

# NPCs Interativos Utilizando LLMs: uma Revisão Sistemática da Literatura

Luísa Rodrigues Foppa  
Curso de Design de Games  
Escola Politécnica /UNIVALI  
Florianópolis, SC, Brasil  
[luisa.foppa@edu.univali.br](mailto:luisa.foppa@edu.univali.br)

Dennis Kerr Coelho  
Curso de Design de Games  
Escola Politécnica /UNIVALI  
Florianópolis, SC, Brasil  
[dennis@palmsoft.com.br](mailto:dennis@palmsoft.com.br)

Anita Fernandes  
Curso de Ciência da Computação  
Campus Kobrasol  
Escola Politécnica /UNIVALI  
São José, SC, Brasil  
[anita.fernandes@univali.br](mailto:anita.fernandes@univali.br)

## ABSTRACT

Emerging technologies present opportunities within the digital gaming market, however, they must be explored and understood to be effectively employed. This paper aims to investigate the application of LLMs (Large Language Models) in NPCs (Non-Playable Characters) to enhance interactivity and improve the overall player experience. By utilizing PRISMA 2020 methodology to conduct a systematic literature review, this study identified and analyzed methods and patterns employed to improve NPC and player dialogue. The findings synthesized recurring challenges such as hallucination, repetition, lack of memory and inconsistency, alongside various strategies employed to improve the output generated by the NPC, most of which involved prompt engineering techniques. Finally, this study addressed the lack of research focused on manipulating LLM outputs via code and emphasized how the high computational cost remains a significant barrier to their widespread integration in games.

## KEYWORDS

LLMs, NPCs, Video Games, Interactivity

## 1 INTRODUÇÃO

O desenvolvimento da arquitetura de transformadores permitiu o advento dos grandes modelos de linguagens (LLMs), que são modelos de inteligência artificial (IA) treinados em vastas quantidades de texto, por conta disso, são capazes de capturar a complexidade da linguagem humana e de gerar e interpretar textos ou imagens de maneira eficiente. [1].

Com o lançamento do modelo ChatGPT da OpenAI, marcou-se o início de uma nova era na aplicação desses modelos em diferentes áreas [2]. Essa tendência é amplificada pela busca das empresas por inovação, especialmente na indústria de jogos, onde a novidade é essencial para atrair consumidores [3]. Os LLMs, capazes de gerar diálogos contextualmente relevantes e personalizados, oferecem uma nova fronteira para a criação de experiências de jogo mais imersivas e envolventes. Essa capacidade pode ser utilizada para formular personagens que desempenham um papel crucial na construção da narrativa e na imersão do jogador.

Diante desse potencial, este trabalho tem como foco a utilização dos LLMs aplicados a personagens não-jogáveis (NPCs) em jogos digitais. Para isso, a realização de uma Revisão Sistemática da

Literatura (RSL), permite avaliar estudos e métodos que apliquem esses modelos para aprimorar a imersão e interação do jogador. Esta análise visa identificar padrões de uso, benefícios e possíveis limitações desses modelos nos jogos.

Devido à natureza emergente do tema, há ainda muitas oportunidades de pesquisa sobre o uso de LLMs, especialmente no contexto de jogos digitais, com ênfase na melhoria de interações com NPCs.

Para a organização da pesquisa, emprega-se a estratégia de desmembramento PerSPECTiF [4], ela é utilizada para detalhar os objetivos desta revisão. A seguir, é apresentada a tabela correspondente a aplicação dessa estratégia neste estudo:

Tabela 1: Desmembramento dos objetivos desta revisão.

| Per | Perspectiva           | Desenvolvedores e designers de Jogos Digitais                                                                                                                                         |
|-----|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| S   | Contexto              | Surgimento e rápida popularização no uso de LLMs a partir de 2018, impulsionado por avanços em arquitetura de redes neurais e disponibilidade de grandes conjuntos de dados textuais. |
| P   | Fenômeno de interesse | Aumentar a imersão do jogador através de diálogos mais naturais e reativos com NPCs                                                                                                   |
| E   | Ambiente              | Jogos digitais                                                                                                                                                                        |
| C   | Comparação (opcional) | Não há comparação de grupos                                                                                                                                                           |
| Ti  | Tempo/temporalidade   | 2022 a 2024                                                                                                                                                                           |
| F   | Resultado             | Diferentes <u>técnicas</u> e/ou <u>implementações</u> de LLMs em jogos digitais                                                                                                       |

A coleta de evidências a respeito da aplicação de LLMs possibilita a identificação de padrões e práticas operacionais mais eficientes e vantajosas, além de também permitir reconhecer melhorias e a formulação de estratégias para mitigar eventuais limitações dos modelos.

