

PROMPT ENGINEERING E MODELOS DE LINGUAGEM EM JOGOS DIGITAIS: UM MAPEAMENTO SISTEMÁTICO DA LITERATURA

PROMPT ENGINEERING AND LANGUAGE MODELS IN DIGITAL GAMES: A SYSTEMATIC MAPPING STUDY

INGENIERÍA DE PROMPTS Y MODELOS DE LENGUAJE EN JUEGOS DIGITALES: UN MAPEO SISTEMÁTICO DE LA LITERATURA

Luis Cuevas Rodríguez¹
Cristina Souza de Araújo²
Clairon Lima Pinheiro³
Jucimar Maia da Silva Júnior⁴

RESUMO: Este artigo apresenta um mapeamento sistemático da literatura sobre o uso de *Prompt Engineering*, modelos de linguagem de grande escala (*Large Language Models* – LLMs) e Inteligência Artificial generativa em jogos digitais. O objetivo é identificar aplicações, técnicas de *prompting*, modelos, ferramentas, benefícios, desafios e lacunas de pesquisa relacionados ao uso dessas tecnologias no desenvolvimento de jogos. A estratégia de busca considerou estudos publicados entre 2023 e maio de 2026 e resultou em um corpus preliminar de 35 trabalhos, dos quais 15 foram selecionados para análise detalhada. Os resultados permitiram organizar uma taxonomia das aplicações, contemplando narrativa e *storytelling* interativo, NPCs e diálogos dinâmicos, *quests* e conteúdo para RPGs, geração procedural de conteúdo jogável, *game design* e co-criação, *pipelines* de desenvolvimento de jogos, jogos educacionais e *gameplay* com agentes baseados em LLMs. A análise indica que o *Prompt Engineering* atua como mecanismo de especificação, controle, validação, recuperação de contexto e integração com motores de jogo. Também foram identificados desafios relacionados à coerência narrativa, *hallucination*, latência, custo computacional, autoria humana e ausência de métricas padronizadas de avaliação. Como contribuição, o artigo organiza os estudos analisados e propõe uma taxonomia das aplicações identificadas.

Palavras-chave: Prompt Engineering. Modelos de Linguagem. Jogos Digitais.

¹Doutor em Ciência da Computação e Inteligência Artificial, Universidade do Estado do Amazonas, Manaus, Amazonas, Brasil.

²Doutora em Engenharia de Software, Universidade do Estado do Amazonas, Manaus, Amazonas, Brasil.

³Doutor em Engenharia Química, Universidade do Estado do Amazonas, Manaus, Amazonas, Brasil.

⁴Doutor em Engenharia Elétrica, Universidade do Estado do Amazonas, Manaus, Amazonas, Brasil.

ABSTRACT: This paper presents a systematic mapping study on the use of Prompt Engineering, Large Language Models (LLMs), and generative artificial intelligence in digital games. The study aims to identify applications, prompting techniques, models, tools, benefits, challenges, and research gaps related to the use of these technologies in game development. The search strategy considered studies published between 2023 and May 2026 and resulted in a preliminary corpus of 35 works, of which 15 were selected for detailed analysis. The results supported the organization of a taxonomy of applications covering interactive narrative and storytelling, dynamic NPCs and dialogues, quests and RPG content, procedural generation of playable content, game design and co-creation, game development pipelines, educational games, and gameplay with LLM-based agents. The analysis indicates that Prompt Engineering acts as a mechanism for specification, control, validation, context retrieval, and integration with game engines. Challenges were also identified regarding narrative coherence, hallucination, latency, computational cost, human authorship, and the lack of standardized evaluation metrics. As a contribution, the paper organizes the analyzed studies and proposes a taxonomy of the identified applications.

Keywords: Prompt Engineering. Large Language Models. Digital Games.

RESUMEN: Este artículo presenta un mapeo sistemático de la literatura sobre el uso de Ingeniería de Prompts, modelos de lenguaje de gran escala (*Large Language Models* – LLMs) e Inteligencia Artificial generativa en juegos digitales. El objetivo es identificar aplicaciones, técnicas de *prompting*, modelos, herramientas, beneficios, desafíos y vacíos de investigación relacionados con el uso de estas tecnologías en el desarrollo de juegos. La estrategia de búsqueda consideró estudios publicados entre 2023 y mayo de 2026 y dio como resultado un corpus preliminar de 35 trabajos, de los cuales 15 fueron seleccionados para análisis detallado. Los resultados permitieron organizar una taxonomía de aplicaciones que incluye narrativa y *storytelling* interactivo, NPCs y diálogos dinámicos, *quests* y contenido para RPGs, generación procedimental de contenido jugable, *game design* y co-creación, *pipelines* de desarrollo de juegos, juegos educativos y *gameplay* con agentes basados en LLMs. El análisis indica que la Ingeniería de Prompts actúa como mecanismo de especificación, control, validación, recuperación de contexto e integración con motores de juego. También se identificaron desafíos relacionados con coherencia narrativa, alucinaciones, latencia, costo computacional, autoría humana y ausencia de métricas estandarizadas de evaluación. Como contribución, el artículo organiza los estudios analizados y propone una taxonomía de las aplicaciones identificadas.

Palabras clave: Ingeniería de Prompts. Modelos de Lenguaje. Juegos Digitales.

