: temporal profile (RQ1) shows rapid post-2022 growth; community structure (RQ2) is highly integrated with three loosely

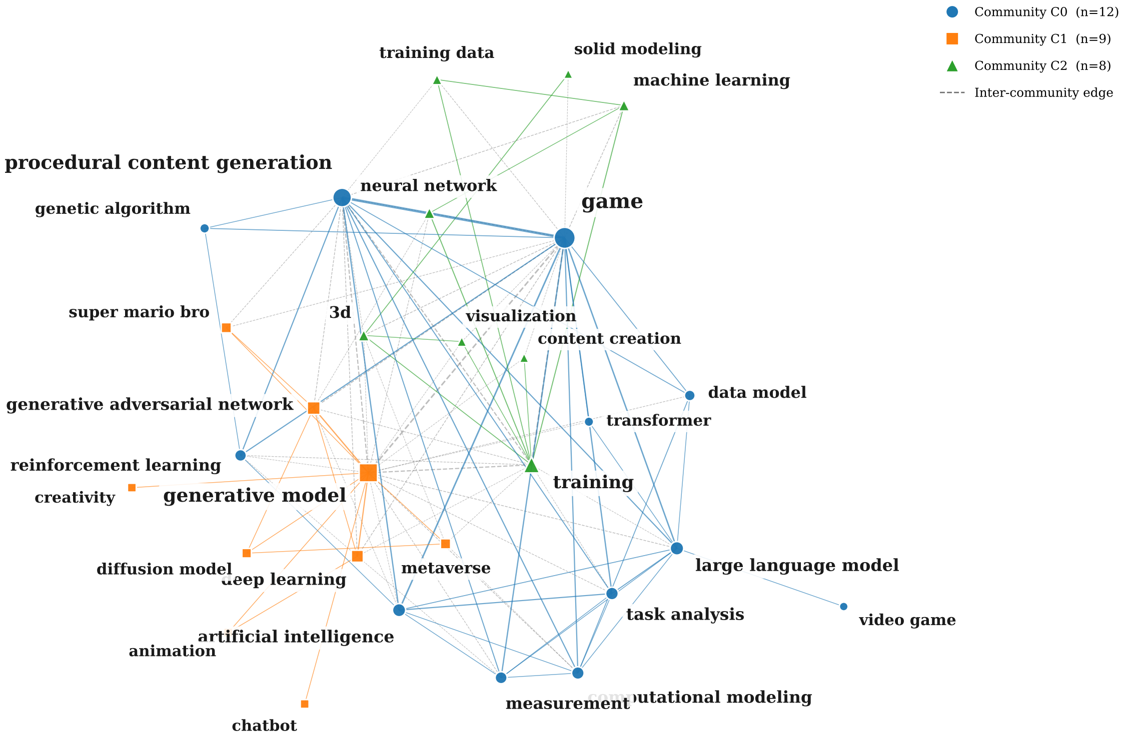

Figure 3. Keyword co-occurrence network (29 nodes, 103 edges). Node color: Louvain community; node size: frequency; edge thickness: co-occurrence weight; dashed edges: cross-community.

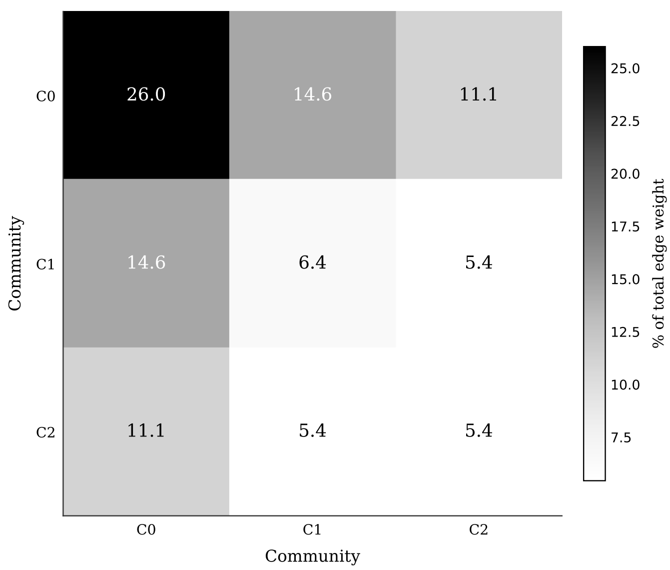

Figure 4. Inter- and intra-community edge-weight matrix (3×3), as percentages of total edge weight. Diagonal: intra-community; off-diagonal pairs contribute twice to the 62.1% inter-community total.

differentiated tendencies; lexical shift (RQ3) confirms the rise of LLMs and diffusion models; hub analysis (RQ4) reveals two anchors—*game* (domain) and *generative model* (technology). Structurally, a dense, transitive network ($d = 0.254$, $C = 0.675$) with low modularity ($Q = 0.136$) and 62.1% inter-community edge weight characterizes a highly integrated, non-fragmented field. This is explained by the two-anchor architecture: *game* (domain anchor, highest degree/eigenvector) and *generative model* (technology anchor, betweenness 2.8$\times$ higher), mediating heterogeneity across LLM, GAN, diffusion, and PCG literatures. This refines existing mappings [Xia 2025, Čišić et al. 2025, Lao 2025], showing post-2020 acceleration coexists with low-modularity structure whose functional meaning (two anchors) is not recoverable from cluster membership alone.