## 2 METODOLOGIA

A coleta de evidências a respeito da aplicação de LLMs possibilita a identificação de padrões e práticas operacionais mais eficientes e

vantajosas, além de também permitir reconhecer melhorias e a formulação de estratégias para mitigar eventuais limitações dos modelos.

Para a realização desta RSL foi empregada a metodologia PRISMA 2020, que tem como finalidade apoiar o planejamento e condução de revisões sistemáticas [5]. Nesta RSL foi empregada a base de dados Scopus, a qual é amplamente reconhecida na comunidade científica, cujo acesso ocorreu no dia 21 de agosto de 2024. A seguir, são apresentadas as questões de pesquisa formuladas para direcionar este estudo:

Tabela 2: Questões de pesquisa.

| ID  | Questões de pesquisa                                                                                                      |
|-----|---------------------------------------------------------------------------------------------------------------------------|
| Q01 | Quais são os modelos utilizados?                                                                                          |
| Q02 | Quais os meios de aplicação (framework, protótipo, jogo já existente) utilizados?                                         |
| Q03 | Quais são os desafios frequentemente mencionados na utilização de LLMs em NPCs?                                           |
| Q04 | Quais são as táticas empregadas para superar as limitações e otimizar a eficácia nas interações?                          |
| Q05 | É mencionado a utilização de chamadas de funções ou variáveis para produzir diálogos melhores ou afetar o estado do jogo? |
| Q06 | Quais trabalhos utilizam e focam na coleta de dados anterior ou posterior ao diálogo/ação para aprimorar o NPC?           |

Após a formulação das questões, permitiu-se a construção da string de busca que partiu de uma extensão de outra já utilizada em uma revisão preliminar de escopo realizada no início de 2024, tal artigo tem como objetivo categorizar e identificar pesquisas que mencionam a utilização de LLM em jogos digitais. A comparação da string de busca pode ser visualizada na tabela a seguir:

Tabela 3. Adaptação de string de busca

| "Large Language Models and Video Games: A Preliminary Scoping Review" [14]                                                                                  | Adaptação                                                                                                                                                                                                                                                                      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ((llm OR llms OR (large AND language AND model) OR (large AND language AND models)) AND ((video AND game) OR (video AND games) OR videogame OR videogames)) | ((llm OR llms OR (large AND language AND model) OR (large AND language AND models)) AND ((video AND game) OR (video AND games) OR videogame OR videogames)) AND ( npc OR npcs OR (non-player AND character) OR (non-player AND characters) OR (non AND player AND characters)) |

A extensão na string ocorreu de forma a acrescentar NPCs, levando em consideração a palavra no singular, plural, com hífen e a forma extensa, sem acrônimos. Essa adaptação permitiu

identificar um número maior de estudos relevantes para o objetivo. Tal pesquisa foi feita no período de 2022 e 2024, que retornou 210 documentos. Para selecionar apenas trabalhos relevantes para este artigo, foi desenvolvido critérios de inclusão e exclusão, presentes na seguinte tabela:

Tabela 4: Critérios de inclusão e exclusão.

| Critérios de inclusão (CI)                                                                                          | Critérios de exclusão (CE)                           |
|---------------------------------------------------------------------------------------------------------------------|------------------------------------------------------|
| CI01: Trabalhos que abordem sobre a aplicação de LLMs em jogo, protótipo ou ferramenta para entretenimento digital. | CE01: Trabalhos com foco em educação.                |
| CI02: Estudos que objetivem aprimorar a interação do jogador com um NPC.                                            | CE02: Artigos cujo acesso implicava custos.          |
|                                                                                                                     | CE03: Estudos de revisão e pesquisas bibliográficas. |