INTRODUÇÃO

A Inteligência Artificial (IA) possui uma trajetória consolidada no desenvolvimento de jogos digitais, sendo empregada no controle de personagens não jogáveis, adaptação de dificuldade, tomada de decisão, geração procedural de conteúdo e suporte a processos de design. Os jogos constituem um domínio de aplicação e experimentação de métodos de IA por combinarem regras, interação, simulação e experiência do usuário (YANNAKAKIS GN e TOGELIUS J, 2025). Entre essas aplicações, a Geração Procedural de Conteúdo, conhecida como *Procedural Content Generation* (PCG), utiliza métodos computacionais para produzir ou apoiar a criação de elementos como níveis, mapas, itens, personagens, narrativas e regras (HENDRIKX M, et al., 2013).

O desenvolvimento recente dos modelos de linguagem de grande escala, ou *Large Language Models* (LLMs), ampliou as possibilidades de uso da linguagem natural em sistemas computacionais. Grande parte desses modelos utiliza como base a arquitetura Transformer, proposta por Vaswani A, et al. (2017), que introduziu mecanismos de atenção para o processamento de sequências. Posteriormente, modelos como GPT-3 demonstraram a capacidade de realizar diferentes tarefas a partir de instruções e exemplos fornecidos no próprio contexto de entrada, característica associada ao *in-context learning* (BROWN TB, et al., 2020). Atualmente, modelos como GPT, LLaMA e Gemini vêm sendo utilizados em tarefas de geração de texto, produção de código, síntese de informações e apoio a processos criativos.

Nesse contexto, o *Prompt Engineering* pode ser compreendido como o processo de elaboração, estruturação e refinamento das instruções fornecidas aos modelos de linguagem. A formulação do *prompt* pode especificar contexto, restrições, formato de saída, exemplos e critérios esperados para uma resposta. Entre as estratégias discutidas na literatura estão o *zero-shot prompting*, no qual a tarefa é apresentada sem exemplos; o *few-shot prompting*, que incorpora exemplos no contexto; e o *Chain-of-Thought*, que busca orientar a decomposição de problemas em etapas intermediárias de raciocínio (BROWN TB, et al., 2020; WEI J, et al., 2022). Dessa forma, o *Prompt Engineering* atua como mecanismo de mediação entre a intenção do usuário e o comportamento do modelo.

No desenvolvimento de jogos digitais, essa interação por linguagem natural vem sendo investigada em tarefas como geração de narrativas, diálogos de personagens não jogáveis (NPCs), criação de *quests*, geração procedural de níveis, produção de *scripts*, prototipação e apoio ao *game design*. Estudos recentes apresentam, por exemplo, sistemas para geração de cenas narrativas interativas, conteúdo para RPGs, NPCs conversacionais, níveis jogáveis e ferramentas de co-criação baseadas em LLMs (ASHBY T, et al., 2023; KUMARAN V, et al., 2023; SISSLER J, 2024; BUONGIORNO S, et al., 2024; KUMARAN V, et al., 2024). Além da geração direta de conteúdo, alguns trabalhos integram modelos de linguagem com mecanismos de memória, recuperação de contexto, *Retrieval-Augmented Generation* (RAG), *function calling*, sistemas de validação e motores de jogo, buscando controlar as saídas produzidas e adequá-las ao estado e às regras do jogo (KOSTILAINEN S, 2024; GALLOTTA R, et al., 2024).

Essas aplicações também apresentam limitações. Respostas inconsistentes, dificuldade de manutenção de contexto, *hallucination*, latência, custo computacional e controle do conteúdo gerado podem comprometer a coerência narrativa, a jogabilidade e a integração com sistemas

de jogo. Além disso, o emprego de IA generativa em processos criativos levanta questões relacionadas à autoria, curadoria, avaliação e participação humana na produção dos artefatos. A variedade de aplicações e de formas de avaliação torna necessária uma análise que permita identificar como essas tecnologias estão sendo utilizadas e quais estratégias vêm sendo adotadas para lidar com essas limitações.

A literatura recente já apresenta esforços de sistematização do tema. Yang D, et al. (2026) realizam uma revisão de escopo sobre aplicações de GPT em jogos, contemplando geração procedural de conteúdo, design e *gameplay mixed-initiative*, agentes e pesquisas com usuários. Salmaze PH, et al. (2026), por sua vez, apresentam um mapeamento sistemático voltado especificamente à geração de narrativas e diálogos, identificando desafios como incoerência narrativa, repetitividade, limitações de memória, latência e ausência de métodos padronizados de avaliação. Embora esses trabalhos contribuam para a organização da literatura, permanece espaço para uma análise centrada especificamente no papel do *Prompt Engineering* e nas diferentes formas pelas quais as estratégias de *prompting* são utilizadas na criação, controle, validação e integração de conteúdos e sistemas em jogos digitais.

Diante desse cenário, este artigo apresenta um mapeamento sistemático da literatura com o objetivo de identificar as aplicações de *Prompt Engineering*, LLMs e Inteligência Artificial generativa em jogos digitais, as técnicas de *prompting* empregadas, os modelos e ferramentas utilizados, bem como os benefícios, desafios, limitações e lacunas apontados pelos estudos. Como contribuição, o trabalho organiza as aplicações identificadas em uma taxonomia e analisa o *Prompt Engineering* não apenas como forma de elaboração de instruções, mas também como mecanismo de especificação, controle, recuperação de contexto, validação e integração entre modelos generativos, motores de jogo e processos de desenvolvimento.

MÉTODOS

Este estudo foi conduzido como um mapeamento sistemático da literatura, estruturado com base nas diretrizes propostas por Petersen K, et al. (2015), com o objetivo de organizar e analisar pesquisas relacionadas ao uso de *Prompt Engineering*, modelos de linguagem de grande escala (*Large Language Models* – LLMs) e Inteligência Artificial generativa em jogos digitais. A unidade de análise foi composta por estudos científicos e acadêmicos que abordam essas tecnologias em atividades de desenvolvimento, design, geração de conteúdo ou interação em jogos.