Practically, since the bridge is *generative model* rather than a specific technique, conceptual traffic passes through an umbrella category, not mature technique-to-application mappings. Entry points are thus organized by problem domain and generic technology, not consolidated technique–application pairs; no dense cluster binds diffusion models to asset pipelines or LLMs to narrative systems. Such pairs crystallize in maturing fields; their absence structurally correlates with low modularity—the field is integrated because it has not yet specialized.

Lexical analysis supports a paradigm shift: LLMs (+3.47 pp) and diffusion models (+1.28 pp) emerge, while GANs (-2.42 pp) and generic deep-learning vocabulary dilute. These magnitudes are indicative (given the 4 articles in the first sub-period), but the direction is supported by the 87 articles of 2023–2025. Declining terms remain in absolute counts, indicating broadening, not replacement.

The structural map exposes under-explored regions (RQ5). The training/3D community is smallest and least central, indicating training pipelines, data curation, and 3D asset generation remain peripheral to the conceptual core despite their weight in production practice—a gap between literature and game development consumption.

6. Conclusion

This study applied an integrated bibliometric and complex-network protocol to map Generative AI in digital games (2018–2025). From 111 peer-reviewed articles (PRISMA-adapted pipeline, LLM screening, human validation), we built a keyword co-occurrence network with multi-measure centrality and Louvain detection. The field is structurally integrated (high density/clustering, low modularity, high inter-community edge weight) around two anchors: *game* (domain) and *generative model* (technology). Lexically, it shifts from GANs/deep-learning toward LLM/diffusion terminology.

Contributions are threefold: a reproducible pipeline (PRISMA, LLM, thesaurus, network analysis); structural characterization across three sub-periods; and identification of the two anchors. Practically, the map orients newcomers to conceptual bridges between generative methods and game production.

Limitations involve coverage and granularity. Reliance on four general databases misses specialized venues (SBGames, AIIDE, IJCAI), potentially excluding relevant work — a known database-mapping limitation that does not affect structural relations computed over retrieved articles. Scopus contributed no unique articles, the first sub-period is small ($n = 4$), and analysis depends on author-supplied keywords. Future work includes venue-level searches of SBGames, AIIDE, IJCAI; corpus expansion; collaboration/citation network incorporation; and tracking hub concepts to game engines.

References

Aria, M. e Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4):959–975.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., e Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.

Bontrager, P. e Togelius, J. (2021). Learning to generate levels from nothing. In *2021 IEEE Conference on Games (CoG)*, pages 1–8.

Buakhaw, P., Kerdthaisong, K., Phenhiran, P., Khlaisamniang, P., Vorathammathorn, S., Ittichaiwong, P., e Yongsatianchot, N. (2025). Deflanderization for game dialogue: Balancing character authenticity with task execution in LLM-based NPCs. arXiv preprint arXiv:2510.13586.

Buongiorno, S., Klinkert, L. J., Chawla, T., Zhuang, Z., e Clark, C. (2024). PANGeA: Procedural artificial narrative using generative AI for turn-based video games. arXiv preprint arXiv:2404.19721.

Callison-Burch, C., Tomar, G. S., Ippolito, D., Martin, L. J., Bailis, S., e Reitter, D. (2022). Dungeons and Dragons as a dialog challenge for artificial intelligence. arXiv preprint arXiv:2210.07109.

Chen, D. Z. et al. (2023). Text2Tex: Text-driven texture synthesis via diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*.

Čišić, D., Drezgić, S., e Čegar, S. (2025). A bibliometric analysis of artificial intelligence in human–computer interaction. In *2025 48th MIPRO ICT and Electronics Convention (MIPRO)*. IEEE.

Clauset, A., Newman, M. E. J., e Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, 70(6):066111.

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., e Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, 62(7):1382–1402.

Ding, Y., Chowdhury, G. G., e Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information Processing and Management*, 37(6):817–842.

Fonseca, C. V. C., Predo, R. M., e Blikstad, N. M. D. (2025). Defining a research agenda for Industry 4.0 technologies in corporate strategies: insights from complex networks and machine learning. *International Journal of Production Research*, 63(23):8743–8760.

Gabani, K., Pathak, H., Thakkar, I., Jain, N., Rajendran, S., e Mukherjee, D. (2023). Dynamic terrain generation with deep GANs. pages 61–69.

Ho, J., Jain, A., e Abbeel, P. (2020). Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 6840–6851.

Hu, X., Xing, Y., Cai, X., Zhao, Y., Cook, M., Borgo, R., e Neate, T. (2025). Designing interactions with generative AI for art and creativity: A systematic review and taxonomy. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference (DIS)*. ACM.