A partir desses critérios foram excluídos 184 estudos desta revisão, essa primeira filtragem foi feita a partir da leitura do título e do resumo dos trabalhos. Os 26 restantes foram lidos por completo em setembro de 2024 e filtrados utilizando os CE e CI para que então se chegasse a 12 artigos finais. Segue um fluxograma representando esse processo:

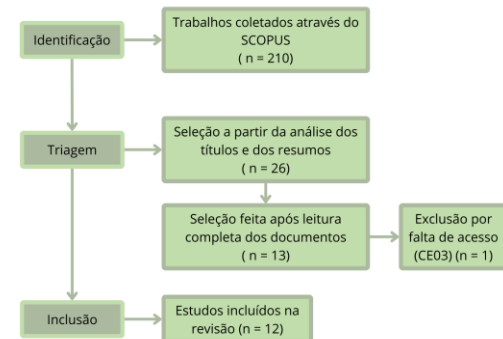


Figura 1: Fluxograma do PRISMA desenhando o processo de seleção dos artigos

Após finalizar essa seleção os 12 estudos inclusos estão presentes na tabela a seguir:

Tabela 5. Estudos incluídos

| ID | Título                                                                                                                     | Ano  | Autores                                                                       |
|----|----------------------------------------------------------------------------------------------------------------------------|------|-------------------------------------------------------------------------------|
| E1 | Personalized Quest and Dialogue Generation in Role-Playing Games: A Knowledge Graph- and Language Model-based Approach [6] | 2023 | Ashby, Trevor; Webb, Braden K.; Knapp, Gregory; Searle, Jackson; Fulda, Nancy |
| E2 | Prototyping Slice of Life: Social Physics with Symbolically Grounded LLM-based                                             | 2024 | Treanor, Mike; Samuel, Ben; Nelson, Mark J.                                   |

|     |                                                                                                                              |      |                                                                                                                                                 |
|-----|------------------------------------------------------------------------------------------------------------------------------|------|-------------------------------------------------------------------------------------------------------------------------------------------------|
|     | Generative Dialogue [7]                                                                                                      |      |                                                                                                                                                 |
| E3  | Craft an Iron Sword: Dynamically Generating Interactive Game Characters by Prompting Large Language Models Tuned on Code [8] | 2022 | Volum, Ryan; Rao, Sudha; Xu, Michael; DesGarennes, Gabriel; Brockett, Chris; Van Durme, Benjamin; Deng, Olivia; Malhotra, Akanksha; Dolan, Bill |
| E4  | FedNPC: Learning A Federated Framework for Large Language Models in Game NPCs [9]                                            | 2023 | Hong, Mengzea; Zhang, Kunb; Zhang, Shuning; He, Zhihangb;                                                                                       |
| E5  | The Turing Quest: Can Transformers Make Good NPCs? [10]                                                                      | 2023 | Gao, Qi Chen; Emami, Ali                                                                                                                        |
| E6  | Conversational Interactions with NPCs in LLM-Driven Gaming: Guidelines from a Content Analysis of Player Feedback [11]       | 2024 | Cox, Samuel Rhys; Ooi, Wei Tsang                                                                                                                |
| E7  | Simulating life: the application of generative agents in virtual environments [12]                                           | 2024 | Omigraliyev, Ruslan; Kenzhe, Damir; Mirambekov, Suienish                                                                                        |
| E8  | Balancing Game Satisfaction and Resource Efficiency: LLM and Pursuit Learning Automata for NPC Dialogues [13]                | 2024 | Jahangiri, Mohammed Mahdi; Rahmani, Parisa                                                                                                      |
| E9  | Investigating Agency of LLMs in Human-AI Collaboration Tasks [14]                                                            | 2024 | Sharma, Ashisha; Rao, Sudhab; Brockett, Chris; Malhotra, Akankshab; Jojic, Nebojsab; Dolan, Billb                                               |
| E10 | NarrativePlay: Interactive Narrative Understanding [15]                                                                      | 2023 | Zhao, Runcong; Zhang, Wenjia; Li, Jiazheng; Zhu, Lixing; Li, Yanran; He, Yulana; Gui, Lina                                                      |

|     |                                                                                                                |      |                                                                                |
|-----|----------------------------------------------------------------------------------------------------------------|------|--------------------------------------------------------------------------------|
| E11 | Bringing Stories to Life in 1001 Nights: A Co-creative Text Adventure Game Using a Story Generation Model [16] | 2022 | Sun, Yuqian; Ni, Xuran; Feng, Haozhen; Lc, Ray; Lee, Chang Hee; Asadipour, Ali |
| E12 | A Framework for Exploring Player Perceptions of LLM-Generated Dialogue in Commercial Video Games [17]          | 2023 | Akoury, Nader; Yang, Qian; Iyye, Mohit                                         |

