

Uma Abordagem com Multiagentes Baseados em LLMs para o Design de Jogos Educativos Endógenos

*An Approach Using LLM-Based Multi-Agents for
Endogenous Educational Game Design*

Caio Silva Azeredo¹, Geraldo Bonorino Xexéo¹

¹Programa de Engenharia de Sistemas e Computação – COPPE/UFRJ
Rio de Janeiro – RJ – Brasil

{caio.azeredo, xexeo}@cos.ufrj.br

Abstract. Introduction: The development of endogenous educational games requires specialized knowledge in both game design and pedagogy, creating barriers that limit the adoption of these methodologies. **Objective:** This work aims to develop and validate a multi-agent system based on Large Language Models (LLMs) to optimize the design process of endogenous educational games through assisted completion of the Endo-GDC (Game Design Canvas for Endogenous Educational Games). **Method:** Following Design Science Research methodology, the system was developed integrating four specialized modules: AI-assisted Brainstorming, Socratic Agent for conceptual refinement, assisted Endo-GDC completion with specialized agents (Coordinator, Mechanics, Narrative, and Engagement), and educational objectives classification based on Bloom's Revised Taxonomy. Validation was conducted through 26 case studies assessed via TAM, TTF, AX, and ADDIE frameworks. **Results:** The system evidenced high technological acceptance, with perceived usefulness of 6.23 on a 7-point scale, task-technology fit of 5.95, satisfactory algorithmic experience of 5.80, and processual effectiveness of 5.89, confirming system efficacy in reducing development time while maintaining pedagogical quality.

Keywords: LLM, Multi-Agent Systems, Educational Game Design, Bloom's Taxonomy, Design Science Research.

Resumo. Introdução: O desenvolvimento de jogos educacionais endógenos requer conhecimento especializado tanto em design de jogos quanto em pedagogia, criando barreiras que limitam a adoção destas metodologias. **Objetivo:** Este trabalho tem como objetivo desenvolver e validar um sistema multiagente baseado em Grandes Modelos de Linguagem (LLMs) para otimizar o processo de design de jogos educacionais endógenos, por meio do preenchimento assistido do Endo-GDC (Game Design Canvas para Jogos Educativos Endógenos). **Método:** Seguindo a metodologia Design Science Research, o sistema foi desenvolvido integrando quatro módulos especializados: Brainstorming assistido por IA, Agente Socrático para refinamento conceitual, preenchimento assistido do Endo-GDC com agentes especializados (Coordenador, Mecânicas, Narrativa e Engajamento) e classificação de objetivos educacionais baseada na Taxonomia de Bloom Revisada. A validação foi conduzida através de 26 estudos de caso avaliados

pelos frameworks TAM, TTF, AX e ADDIE. Resultados: O sistema evidenciou alta aceitação tecnológica, com utilidade percebida de 6,23 em escala de 7 pontos, adequação tarefa-tecnologia de 5,95, experiência algorítmica satisfatória de 5,80 e efetividade processual de 5,89, confirmando a eficácia do sistema em reduzir tempo de desenvolvimento mantendo qualidade pedagógica.

Palavras-chave: LLM, Sistemas Multiagentes, Design de Jogos Educativos, Taxonomia de Bloom, Design Science Research.

1. Introdução

O desenvolvimento de jogos educacionais enfrenta desafios relacionados à complexidade na integração de elementos lúdicos com objetivos pedagógicos e ao tempo necessário para criação de experiências de aprendizagem [Prensky 2001]. Designers educacionais necessitam conhecimento especializado tanto em design de jogos quanto em pedagogia, além de enfrentarem dificuldades na classificação e definição de objetivos educacionais adequados [Taucei 2019].

Jogos educacionais endógenos, nos quais o conteúdo educacional é integrado às mecânicas centrais ao invés de ser adicionado como camada superficial [Taucei 2019], representam uma abordagem promissora para resolver a tensão entre objetivos pedagógicos e engajamento lúdico. Contudo, o design destes jogos é particularmente desafiador, pois requer identificação criativa de formas autênticas de integração entre domínio educacional e mecânicas de jogo. O Endo-GDC constitui uma ferramenta estruturada para este fim, porém seu preenchimento manual demanda conhecimento especializado e tempo considerável [Taucei 2019].

Os Grandes Modelos de Linguagem (LLMs) apresentam capacidades de processamento de linguagem natural que podem ser aplicadas em sistemas multiagentes para automação de processos complexos [Zhang et al. 2024]. A integração destas tecnologias no contexto educacional pode contribuir para redução do tempo de desenvolvimento mantendo qualidade pedagógica.