Korsah, A. (2025). The impact of generative artificial intelligence in game development: A scoping review. Master's thesis, University of Oulu, Oulu, Finland.

Kumaran, V., Rowe, J., Mott, B., e Lester, J. (2023). SceneCraft: Automating interactive narrative scene generation in digital games with large language models. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 19(1):86–96.

Lao, A. (2025). A systematic review of AI-driven game design and user experience research. *Humanities and Social Science Research*, 8(4):229.

Li, Z. (2025). Design and application of a dynamic NPC system based on local LLMs. *Transactions on Computer Science and Information Systems Research*.

Liu, J., Snodgrass, S., Khalifa, A., Yannakakis, G. N., e Togelius, J. (2021). Deep learning for procedural content generation. *Neural Computing and Applications*, 33:19–37.

Liu, X. et al. (2025). Cross-modal AIGC integration in Unity3D for game art creation. *Electronics*, 14(6).

Meira-Rodríguez, J. (2025). Creación de ambientes para niveles de videojuegos utilizando Stable Diffusion. *European Public and Social Innovation Review*.

Muñoz-Leiva, F., Viedma-del Jesús, M. I., Sánchez-Fernández, J., e López-Herrera, A. G. (2012). An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective. *Quality and Quantity*, 46(4):1077–1095.

Ni, X., Liu, J., Cao, X., e Wang, C. (2019). Tracking the dynamics of co-word networks for emerging topic identification. *Information Processing and Management*, 56(4):102043.

Nie, Y., Middleton, M., Merino, T., Kanagaraja, N., Kumar, A., Zhuang, Z., e Togelius, J. (2025). Moonshine: Distilling game content generators into steerable generative models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(13):14344–14351.

NVIDIA (2024). Edify 3D: Scalable high-quality 3D asset generation. arXiv preprint arXiv:2411.07135.

Özkan, M. B. (2025). Procedural game level design with deep reinforcement learning. arXiv preprint arXiv:2510.15120.

Radhakrishnan, S., Erbis, S., Isaacs, J. A., e Kamarthi, S. (2017). Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature. *PLoS ONE*, 12(3):e0172778.

Schrum, J. et al. (2020). Interactive evolution and exploration within latent level-design space of generative adversarial networks. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*. ACM.

Singh, A. et al. (2024a). Systematic review of GAN for enhancing efficiency in AI in gaming. In *Proceedings of the IEEE ACROSET*.

Singh, R. et al. (2024b). Enhancing game development: A comparative analysis of image generative AI. In *Proceedings of the IEEE MetroXRAINE*.

Sudhakaran, S., González-Duque, M., Glanois, C., Najarro, E., Freiberger, S., e Risi, S. (2023). MarioGPT: Open-ended text2level generation through large language models. arXiv preprint arXiv:2302.05981.

Sweetser, P. (2024). Large language models and video games: A preliminary scoping review. arXiv preprint arXiv:2403.02613.

Ternar, D.-A., Denisova, A., Cunha, J. M., Kultima, A., e Guckelsberger, C. (2026). Generative AI in game development: A qualitative research synthesis. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. ACM.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7):467–473.

Trukhin, R., Frenkel Isaksson, J., e Eladhari, M. P. (2025). LLM-powered NPCs: Evaluating the impact of large language models on NPC design. In *Interactive Storytelling (ICIDS)*, volume 16375 of *Lecture Notes in Computer Science*, pages 171–189. Springer.

van Eck, N. J. e Waltman, L. (2014). Visualizing bibliometric networks. In Ding, Y., Rousseau, R., e Wolfram, D., editors, *Measuring Scholarly Impact: Methods and Practice*, pages 285–320. Springer, Cham.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., e Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 5998–6008.

Wu, X. (2025). Leveraging Stable Diffusion for enhanced game asset generation pipelines. *International Journal of Advance in Applied Science Research*, 4(12).

Wu, Z., Chen, Z., Zhu, D., Mousas, C., e Kao, D. (2025). A systematic review of generative AI on game character creation: Applications, challenges, and future trends. *IEEE Transactions on Games*.

Xia, Y. (2025). Generative AI in game development: A bibliometric study with an auditable computing workflow. In *Proceedings of the 2025 International Conference on Artificial Intelligence, Virtual Reality and Interaction Design*. ACM.

Yang, D., Kleinman, E., e Harteveld, C. (2024). GPT for games: A scoping review (2020–2023). In *2024 IEEE Conference on Games (CoG)*, pages 1–8.

Yang, D., Kleinman, E., e Harteveld, C. (2025). GPT for games: An updated scoping review (2020–2024). *IEEE Transactions on Games*.

Zhang, J., Xie, J., Hou, W., Tu, X., Xu, J., Song, F., Wang, Z., e Lu, Z. (2012). Mapping the knowledge structure of research on patient adherence: knowledge domain visualization based co-word analysis and social network analysis. *PLoS ONE*, 7(4):e34497.