A escolha pelo mapeamento sistemático foi motivada pela diversidade de aplicações, técnicas, ferramentas e formas de avaliação encontradas na literatura recente. Esse tipo de estudo permite identificar tendências, categorias de aplicação, lacunas e oportunidades de pesquisa em áreas ainda em consolidação. Como suporte contextual e comparativo, também foram considerados estudos secundários recentes sobre LLMs e jogos digitais, sem incorporá-los ao corpus principal analisado.

Questões de pesquisa

Para orientar o processo de busca, seleção, extração e análise dos estudos, foram definidas seis questões de pesquisa:

QP1: Quais aplicações de *Prompt Engineering*, modelos de linguagem e IA generativa são identificadas no contexto de jogos digitais?

QP2: Quais técnicas de *prompting* são identificadas nos estudos relacionados a jogos digitais?

QP3: Quais modelos de linguagem, ferramentas, plataformas e mecanismos de integração são empregados nos trabalhos analisados?

QP4: Quais benefícios e contribuições do uso de *Prompt Engineering*, modelos de linguagem e IA generativa são relatados para o desenvolvimento de jogos digitais?

QP5: Quais desafios, limitações e riscos são apontados nos estudos relacionados ao uso de IA generativa em jogos digitais?

QP6: Quais lacunas de pesquisa e oportunidades para trabalhos futuros podem ser identificadas na literatura analisada?

Estratégia de busca

As buscas foram realizadas em maio de 2026 nas bases ACM Digital Library, IEEE Xplore, SpringerLink, Scopus e ScienceDirect. O Google Scholar foi utilizado como mecanismo complementar de busca. O recorte temporal adotado para a composição do corpus compreendeu o período entre 2023 e maio de 2026.

A estratégia de busca foi construída a partir da combinação de termos relacionados a *Prompt Engineering*, modelos de linguagem e jogos digitais. A *string* principal utilizada é apresentada a seguir (Tabela 1).

Tabela 1 - *String* de busca utilizada no estudo.

<i>String</i> de Busca
("prompt engineering" OR "prompting techniques" OR "large language models" OR "LLMs") AND ("digital games" OR "video games" OR "game development" OR "game design")

Fonte: Elaborada pelos autores (2026).

Além da *string* principal, foram utilizadas variações complementares durante o processo exploratório, incluindo termos relacionados à geração narrativa, diálogos de NPCs, narrativa interativa, Inteligência Artificial generativa e desenvolvimento assistido por IA. O objetivo dessa estratégia foi ampliar a cobertura da busca diante da diversidade terminológica encontrada na área.

O recorte temporal de 2023 a maio de 2026 foi aplicado somente aos estudos candidatos à composição do corpus do mapeamento. Referências anteriores a esse período foram utilizadas exclusivamente para fundamentação conceitual. Estudos secundários, como revisões e outros mapeamentos da literatura, também não foram incluídos no conjunto de estudos primários analisados em detalhe, sendo utilizados apenas para contextualização e comparação dos resultados.

6

Critérios de inclusão e exclusão

Foram incluídos estudos relacionados ao uso de *Prompt Engineering*, técnicas de *prompting*, LLMs ou IA generativa aplicados ao contexto de jogos digitais, desenvolvimento de jogos, geração de conteúdo ou interação baseada em linguagem natural. Também foram considerados trabalhos completos publicados em periódicos, conferências ou *workshops*, capítulos de livros e dissertações acadêmicas, em português ou inglês, desde que apresentassem descrição metodológica, aplicação prática, análise ou discussão relacionada ao tema.

Foram excluídos trabalhos duplicados; estudos sem relação direta com jogos digitais; trabalhos sobre *Prompt Engineering* aplicados a outros domínios; documentos sem conteúdo técnico-científico suficiente; trabalhos sem acesso ao texto completo; e estudos que apresentassem informações insuficientes para responder às questões de pesquisa.

Processo de seleção dos estudos

O processo de seleção foi realizado em etapas sucessivas. Inicialmente, foram conduzidas buscas exploratórias nas fontes selecionadas, considerando o recorte temporal definido. Em seguida, os resultados foram avaliados por meio da leitura de títulos, resumos e palavras-chave, observando sua relação com jogos digitais, modelos de linguagem, *Prompt Engineering*, IA generativa, geração de conteúdo, narrativas, NPCs, *quests*, níveis ou desenvolvimento assistido por IA.

Após essa triagem inicial, foi constituído um corpus preliminar de 35 estudos considerados potencialmente relacionados ao escopo da pesquisa. Esses trabalhos foram submetidos aos critérios de inclusão e exclusão. Como resultado, 20 estudos não foram incluídos na análise detalhada por não atenderem integralmente aos critérios estabelecidos ou por apresentarem relação indireta com as questões de pesquisa. Ao final, 15 estudos primários foram selecionados para análise detalhada. A triagem inicial e a aplicação dos critérios de inclusão e exclusão foram realizadas pelo primeiro autor e posteriormente revisadas pelos coautores. Eventuais divergências na seleção dos estudos foram discutidas até a obtenção de consenso.

Além desses trabalhos, foram considerados dois estudos secundários recentes sobre LLMs e jogos digitais: a revisão de escopo de Yang D, et al. (2026), sobre aplicações de GPT em jogos, e o mapeamento sistemático de Salmaze PH, et al. (2026), voltado à geração de narrativas e diálogos. Esses estudos foram utilizados exclusivamente como apoio à contextualização e comparação dos achados, sem integrar o corpus principal (Figura 1).