O objetivo principal desta pesquisa é **desenvolver e validar uma abordagem que incorpore técnicas de Inteligência Artificial para otimizar e automatizar o processo de design de jogos educacionais**, focando no preenchimento assistido do Endo-GDC. Especificamente, a pesquisa busca: (i) investigar métodos de assistência inteligente para geração de ideias educacionais; (ii) desenvolver abordagens para refinamento conceitual via método socrático automatizado; (iii) examinar estratégias de automação para classificação de objetivos educacionais alinhados à Taxonomia de Bloom Revisada; e (iv) analisar abordagens multiagentes para suporte ao design instrucional, avaliando sua aceitação e eficácia.

2. Trabalhos Relacionados

A criação de jogos educacionais envolve a integração de múltiplos elementos de design e pedagogia [Kalmpourtzis 2018], demandando metodologias específicas para garantir efetividade pedagógica [Boller e Kapp 2017]. A Taxonomia de Bloom, em sua versão revisada [Anderson e Krathwohl 2001], fornece um framework estruturado

para classificação de objetivos educacionais em seis níveis cognitivos. Churches [Churches 2008] estabeleceu conexões entre ferramentas digitais e estes níveis cognitivos.

No âmbito de sistemas multiagentes baseados em LLM, agentes percebem e interagem através de mensagens ambientais, de interação e de auto-reflexão [Zhang et al. 2023]. A integração de memória e conhecimento externo permite que agentes mantenham contexto histórico e acessem informações especializadas [Park et al. 2023, Zhao et al. 2024]. Frameworks como AutoAgents [Chen et al. 2024], ChatDev [Qian et al. 2023] e MetaGPT [Hong et al. 2023] demonstram a viabilidade de sistemas multiagentes baseados em LLMs para execução de tarefas complexas. A comunicação entre agentes segue modelos baseados em intenções e compromentimentos [Cohen e Levesque 1990].

3. Metodologia

Esta pesquisa adota o Design Science Research (DSR) como abordagem metodológica, seguindo as etapas propostas por Dresch et al. [Dresch et al. 2020]. O DSR constitui uma abordagem orientada à concepção de artefatos que resolvam problemas relevantes, contribuindo simultaneamente para o avanço do conhecimento científico. A escolha do DSR deve-se à sua característica de conectar a relevância prática com o rigor científico, sendo particularmente adequada para pesquisas em computação aplicada à educação.

O processo seguiu cinco etapas: (1) **Conscientização do problema**, envolvendo a identificação das dificuldades na integração de elementos lúdicos com objetivos pedagógicos, das limitações temporais e da escassez de ferramentas especializadas; (2) **Proposição do artefato**, definindo quatro módulos funcionais: Brainstorming, Agente Socrático, Endo-GDC e Taxonomia de Bloom; (3) **Desenvolvimento**, seguindo abordagem iterativa e incremental com ciclos de implementação, teste e refinamento; (4) **Avaliação**, realizada através de 26 estudos de caso em diferentes contextos educacionais, utilizando quatro frameworks estabelecidos: Technology Acceptance Model – TAM [Davis 1989], Task-Technology Fit – TTF [Goodhue e Thompson 1995], Algorithmic Experience – AX [Shin 2021] e o modelo ADDIE [Branch 2009]; e (5) **Comunicação dos resultados**.

As conjecturas teóricas que orientaram o desenvolvimento foram: (C1) múltiplos agentes especializados otimizam limitações de sistemas monolíticos no design educacional; (C2) assistência por IA reduz tempo de design mantendo qualidade pedagógica; (C3) processo assistido auxilia na obtenção de fundamentação pedagógica sólida; e (C4) método socrático automatizado conduz a designs pedagogicamente mais robustos.

4. Sistema Proposto

A arquitetura do sistema foi concebida para suportar múltiplos agentes especializados categorizados em quatro tipos: Coordenador, Especialista em Mecânicas, Especialista em Narrativa e Especialista em Engajamento. O sistema é organizado em quatro módulos funcionais, cujas interfaces são apresentadas na Figura 1.

O **Módulo de Brainstorming** (Figura 1a) facilita sessões assistidas por IA para geração de ideias iniciais. O usuário define parâmetros do projeto e insere ideias