E a seguir, a tabela 6 demonstra como esses estudos apontados respondem as questões da pesquisa mencionadas anteriormente:

**Tabela 6. Estudos com respostas das questões de pesquisa Q01, Q02, Q05 e Q06**

| Título                               | Modelos                        | Aplicação                        | Código | Coleta de dados |
|--------------------------------------|--------------------------------|----------------------------------|--------|-----------------|
| Personalized Quest... (E1)           | GPT-2                          | Framework                        | Não    | Sim             |
| Prototyping Slice of Life... (E2)    | Gemini                         | Protótipo de jogo                | Sim    | Não             |
| Craft an Iron Sword ... (E3)         | GPT-3                          | Jogo existente                   | Sim    | Sim             |
| FedNPC ... (E4)                      | -                              | Framework                        | Não    | Sim             |
| The Turing Quest... (E5)             | GPT-3                          | Pipeline narrativa               | Não    | Sim             |
| Conversational Interactions... (E6)  | -                              | Jogo existente                   | Não    | Não             |
| Simulating life... (E7)              | Llama-2                        | Protótipo de jogo                | Não    | Sim             |
| Balancing Game Satisfaction... (E8)  | -                              | Protótipo de jogo                | Não    | Não             |
| Investigating Agency of LLMs... (E9) | GPT-4, GPT-3, Llama-2, Guanaco | Framework                        | Não    | Sim             |
| NarrativePlay... (E10)               | GPT-3, GPT-4, Llama-2          | Protótipo de jogo (novel system) | Sim    | Sim             |
| Bringing Stories to Life ... (E11)   | GPT-2                          | Protótipo de jogo                | Não    | Sim             |
| A Framework for ... (E12)            | GPT-4                          | Framework                        | Sim    | Sim             |

**Tabela 7. Estudos com respostas das questões de pesquisa Q03 e Q04**

| Título                               | Desafios                                                                      | Táticas                                                                                     |
|--------------------------------------|-------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| Personalized Quest... (E1)           | Inconsistência                                                                | Fine-tuning, prompt                                                                         |
| Prototyping Slice of Life... (E2)    | Alucinação; Inconsistência; Imprevisibilidade                                 | Prompt, instruction-tuning                                                                  |
| Craft an Iron Sword ... (E3)         | Alucinação; Perca de contexto; Inconsistência; Repetição; Viés de recência    | Few-shot prompt, RAG, programação meta-prompt, multi-stage prompt                           |
| FedNPC ... (E4)                      | Custo de comunicação; Dados não distribuídos de forma idêntica entre modelos  | Modelo global compartilhado, ciclo fechado de treinamento, atualização automática do modelo |
| The Turing Quest... (E5)             | Dificuldade em gerar primeiras frases; Repetição; Janela de contexto          | Prompt, Feature Characterization Schema                                                     |
| Conversational Interactions... (E6)  | Alucinação, inconsistência, falta de memória                                  | -                                                                                           |
| Simulating life... (E7)              | -                                                                             | Fluxo de memória                                                                            |
| Balancing Game Satisfaction... (E8)  | -                                                                             | Autômatos de aprendizagem de busca (PLA)                                                    |
| Investigating Agency of LLMs... (E9) | -                                                                             | Agência, Prompt, fine-tuning, In-Context Learning (ICL)                                     |
| NarrativePlay... (E10)               | -                                                                             | Prompt, few-shot prompting, formatação JSON                                                 |
| Bringing Stories to Life ... (E11)   | Falta de detecção de palavras que não se encaixam com o contexto da narrativa | Prompt, few-shot prompting                                                                  |
| A Framework for ... (E12)            | Inconsistência, menos adequação ao fluxo da narrativa                         | Fine-tuning, Few-shot prompting                                                             |

### 3 RESULTADOS

Os resultados dessa revisão, conforme apresentados na tabela 6, indicam que os modelos mais utilizados nos estudos pertencem a família GPT (GPT-2, GPT-3, GPT-4). Além desses, outros são mencionados, como o Gemini, Llama 2 e Guanaco, sendo esse último, mencionado no artigo E9, uma adaptação do Llama.