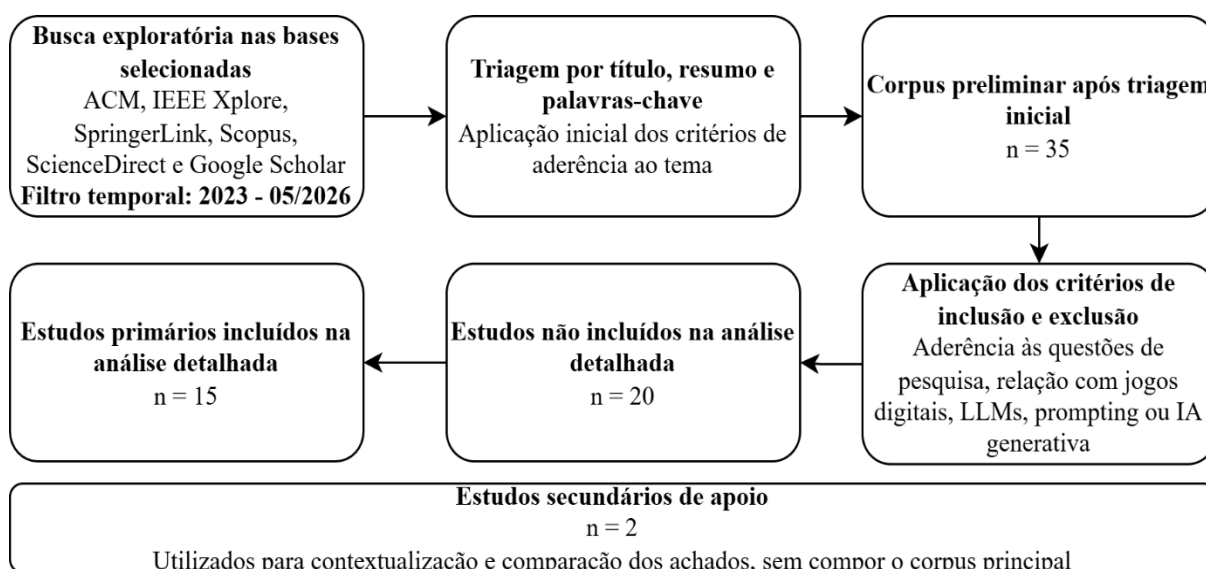


Figura 1 - Processo de composição do corpus analisado.

Fonte: Elaborada pelos autores (2026).

Extração e análise dos dados

Após a seleção dos estudos, os dados foram extraídos com base nas questões de pesquisa. Para cada trabalho selecionado, foram coletadas informações referentes ao ano de publicação, tipo de publicação, contexto de aplicação, técnicas de *prompting*, modelos de linguagem, ferramentas ou plataformas utilizadas, aplicações em jogos digitais, benefícios relatados, desafios, limitações e oportunidades de pesquisa futura.

Os dados foram organizados em tabelas comparativas e analisados de forma qualitativa e descritiva. A partir dessa análise, os estudos foram agrupados de forma indutiva em oito categorias: narrativa e *storytelling* interativo; NPCs e diálogos dinâmicos; *quests* e conteúdo para RPGs; geração procedural de conteúdo jogável; *game design* e co-criação; *pipelines* de desenvolvimento de jogos; jogos educacionais; e *gameplay* com agentes baseados em LLMs. As categorias foram construídas considerando o foco dos estudos, os artefatos produzidos, o papel desempenhado pelos modelos de linguagem e a etapa do desenvolvimento de jogos em que as tecnologias foram empregadas.

Por se tratar de um estudo exclusivamente bibliográfico, baseado em publicações científicas e trabalhos acadêmicos previamente disponíveis, não houve participação de indivíduos nem realização de procedimentos experimentais envolvendo participantes.

RESULTADOS

Após a aplicação da estratégia de busca, do recorte temporal entre 2023 e maio de 2026 e da triagem inicial dos resultados, foi constituído um corpus preliminar de 35 estudos. Após a aplicação dos critérios de inclusão e exclusão, 15 estudos foram selecionados para análise detalhada. Desse conjunto, quatro estudos foram publicados em 2023, oito em 2024 e três em 2025.

Os trabalhos selecionados abordam diferentes aplicações de *Prompt Engineering*, modelos de linguagem de grande escala e Inteligência Artificial generativa no contexto de jogos digitais. Entre os objetos investigados estão narrativas e diálogos, NPCs, *quests*, geração procedural de níveis, agentes de jogo, prototipação, ferramentas de apoio ao *game design* e integração de modelos generativos ao processo de desenvolvimento. Os estudos analisados e seus respectivos focos são apresentados a seguir (Tabela 2).

Tabela 2 - Estudos selecionados para análise detalhada.