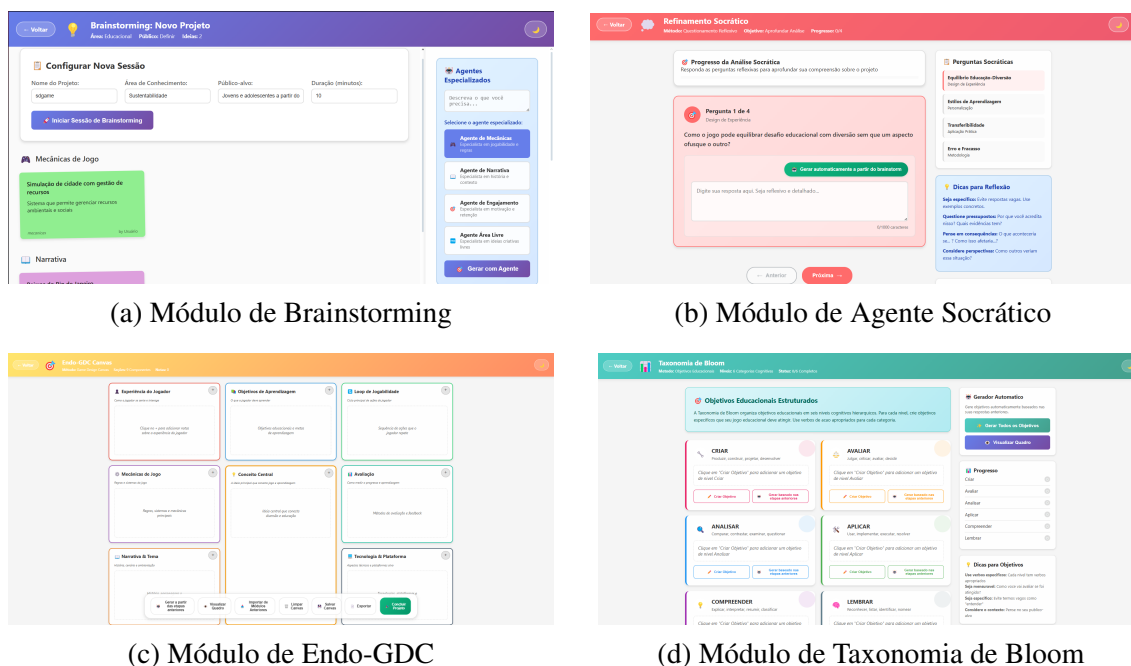


Figura 1. Interfaces dos quatro módulos do sistema: (a) Brainstorming com agentes especializados; (b) Refinamento sócrático com questionamento reflexivo; (c) Endo-GDC Canvas com nove seções; (d) Taxonomia de Bloom com seis níveis cognitivos.

organizadas por categoria. O painel lateral permite selecionar agentes especializados (Mecânicas, Narrativa, Engajamento ou Área Livre) para geração de sugestões direcionadas. Nos 26 estudos, o número médio de ideias iniciais foi 3,4, enquanto o sistema gerou em média 5,8 sugestões complementares (expansão de 70%).

O **Módulo de Agente Sócrático** (Figura 1b) aplica o método sócrático para refinar ideias e identificar problemas. A interface apresenta questões reflexivas organizadas por categoria (Equilíbrio Educação-Diversão, Estilos de Aprendizagem, Transferibilidade, Erro e Fracasso), com geração automática a partir do brainstorm. Em 89% dos casos, o processo resultou em refinamentos significativos das ideias iniciais.

O **Módulo de Endo-GDC** (Figura 1c) oferece um canvas interativo com nove seções: Experiência do Jogador, Objetivos de Aprendizagem, Loop de Jogabilidade, Mecânicas de Jogo, Conceito Central, Avaliação, Narrativa & Tema e Tecnologia & Plataforma. O Agente Coordenador orquestra as contribuições dos demais agentes, com 92% das sugestões consideradas relevantes.

O **Módulo de Taxonomia de Bloom** (Figura 1d) auxilia na formulação de objetivos educacionais segundo os seis níveis da Taxonomia Revisada [Anderson e Krathwohl 2001]. A interface exhibe cards para cada nível com verbos de ação apropriados, permitindo criação manual ou geração automática. A distribuição média nos 26 estudos foi: Lembrar (18%), Entender (22%), Aplicar (19%), Analisar (16%), Avaliar (13%) e Criar (12%).

A implementação utilizou ASP.NET Core no backend, React no frontend, Entity Framework Core com SQL Server, e integração com APIs de LLM (OpenAI e Groq).

5. Resultados

A Tabela 1 apresenta os resultados da avaliação com os quatro frameworks, operacionalizada em escala Likert de 7 pontos, aplicada aos 26 estudos de caso em contextos que abrangem desde ensino fundamental até formação corporativa.

Tabela 1. Resultados dos frameworks de avaliação (escala Likert 1–7)

Constructo	Média	DP
<i>Technology Acceptance Model (TAM)</i>		
Utilidade Percebida	6,23	0,41
Facilidade de Uso	5,84	0,35
Intenção Comportamental	5,90	0,25
<i>Task-Technology Fit (TTF)</i>		
Adequação Tarefa-Tecnologia	5,95	0,12
Impacto na Performance	5,89	0,22
<i>Algorithmic Experience (AX)</i>		
Transparência Algorítmica	5,58	0,31
Confiança Algorítmica	5,81	0,22
Satisfação Geral AX	5,80	0,20
<i>ADDIE Process Evaluation</i>		
Eficiência do Desenvolvimento	5,92	0,21
Satisfação Geral do Processo	5,89	0,17