Esses modelos foram aplicados em um contexto de jogo através de frameworks (E1, E4, E9, E12), protótipos (E2, E7, E8, E10, E11), aplicação em jogos comerciais (E3, E6) e criação de pipeline narrativa (E5). Os estudos dos frameworks (E1, E4, E9), da pipeline narrativa (E5) e da técnica aplicada em protótipo (E8) apresentam abordagens que podem ser aplicados em qualquer gênero de jogo, sugerindo uma flexibilidade nas implementações.

Os quatro trabalhos que propõem a construção de um framework apresentam objetivos específicos e distintos. O estudo E1 investiga a ampliação da capacidade dos NPCs por meio da geração de missões, enquanto os demais concentram-se na otimização do desempenho do modelo. Nesse sentido, o E4 sugere a criação de um modelo global compartilhado, o E9 analisa a implementação de agência em modelos, que inclui a capacidade de influenciar eventos, enquanto o E12 aborda uma avaliação dinâmica para a gestão de sistemas de diálogos que orientam interações baseadas em tarefas.

Dentre os protótipos, os estudos E2 e E7 apresentam uma característica em comum: a ênfase na interação entre NPCs e na simulação de relações sociais. Ambos desenvolvem protótipos com o intuito de investigar esse fenômeno, contribuindo para a discussão sobre o potencial dos modelos de linguagem em simular interações complexas em ambientes de jogo.

Já quanto aos outros estudos com protótipo, E8 realiza uma integração narrativa de um LLM com Pursuit Learning Automata (PLA), uma técnica de aprendizado por reforço, o E10 cria um sistema de narrativa sincronizando ferramentas e o E11 projeta um jogo em que o diálogo gerado pelo NPC afeta outras mecânicas e, consequentemente, se torna uma condição de vitória.

Dentre os problemas mais recorrentes observados, estão a alucinação dos modelos, inconsistência nas respostas e limitações relacionadas à janela de contexto. Diversas pesquisas investigam a manipulação de prompts como uma estratégia para otimizar a interação dos NPCs e resolver esses eventuais problemas. Trabalhos como (E2, E3 E5, E10) utilizam prompts que incluem o estado atual do jogo, sendo que o E3 faz o uso da técnica de RAG (Retrieval-Augmented Generation) para acessar essas informações e como solução para o viés de recência.

Apesar das inovações, poucos trabalhos exploram a manipulação do chat por meio de variáveis e funções. Em contrapartida, observou-se que a maioria dos estudos utilizou alguma forma de base de dados para treinamento dos modelos.

### 4 DISCUSSÃO

Os resultados reunidos nesta revisão revelam uma preocupação crescente com a qualidade da interação, em detrimento da exploração de outras funcionalidades que poderiam ser desenvolvidas a partir dos outputs gerados. Isso não é surpreendente, considerando que esses modelos são ferramentas imprevisíveis e resultados indesejáveis podem impactar negativamente a experiência do jogador.

Com isso algumas estratégias foram desenvolvidas a partir dessa necessidade de controle sobre os outputs gerados, minimizando possíveis impactos. No trabalho E2, os modelos são

treinados com instruções e os prompts enviados são compostos pelo estado do jogo e solicitações imperativas. Já o estudo E8, que faz uso do PLA, relata que essa abordagem resultou em feedbacks positivos dos jogadores. No E4, o cliente recebe instâncias do modelo global e as interações realizadas geram dados que são retransmitidos ao servidor, permitindo a otimização contínua do modelo.

Em relação ao controle de memória e contexto, o E6 propõe a sumarização de diálogos nos prompts, enquanto o artigo E10 utiliza formatação JSON para tornar o prompt mais conciso e fornecer um contexto mais claro ao modelo. Similarmente, o E5 utiliza esquemas que incluem dados sobre o mundo do jogo e personagens, incorporando-os aos prompts. Já para lidar com o problema de repetição, o E5 implementa uma penalidade de frequência dinâmica, que aumenta progressivamente a cada repetição.