ID	Estudo	Ano	Foco
Eo1	Buongiorno S, et al. – PANGeA	2024	Narrativa procedural em RPGs
Eo2	Kumaran V, et al. – SceneCraft	2023	Geração de cenas narrativas interativas
Eo3	Ashby T, et al. – Personalized Quest and Dialogue Generation	2023	Geração personalizada de <i>quests</i> e diálogos em RPGs
Eo4	Hardiman JPW, et al. – Diálogos e <i>quests</i> com Gemini e Sentence-BERT	2024	Geração de diálogos e <i>quests</i>
Eo5	Gallotta R, et al. – LLMaker	2024	Geração de conteúdo com <i>function calling</i>
Eo6	Sissler J – NPCs em Unity 3D com GPT-3.5	2024	NPCs conversacionais integrados ao Unity 3D
Eo7	Kostilainen S – NPC Dialogue com RAG	2024	NPCs responsivos com RAG e dados do jogador
Eo8	Taveekitworachai P, et al. – ChatGPT4PCG	2023	Geração procedural de níveis para Science Birds
Eo9	Earle S, et al. – ScriptDoctor	2025	Geração automática de jogos com LLMs e busca em árvore
E10	Earle S, et al. – DreamGarden	2025	Assistência ao design de jogos baseada em <i>prompt</i>
E11	Grow A e Khosmood F – ChatGPT Game Jam	2023	Uso de LLMs em <i>game jams</i> e prototipação
E12	Alharthi SA – Generative AI in Game Design	2025	IA generativa aplicada ao <i>game design</i>
E13	Begemann A e Hutson J – AI-assisted Game Development	2024	IA generativa em <i>pipelines</i> de desenvolvimento
E14	Kumaran V, et al. – Procedural Level Generation in Educational Games	2024	Geração de níveis educacionais por linguagem natural
E15	Sidji M e Stephenson M – Prompt Engineering for Codenames	2024	Técnicas de <i>prompting</i> aplicadas a agentes de jogo

Fonte: Elaborada pelos autores (2026).

Síntese dos achados por questão de pesquisa

Em relação às aplicações de *Prompt Engineering*, modelos de linguagem e IA generativa em jogos digitais (QP₁), foram identificadas oito categorias: narrativa e *storytelling* interativo; NPCs e diálogos dinâmicos; *quests* e conteúdo para RPGs; geração procedural de conteúdo

jogável; *game design* e co-criação; *pipelines* de desenvolvimento de jogos; jogos educacionais; e *gameplay* com agentes baseados em LLMs.

Quanto às técnicas de *prompting* (QP₂), foram identificadas estratégias de *zero-shot prompting*, *few-shot prompting*, *Chain-of-Thought*, *Self-Refine*, refinamento iterativo e *prompts* estruturados. Essas estratégias aparecem tanto de forma explicitamente nomeada nos estudos quanto incorporadas aos procedimentos de interação com os modelos.

Em relação aos modelos, ferramentas, plataformas e mecanismos de integração (QP₃), foram identificados modelos generativos das famílias GPT, incluindo GPT-2, GPT-3.5, GPT-4 e GPT-4o, além de Llama-3 8B e Gemini 1.0 Pro. Sentence-BERT aparece como modelo de representação semântica. Também foram empregados mecanismos como RAG, *function calling*, *knowledge graphs*, *embeddings*, sistemas de memória e procedimentos de validação. Entre as ferramentas e plataformas aparecem Unity, Unreal Engine, PuzzleScript, Science Birds, Future Worlds, LangChain, PyTorch, Transformers e APIs de modelos de linguagem.

Os benefícios e contribuições relatados (QP₄) incluem apoio à geração de conteúdo narrativo e jogável, criação de diálogos e *quests*, prototipação, geração de níveis, suporte ao processo de design, geração de código e integração de conteúdos aos motores de jogo.

Entre os desafios, limitações e riscos (QP₅), foram identificadas questões relacionadas à coerência narrativa, manutenção de contexto e memória, *hallucination*, latência, custo computacional, controle das saídas, autoria e avaliação dos conteúdos gerados.

Quanto às lacunas e oportunidades de pesquisa (QP₆), os estudos apontam ausência de métricas padronizadas para avaliação dos conteúdos gerados, necessidade de ampliar avaliações com jogadores e usuários finais e necessidade de comparações sistemáticas entre estratégias de *prompting*. Uma síntese dos resultados de acordo com as questões de pesquisa é apresentada a seguir (Tabela 3).

Tabela 3 - Síntese dos resultados por questão de pesquisa.

Questão	Dimensão analisada	Resultados identificados
QP ₁	Aplicações	Narrativa e <i>storytelling</i> interativo; NPCs e diálogos dinâmicos; <i>quests</i> e conteúdo para RPGs; geração procedural de conteúdo jogável; <i>game design</i> e co-criação; <i>pipelines</i> de desenvolvimento de jogos; jogos educacionais; <i>gameplay</i> com agentes baseados em LLMs
QP ₂	Técnicas de <i>prompting</i>	<i>Zero-shot</i> ; <i>few-shot</i> ; <i>Chain-of-Thought</i> ; <i>Self-Refine</i> ; refinamento iterativo; <i>prompts</i> estruturados
QP ₃	Modelos, ferramentas, plataformas e integração	GPT-2, GPT-3.5, GPT-4, GPT-4o, Llama-3 8B, Gemini 1.0 Pro; Sentence-BERT; RAG; <i>function calling</i> ; <i>knowledge graphs</i> ; <i>embeddings</i> ; sistemas de memória; Unity; Unreal Engine; PuzzleScript; Science Birds; Future Worlds; LangChain; PyTorch e Transformers

QP4	Benefícios e contribuições	Geração de conteúdo, diálogos, <i>quests</i> e níveis; prototipação; suporte ao design; geração de código; integração com motores de jogo
QP5	Desafios, limitações e riscos	Coerência narrativa, memória e contexto, <i>hallucination</i> , latência, custo computacional, controle das saídas, autoria e avaliação
QP6	Lacunas de pesquisa	Ausência de métricas padronizadas; necessidade de ampliar avaliações com usuários; necessidade de comparações sistemáticas entre técnicas de <i>prompting</i>

Fonte: Elaborada pelos autores (2026).