Os resultados do TAM indicam aceitação substancial, com a Utilidade Percebida apresentando a maior média ($M=6,23$) e forte correlação com Intenção Comportamental ($r = 0,78, p < 0,001$), confirmando a conjectura C2. A análise TTF revelou alinhamento satisfatório ($M=5,95$, $DP=0,12$), com baixa variabilidade indicando consenso entre participantes. A correlação entre TTF e TAM ($r = 0,742, p < 0,001$) confirma complementaridade entre adequação técnica e aceitação comportamental. A Transparência Algorítmica ($M=5,58$) indicou área para melhorias, enquanto a correlação entre Controlabilidade e Confiança ($r = 0,67$) sugere que maior controle percebido fortalece a confiança algorítmica. A avaliação ADDIE demonstrou efetividade consistente, corroborando as conjecturas C3 e C4. A consistência interna foi satisfatória em todos os frameworks: TAM ($\alpha = 0,87$), TTF ($\alpha = 0,83$), AX ($\alpha = 0,81$) e ADDIE ($\alpha = 0,85$).

6. Conclusão

Esta pesquisa desenvolveu e validou, via Design Science Research, um sistema multiagente baseado em LLMs para o design de jogos educacionais endógenos. A validação com 26 estudos de caso confirmou as quatro conjecturas teóricas propostas, evidenciando que a abordagem multiagente é eficaz para reduzir tempo de desenvolvimento mantendo qualidade pedagógica. As limitações incluem dependência de APIs externas de LLM e necessidade de melhorias na transparência algorítmica. A pesquisa contribui ao estabelecer uma ponte entre sistemas multiagentes e design educacional, oferecendo uma ferramenta que democratiza o acesso ao design de jogos educacionais endógenos. Como trabalhos futuros, pretende-se explorar arquiteturas multiagentes mais sofisticadas e integração com ferramentas de prototipagem rápida.

Referências

- Anderson, L. W. e Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition*. Addison Wesley Longman, Inc.
- Boller, S. e Kapp, K. (2017). *Play to learn: Everything you need to know about designing effective learning games*. Association for talent development.
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*, volume 722. Springer Science & Business Media.
- Chen, G., Dong, S., Shu, Y., Zhang, G., Sesay, J., Karlsson, B. F., Fu, J., e Shi, Y. (2024). Autoagents: A framework for automatic agent generation. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pages 22–30.
- Churches, A. (2008). Bloom's digital taxonomy. *Tech & Learning*.
- Cohen, P. R. e Levesque, H. J. (1990). Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340.
- Dresch, A., Lacerda, D. P., e Junior, J. A. V. A. (2020). *Design science research: método de pesquisa para avanço da ciência e tecnologia*. Bookman Editora.
- Goodhue, D. L. e Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS quarterly*, pages 213–236.
- Hong, S., Zheng, X., Chen, J., Cheng, Y., Wang, J., Zhang, C., Wang, Z., Yau, S. K. S., Lin, Z., Zhou, L., et al. (2023). Metagpt: Meta programming for multi-agent collaborative framework. *arXiv preprint arXiv:2308.00352*, 3(4):6.
- Kalmpourtzis, G. (2018). *Educational Game Design Fundamentals: A Journey to Creating Intrinsically Motivating Learning Experiences*. CRC Press, Boca Raton.
- Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., e Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pages 1–22.
- Prensky, M. (2001). *Digital Game-Based Learning*. McGraw-Hill, New York.
- Qian, C., Liu, W., Liu, H., Chen, N., Dang, Y., Li, J., Yang, C., Chen, W., Su, Y., Cong, X., et al. (2023). Chatdev: Communicative agents for software development. *arXiv preprint arXiv:2307.07924*.
- Shin, D. (2021). The algorithmic transparency paradox: Explaining and justifying algorithmic decision-making. *AI & Society*, 36(4):1185–1200.
- Taucei, B. (2019). Endo-gdc: Desenvolvimento de um game design canvas para concepção de jogos educativos endógenos. Dissertação de mestrado, Programa de Engenharia de Sistemas e Computação, COPPE/UFRJ, Rio de Janeiro, Brasil.
- Zhang, H., Du, W., Shan, J., Zhou, Q., Du, Y., Tenenbaum, J. B., Shu, T., e Gan, C. (2023). Building cooperative embodied agents modularly with large language models. *arXiv preprint arXiv:2307.02485*.

- Zhang, Y., Yang, S., Bai, C., Wu, F., Li, X., Wang, Z., e Li, X. (2024). Towards efficient llm grounding for embodied multi-agent collaboration. *arXiv preprint arXiv:2405.14314*.
- Zhao, A., Huang, D., Xu, Q., Lin, M., Liu, Y. J., e Huang, G. (2024). Expel: Llm agents are experiential learners. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(2024):19632–19642.