O estudo E3 menciona o uso da técnica RAG, que se mostrou eficaz na recuperação de informações relevantes de uma base de dados, auxiliando na contextualização do modelo. Por outro lado, o artigo E9 explora uma questão distinta: a ausência de agência dos modelos, isso se refere à capacidade de moldar proativamente eventos, algo inerente às interações humanas. Os autores sugerem que a agência pode ser incorporada por meio de Instruções por Contexto de Linguagem (ICL), utilizando exemplos específicos.

Observa-se que os projetos que utilizaram os modelos Llama-2 e Gemini foram realizados entre 2023 e 2024, evidenciando que a introdução de novas arquiteturas de modelos de linguagem está reduzindo a predominância daqueles da família GPT.

Os modelos Gemini destacam-se pela diversidade e precisão, enquanto os modelos ChatGPT apresentam maior unicidade e criatividade, o que resulta em diálogos com maior fluidez e imaginação. [18] [19]

Um modelo deve ser avaliado de acordo com a tarefa que ele vai executar [20]. No âmbito dos jogos, algumas métricas são mais críticas, como a duplicidade e a alucinação nas respostas, que podem impactar negativamente na experiência do jogador.

Já os modelos Llama têm a vantagem de ser código aberto, o que promove flexibilidade e transparência para desenvolvedores. Embora, isso possa ser visto como uma vulnerabilidade quanto a introdução de vieses, o que exige maior cuidado sobre os dados em fine-tunings [21]. Porém, o Llama não desempenha melhor do que o modelo ChatGPT-4, ele é superado em questões de coerência, relevância e legibilidade. [22]

Um ponto crítico identificado nos estudos é a limitação na manipulação de saídas e prompts por meio de código, evidenciando uma lacuna considerável na exploração das interações com NPCs. Um exemplo que realiza essa abordagem de forma eficaz está presente no artigo E11, onde o NPC, que inicialmente narra uma história, ao mencionar armas, estas podem ser integradas nas mecânicas de combate do jogo. Adicionalmente, o estudo emprega prompts que incorporam as últimas cinco mensagens, técnica que ajuda a manter o contexto nas interações.

Por fim, uma observação interessante é feita no artigo E12, onde os participantes da pesquisa manifestaram preferência por diálogos elaborados pelo game designer em vez daqueles gerados

pelo modelo GPT-4, pois esse último apresenta incoerências em relação à lógica do fluxo e ao estado do jogo. No entanto, vale ressaltar que o framework concebido no estudo não tem como objetivo a geração completa de diálogos pelos modelos, mas sim o reconhecimento do estado do jogo através do input do jogador.

## 5 CONCLUSÃO

Com base nos 12 estudos selecionados, esta revisão sistemática respondeu a todas as questões formuladas na Tabela 2. As táticas e estratégias extraídas estão detalhadas nos resultados e na discussão deste artigo e permitem inferir importantes conclusões sobre o estado atual dos avanços na implementação de LLMs em NPCs.

A gratuidade dos novos modelos contribui com a democratização do acesso a essa tecnologia, mesmo que essas versões possam apresentar limitações. Com a maior acessibilidade há um aumento potencial no volume de pesquisas, incluindo as que exploram o desenvolvimento de NPCs utilizando LLMs. Esse fenômeno será ainda mais favorecido à medida que esses modelos apresentarem uma redução de erros, proporcionando uma sensação de maior confiabilidade para os desenvolvedores.

No entanto, é importante destacar que as pesquisas envolvendo LLMs ainda enfrentam uma limitação significativa devido ao alto custo computacional. Espera-se que, no futuro, com o avanço nas tecnologias de hardware e otimizações de algoritmos, esse custo seja reduzido. Tal progresso possibilitará mais investigações das interações entre NPCs e jogadores, superando as restrições impostas pelos atuais requisitos de infraestrutura.

Uma abordagem promissora é o uso de manipulação de código para gerar interações complexas, ou seja, que impactem diretamente o estado do jogo e para auxiliar no gerenciamento de memória e contexto, facilitando assim a criação de um prompt informativamente rico.