DISCUSSÃO

Os resultados deste mapeamento mostram que o uso de *Prompt Engineering*, modelos de linguagem de grande escala e Inteligência Artificial generativa em jogos digitais abrange diferentes etapas do processo de criação e desenvolvimento. As aplicações identificadas incluem geração de narrativas, diálogos de NPCs, *quests*, níveis, *scripts*, prototipação, apoio ao *game design* e desenvolvimento de agentes de jogo. Nesse conjunto de aplicações, a linguagem natural assume o papel de interface entre a intenção humana e os sistemas generativos, permitindo que designers, desenvolvedores, educadores e jogadores expressem objetivos, restrições e características desejadas para os conteúdos produzidos.

Essa diversidade de aplicações é coerente com a revisão de escopo realizada por Yang D, et al. (2026), que organiza o uso de GPT em jogos em áreas como geração procedural de conteúdo, design *mixed-initiative*, *gameplay mixed-initiative*, agentes que jogam e pesquisas com usuários. As categorias identificadas neste estudo apresentam correspondência com parte dessa organização, especialmente nos usos relacionados à geração procedural, co-criação, apoio ao design e agentes baseados em LLMs. Entretanto, o presente mapeamento concentra-se especificamente no papel do *Prompt Engineering*, observando como a formulação e estruturação de instruções participa não apenas da geração de conteúdo, mas também da especificação, controle e integração dos modelos com outros componentes dos sistemas de jogo.

A análise das técnicas de *prompting* reforça essa função. Estratégias como *zero-shot*, *few-shot*, *Chain-of-Thought*, *Self-Refine* e *prompts* estruturados são empregadas para orientar a interação com os modelos. Em parte dos sistemas analisados, essas estratégias são combinadas com mecanismos complementares de recuperação de contexto e integração, como RAG e *function calling*. Em aplicações como ChatGPT4PCG e Codenames, a própria estratégia de *prompting* constitui uma variável do funcionamento do sistema (TAVEEKITWORACHAI P, et al., 2023; SIDJI M e STEPHENSON M, 2024). Em outros trabalhos, o *prompt* é integrado a

regras, mecanismos de memória, sistemas de recuperação de informação ou processos de validação. Dessa forma, o *Prompt Engineering* deixa de representar apenas a elaboração de comandos em linguagem natural e passa a atuar como parte da arquitetura computacional utilizada para controlar a geração.

Outro resultado observado é a presença de arquiteturas híbridas. Os estudos analisados apresentam diferentes formas de integração dos LLMs com outros componentes computacionais. Entre as configurações identificadas, os LLMs aparecem integrados a motores de jogo, APIs, bases de conhecimento, sistemas de recuperação de contexto, mecanismos de memória, *function calling*, validação automática, *feedback* de compilação e agentes de teste. Exemplos como PANGeA, LLMaker, ScriptDoctor e abordagens baseadas em RAG ilustram essa combinação entre modelos generativos e mecanismos externos de controle (BUONGIORNO S, et al., 2024; GALLOTTA R, et al., 2024; EARLE S, et al., 2025; KOSTILAINEN S, 2024). Essa integração busca limitar inconsistências, incorporar informações específicas do contexto e permitir que os conteúdos gerados possam ser utilizados ou avaliados dentro do ambiente do jogo.

A geração narrativa e a interação com NPCs constituem outro conjunto de aplicações identificado no corpus analisado. Sistemas como PANGeA, SceneCraft, Personalized Quest and Dialogue Generation e abordagens baseadas em RAG exploram a geração de histórias, diálogos, eventos narrativos e respostas contextualizadas (BUONGIORNO S, et al., 2024; KUMARAN V, et al., 2023; ASHBY T, et al., 2023; KOSTILAINEN S, 2024). Esses resultados são compatíveis com o mapeamento de Salmaze PH, et al. (2026), dedicado especificamente à geração de narrativas e diálogos em jogos digitais. Os autores identificam problemas como incoerência narrativa, repetitividade, limitações de memória e latência, desafios também encontrados no corpus deste estudo. Em aplicações interativas, tais limitações assumem importância particular, pois as respostas do modelo precisam permanecer coerentes com personagens, acontecimentos anteriores, regras e estado atual do jogo.

Os benefícios relatados nos estudos incluem apoio à geração de conteúdo, prototipação e interação. O emprego de LLMs permite produzir ou auxiliar a produção de narrativas, diálogos, *quests*, níveis, código e outros artefatos utilizados no desenvolvimento. Em ferramentas voltadas ao processo criativo, os modelos também podem apoiar ideação e exploração de alternativas. Entretanto, esses benefícios coexistem com questões relacionadas à autoria e ao controle humano. Estudos sobre *game jams*, *pipelines* criativos e ferramentas de

design assistido indicam que a automação pode modificar a participação do desenvolvedor no processo de criação, deslocando parte de seu trabalho para atividades de especificação, seleção, refinamento e validação dos conteúdos gerados (GROW A e KHOSMOOD F, 2023; BEGEMANN A e HUTSON J, 2024; ALHARTHI SA, 2025).

Além das questões criativas, foram identificados desafios técnicos relacionados à *hallucination*, latência, custo computacional, manutenção de contexto e controle das saídas. Esses problemas ajudam a explicar a adoção de mecanismos externos de recuperação, memória e validação observada nos estudos. A combinação entre LLMs e componentes adicionais pode reduzir determinadas limitações, mas também introduz maior complexidade na arquitetura dos sistemas. Assim, a avaliação de aplicações baseadas em LLMs em jogos não depende apenas da qualidade textual das respostas, mas também de aspectos como validade do conteúdo, jogabilidade, consistência com o estado do jogo e tempo de resposta.

Outra limitação observada diz respeito aos procedimentos de avaliação. Parte dos trabalhos analisados possui caráter exploratório, experimental ou baseado em protótipos, incluindo *frameworks*, sistemas demonstrativos e estudos de caso. Nem todos realizam avaliações extensas com jogadores, designers profissionais ou usuários finais. Essa característica também é apontada por Salmaze PH, et al. (2026), que destacam a necessidade de métodos padronizados de avaliação e de comparações entre conteúdos produzidos por LLMs e aqueles desenvolvidos por humanos. No corpus deste mapeamento, diferentes critérios são empregados, incluindo coerência, criatividade, fluência, estabilidade, validade, jogabilidade, tempo de resposta, satisfação do usuário e similaridade com objetivos definidos.

A heterogeneidade desses critérios dificulta a comparação direta entre modelos, técnicas de *prompting* e arquiteturas. Nesse sentido, a ausência de métricas e procedimentos padronizados constitui uma lacuna identificada no corpus analisado. Trabalhos futuros podem explorar conjuntos de métricas capazes de combinar características técnicas e avaliações humanas, permitindo comparar diferentes estratégias de *prompting* em tarefas específicas de jogos digitais. Avaliações com jogadores, designers e desenvolvedores também podem contribuir para compreender de forma mais ampla os efeitos dessas tecnologias sobre a qualidade dos conteúdos e sobre o processo de desenvolvimento.

Como limitação deste próprio mapeamento, destaca-se a utilização de uma estratégia de busca baseada em um conjunto específico de termos e a seleção dos estudos de acordo com o escopo estabelecido. Trabalhos relevantes podem utilizar terminologias diferentes, como

generative agents, *interactive narrative generation*, *AI-assisted game development* ou *mixed-initiative design*, e eventualmente não terem sido recuperados pela estratégia adotada. Além disso, parte da literatura recente é composta por protótipos, dissertações ou estudos exploratórios, característica que deve ser considerada na interpretação dos resultados. Os dois estudos secundários utilizados neste trabalho contribuíram para contextualizar e comparar os achados, mas não integraram o corpus de estudos primários.

De modo geral, os resultados indicam que o *Prompt Engineering* vem assumindo diferentes funções no desenvolvimento de jogos digitais. Além de orientar a geração de respostas dos LLMs, os *prompts* são empregados para especificar conteúdos, incorporar contexto, estabelecer restrições e conectar modelos generativos a sistemas computacionais. O desenvolvimento dessa área, entretanto, permanece associado a desafios de avaliação, controle, autoria, desempenho e integração, indicando a necessidade de pesquisas que combinem avanços técnicos com procedimentos sistemáticos de validação em situações reais de desenvolvimento e uso de jogos digitais.

CONCLUSÃO

Este artigo apresentou um mapeamento sistemático da literatura sobre o uso de *Prompt Engineering*, modelos de linguagem de grande escala e Inteligência Artificial generativa em jogos digitais. A partir de um corpus preliminar de 35 estudos, 15 trabalhos foram selecionados para análise detalhada, permitindo identificar aplicações, técnicas de *prompting*, modelos, ferramentas, benefícios, desafios e lacunas de pesquisa.

Os resultados mostram que essas tecnologias vêm sendo empregadas em diferentes contextos, incluindo geração de narrativas, diálogos de NPCs, *quests*, geração procedural de conteúdo, apoio ao *game design*, prototipação, desenvolvimento de jogos educacionais e agentes baseados em LLMs. Também foi possível observar que o *Prompt Engineering* não se limita à formulação de instruções, sendo utilizado como mecanismo de especificação, controle, recuperação de contexto, validação e integração entre modelos generativos e componentes dos sistemas de jogo.

Entre os desafios identificados destacam-se problemas de coerência narrativa, manutenção de contexto, *hallucination*, latência, custo computacional, controle das saídas, autoria e avaliação dos conteúdos gerados. A análise também revelou limitações relacionadas à

ausência de métricas padronizadas e à necessidade de ampliar avaliações com jogadores, designers e usuários finais.

Como trabalhos futuros, recomenda-se ampliar o conjunto de estudos analisados, comparar estratégias de *prompting* em tarefas específicas de jogos digitais e desenvolver procedimentos de avaliação que combinem métricas técnicas e avaliações humanas. Também se mostra relevante investigar de forma mais sistemática a integração entre LLMs, motores de jogo, mecanismos de memória, recuperação de contexto e validação automática.

De modo geral, o estudo contribui para organizar a literatura recente sobre *Prompt Engineering* e LLMs em jogos digitais, propondo uma síntese das aplicações e técnicas identificadas e destacando desafios metodológicos e técnicos que ainda precisam ser investigados para ampliar a adoção dessas abordagens em processos de desenvolvimento de jogos.

DECLARAÇÃO SOBRE O USO DE INTELIGÊNCIA ARTIFICIAL

Durante a elaboração deste manuscrito, foi utilizada a ferramenta ChatGPT (OpenAI) como apoio à estruturação e organização textual, elaboração e reformulação de trechos, revisão linguística, tradução dos resumos e adequação às normas editoriais da revista. As sugestões geradas foram submetidas à revisão e validação humana, permanecendo sob responsabilidade dos autores as decisões metodológicas, a seleção dos estudos, a extração dos dados, a interpretação final dos resultados, as conclusões e a veracidade das informações, citações e referências apresentadas. A ferramenta não foi considerada autora nem utilizada como fonte de evidência científica.

15

REFERÊNCIAS

ALHARTHI, Sultan A. Generative AI in game design: enhancing creativity or constraining innovation? *Journal of Intelligence*, Basel, v. 13, n. 6, art. 60, 2025. DOI: 10.3390/jintelligence13060060.

ASHBY, Trevor et al. Personalized quest and dialogue generation in role-playing games: a knowledge graph- and language model-based approach. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. New York: Association for Computing Machinery, 2023. art. 290, p. 1-20. DOI: 10.1145/3544548.3581441.