Por fim, estudos futuros também devem se concentrar na otimização dos modelos para garantir que NPCs interativos possam ser integrados em jogos de forma eficiente, sem comprometer a jogabilidade. Além disso, é essencial estar atento a novos lançamentos e técnicas utilizadas por grandes empresas como OpenAI<sup>1</sup>, Google<sup>2</sup> e Meta<sup>3</sup>, que estão trabalhando ativamente em soluções para muitos dos problemas mencionados, como eficiência computacional e gerenciamento de contexto.

## 6 CONSIDERAÇÕES ÉTICAS

Ao utilizar LLMs em jogos, é importante informar os jogadores sobre a coleta de dados durante o diálogo e alertá-los para não compartilharem informações pessoais ou sigilosas. Além disso, a utilização excessiva de LLMs na produção de um jogo pode comprometer a originalidade e intencionalidade artística, afetando a narrativa da obra.

<sup>1</sup> <https://openai.com/>

<sup>2</sup> <https://developers.google.com/>

<sup>3</sup> <https://developers.meta.com/horizon>

## REFERÊNCIAS

- [1] Ozdemir, S. Quick start guide to large language models: Strategies and best practices for using ChatGPT and other LLMs. Boston, MA, USA: Addison Wesley, 2023. E-book (não paginado). ISBN-13: 978-0138199197
- [2] Sweetser, Penny. (2024). Large languages models and video games: A preliminary scoping review. Proc. ACM Convers. User Interf. 2024. <https://doi.org/10.1145/3640794.3665582>
- [3] Schell, J. The art of game design: A book of lenses, third edition. 3. ed. Londres, England: Routledge, 2019. ISBN-13: 978-1138632059
- [4] Costa, A., Fontanari, A. e Zoltowski A. (2022). Como escrever um artigo de revisão sistemática: um guia atualizado. In M. I. C. Sampaio, A. A. Z. P. Sabadini, & S. H. Koller (Orgs.), Produção científica: um guia prático (1ª ed., pp. 130-165). DOI: <https://doi.org/10.11606/9786587596280>
- [5] Page, M.J., McKenzie, J.E., Bossuyt, P.M. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 10, 89 (2021). DOI: <https://doi.org/10.1186/s13643-021-01626-4>
- [6] Trevor Ashby, Braden K Webb, Gregory Knapp, Jackson Searle, e Nancy Fulda. 2023. 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 (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 290, 1–20. DOI: <https://doi.org/10.1145/3544548.3581441>
- [7] Mike Treanor, Ben Samuel, e Mark J. Nelson. 2024. Prototyping Slice of Life: Social Physics with Symbolically Grounded LLM-based Generative Dialogue. In Proceedings of the 19th International Conference on the Foundations of Digital Games (FDG '24). Association for Computing Machinery, New York, NY, USA, Article 56, 1–4. DOI: <https://doi.org/10.1145/3649921.3656988>
- [8] Ryan Volum, Sudha Rao, Michael Xu, Gabriel DesGarennes, Chris Brockett, Benjamin Van Durme, Olivia Deng, Akanksha Malhotra, e Bill Dolan. 2022. Craft an iron sword: Dynamically generating interactive game characters by prompting large language models tuned on code. Em Proceedings of the 3rd Wordplay: When Language Meets Games Workshop (Wordplay 2022), 2022. Association for Computational Linguistics, Stroudsburg, PA, USA. DOI: <https://doi.org/10.18653/v1/2022.wordplay-1.3>
- [9] Mengze Hong, Kun Zhang, Shuning Zhang, e Zhihang He. 2023. FedNPC: A federated learning framework for large language models in game NPCs. Em 2023 IEEE 21st Student Conference on Research and Development (SCORED), 2023. IEEE. DOI: <https://doi.org/10.1109/SCORED60679.2023.10563705>
- [10] Qi Chen Gao e Ali Emami. 2023. The Turing quest: Can transformers make good NPCs? Em Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop), 2023. Association for Computational Linguistics, Stroudsburg, PA, USA. DOI: <https://doi.org/10.18653/v1/2023.acl-srw.17>