BEGEMANN, Andrew; HUTSON, James. Empirical insights into AI-assisted game development: a case study on the integration of generative AI tools in creative pipelines. *Metaverse*, v. 5, n. 2, art. 2568, 2024. DOI: 10.54517/m.v5i2.2568.

BROWN, Tom B. et al. Language models are few-shot learners. In: LAROCHELLE, Hugo et al. (ed.). *Advances in Neural Information Processing Systems* 33. Red Hook: Curran Associates, 2020. p. 1877-1901.

BUONGIORNO, Steph et al. PANGeA: procedural artificial narrative using generative AI for turn-based, role-playing video games. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, v. 20, n. 1, p. 156-166, 2024. DOI: 10.1609/aiide.v20i1.31876.

EARLE, Sam; PARAJULI, Samyak; BANBURSKI-FAHEY, Andrzej. DreamGarden: a designer assistant for growing games from a single prompt. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. New York: Association for Computing Machinery, 2025. art. 57, p. 1-19. DOI: 10.1145/3706598.3714233.

EARLE, Sam et al. ScriptDoctor: automatic generation of PuzzleScript games via large language models and tree search. In: *Proceedings of the 2025 IEEE Conference on Games*. New York: Institute of Electrical and Electronics Engineers, 2025. p. 1-5. DOI: 10.1109/CoG64752.2025.11114269.

GALLOTTA, Roberto; LIAPIS, Antonios; YANNAKAKIS, Georgios N. Consistent game content creation via function calling for large language models. In: *Proceedings of the 2024 IEEE Conference on Games*. New York: Institute of Electrical and Electronics Engineers, 2024. p. 1-4. DOI: 10.1109/CoG60054.2024.10645599.

GROW, April; KHOSMOOD, Foaad. ChatGPT GameJam: unleashing the power of large language models for game jams. In: *Proceedings of the 7th International Conference on Game Jams, Hackathons and Game Creation Events*. New York: Association for Computing Machinery, 2023. p. 51-54. DOI: 10.1145/3610602.3610605.

HARDIMAN, Jason Patrick Winarto et al. AI-powered dialogues and quests generation in role-playing games using Google's Gemini and Sentence BERT framework. *Procedia Computer Science*, v. 245, p. 1111-1119, 2024. DOI: 10.1016/j.procs.2024.10.340.

HENDRIKX, Mark et al. Procedural content generation for games: a survey. *ACM Transactions on Multimedia Computing, Communications, and Applications*, v. 9, n. 1, p. 1-22, 2013. DOI: 10.1145/2422956.2422957.

KOSTILAINEN, Sami. Next generation of NPC dialogue: creating responsive NPCs (Non-Player Characters) with Retrieval-Augmented Generation and real-time player data. 2024. Dissertação (Mestrado em Artificial Intelligence and Data Analytics) – Jamk University of Applied Sciences, Jyväskylä, 2024.

KUMARAN, Vikram et al. 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*, v. 19, n. 1, p. 86-96, 2023. DOI: 10.1609/aiide.v19i1.27504.

KUMARAN, Vikram et al. Procedural level generation in educational games from natural language instruction. *IEEE Transactions on Games*, v. 16, n. 4, p. 937-946, 2024. DOI: 10.1109/TG.2024.3392670.

PETERSEN, Kai; VAKKALANKA, Sairam; KUŹNIARZ, Ludwik. Guidelines for conducting systematic mapping studies in software engineering: an update. *Information and Software Technology*, v. 64, p. 1-18, 2015. DOI: 10.1016/j.infsof.2015.03.007.

SALMAZE, Pedro Henrique et al. Narrative and dialogue generation for video-games: a systematic mapping. *Engineering Applications of Artificial Intelligence*, v. 178, art. 115041, 2026. DOI: 10.1016/j.engappai.2026.115041.

SIDJI, Matthew; STEPHENSON, Matthew. Prompt engineering ChatGPT for Codenames. In: *Proceedings of the 2024 IEEE Conference on Games*. New York: Institute of Electrical and Electronics Engineers, 2024. art. 10645591. DOI: 10.1109/CoG60054.2024.10645591.

SISSLER, John. Enhancing non-player characters in Unity 3D using GPT-3.5. *Games: Research and Practice*, v. 2, n. 3, art. 25, p. 1-16, 2024. DOI: 10.1145/3662003.

TAVEEKITWORACHAI, Pittawat et al. ChatGPT4PCG Competition: character-like level generation for Science Birds. In: *Proceedings of the 2023 IEEE Conference on Games*. New York: Institute of Electrical and Electronics Engineers, 2023. p. 1-8. DOI: 10.1109/CoG57401.2023.10333206.

VASWANI, Ashish et al. Attention is all you need. In: *Advances in Neural Information Processing Systems* 30. Red Hook: Curran Associates, 2017. p. 5998-6008.

WEI, Jason et al. Chain-of-thought prompting elicits reasoning in large language models. In: *Advances in Neural Information Processing Systems* 35. Red Hook: Curran Associates, 2022. p. 24824-24837.

YANG, Daijin; KLEINMAN, Erica; HARTEVELD, Casper. GPT for games: an updated scoping review (2020-2024). *IEEE Transactions on Games*, v. 18, n. 3, p. 432-447, 2026. DOI: 10.1109/TG.2025.3563780.

YANNAKAKIS, Georgios N.; TOGELIUS, Julian. *Artificial Intelligence and Games*. 2. ed. Cham: Springer, 2025. DOI: 10.1007/978-3-031-83347-2.