- [11] Samuel Rhys Cox and Wei Tsang Ooi. 2024. Conversational Interactions with NPCs in LLM-Driven Gaming: Guidelines from a Content Analysis of Player Feedback. In Chatbot Research and Design: 7th International Workshop, CONVERSATIONS 2023, Oslo, Norway, November 22–23, 2023, Revised Selected Papers. Springer-Verlag, Berlin, Heidelberg, 167–184. DOI: [https://doi.org/10.1007/978-3-031-54975-5\\_10](https://doi.org/10.1007/978-3-031-54975-5_10)
- [12] R. Omirgaliyev, D. Kenzhe and S. Mirambekov, "Simulating Life: The Application of Generative Agents in Virtual Environments," 2024 IEEE AITU: Digital Generation, Astana, Kazakhstan, 2024, pp. 181-187, DOI: <https://doi.org/10.1109/IEEECONF61558.2024.10585387>.
- [13] Mohammed Mahdi Jahangiri e Parisa Rahmani. 2024. Balancing game satisfaction and resource efficiency: LLM and pursuit learning automata for NPC dialogues. Em 2024 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2024. IEEE DOI: <https://doi.org/10.1109/HORA61326.2024.10550450>
- [14] Ashish Sharma, Sudha Rao, Chris Brockett, Akanksha Malhotra, Nebojsa Jovic, e Bill Dolan. 2023. Investigating Agency of LLMs in Human-AI Collaboration Tasks. (Maio 2023). DOI: <https://doi.org/10.48550/arXiv.2305.12815>
- [15] Runcong Zhao, Wenjia Zhang, Jiazheng Li, Lixing Zhu, Yanran Li, Yulan He, e Lin Gui. 2023. NarrativePlay: Interactive Narrative Understanding. arXiv [cs.CL]. DOI: <https://doi.org/10.48550/arXiv.2310.01459>
- [16] Yuqian Sun, Xuran Ni, Haozhen Feng, Ray LC, Chang Hee Lee, and Ali Asadipour. 2022. Bringing Stories to Life in 1001 Nights: A Co-creative Text Adventure Game Using a Story Generation Model. In Interactive Storytelling: 15th International Conference on Interactive Digital Storytelling, ICIDS 2022, Santa Cruz, CA, USA, December 4–7, 2022, Proceedings. Springer-Verlag, Berlin, Heidelberg, 651–672. DOI: [https://doi.org/10.1007/978-3-031-22298-6\\_42](https://doi.org/10.1007/978-3-031-22298-6_42)Conference Name:ACM Woodstock conference
- [17] Nader Akoury, Qian Yang, e Mohit Iyyer. 2023. A Framework for Exploring Player Perceptions of LLM-Generated Dialogue in Commercial Video Games. In Findings of the Association for Computational Linguistics: EMNLP 2023, pages 2295–2311, Singapore. Association for Computational Linguistics. DOI: <https://doi.org/10.18653/v1/2023.findings-emnlp.151>
- [18] Alsajri, A. et al.trans. 2024. Generative Models in Natural Language Processing: A Comparative Study of ChatGPT and Gemini. *Babylonian Journal of Artificial Intelligence*. 2024, (Nov. 2024), 134–145. DOI: <https://doi.org/10.58496/BJAI/2024/015>.
- [19] Rane, N., Choudhary, S. and Rane, J. 2024. Gemini versus ChatGPT: applications, performance, architecture, capabilities, and implementation. *Journal of Applied Artificial Intelligence*. 5, 1 (Mar. 2024), 69–93. DOI: <https://doi.org/10.48185/jaai.v5i1.1052>
- [20] Minaee, S. et al. 2024. Large Language Models: a survey. arXiv (Cornell University). (Feb. 2024). DOI: <https://doi.org/10.48550/arxiv.2402.06196>.
- [21] Ragav Gupta, 2025. Comparative Analysis of DeepSeek R1, ChatGPT, Gemini, Alibaba, and LLaMA: Performance, Reasoning Capabilities, and Political Bias. *Authorea*. February 10, 2025. DOI: <https://doi.org/10.22541/au.173921625.50315230/v1>
- [22] S. R. Bogireddy and N. Dasari. 2024. Comparative Analysis of ChatGPT-4 and LLaMA: Performance Evaluation on Text Summarization, Data Analysis, and Question Answering, 2024 *15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, pp. 1-7, DOI: 10.1109/ICCCNT61001.2024.10725662.