



UMA ABORDAGEM COM MULTIAGENTES BASEADOS EM LLMS PARA O DESIGN DE JOGOS EDUCATIVOS

Caio Silva Azeredo

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Sistemas e Computação.

Orientador: Geraldo Bonorino Xexéo

Rio de Janeiro
Junho de 2025

UMA ABORDAGEM COM MULTIAGENTES BASEADOS EM LLMS PARA O
DESIGN DE JOGOS EDUCATIVOS

Caio Silva Azeredo

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE
MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Orientador: Geraldo Bonorino Xexéo

Aprovada por: Prof. Daniel Serrão Schneider

Prof. Rosa Maria Esteves Moreira da Costa

RIO DE JANEIRO, RJ – BRASIL
JUNHO DE 2025

Silva Azeredo, Caio

Uma abordagem com multiagentes baseados em LLMs para o Design de Jogos Educativos/Caio Silva Azeredo. – Rio de Janeiro: UFRJ/COPPE, 2025.

XVII, 242 p.: il.; 29, 7cm.

Orientador: Geraldo Bonorino Xexéo

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2025.

Referências Bibliográficas: p. 134 – 164.

1. LLM. 2. multiagentes. 3. Design de Jogos. I. Bonorino Xexéo, Geraldo. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

*Aos meus avós: Fernando,
Risalva, Levy e Ana Maria.
Minhas primeiras lembranças
jogando e criando jogos são com
vocês*

Agradecimentos

Gostaria de primeiramente agradecer a Deus.

À minha esposa, Ana Caroline Gonçalves da Silva Azeredo, por todo seu amor e por nunca ter deixado de acreditar em mim, te amo.

Ao meu pai, Marcos Jesse Rios Azeredo, pela inspiração e exemplo que sempre foi para mim.

À minha mãe, Valeria Azevedo Silva Azeredo pela criação e educação que me deu, a qual tentarei resumir na frase que sempre me disse: "Primeiro a obrigação e depois a diversão".

Aos meus sogros Wilson dos Santos Pinto da Silva e Vera Lucia Gonçalves por todo o suporte que me deram.

Ao professor Geraldo Bonorino Xexéo pela paciência e pelos seus ensinamentos que foram muito mais que acadêmicos, foram ensinamentos para vida.

Ao professor Marcus Vinícius Coutinho Parreiras por todo incentivo e motivação

Ao professores da COPPE/UFRJ por compartilharam seus conhecimentos.

Ao CNPQ e a CAPES por terem financiando meu mestrado.

À Coppetec e RNP pelas bolsas de pesquisa fornecidas ao longo do mestrado.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

UMA ABORDAGEM COM MULTIAGENTES BASEADOS EM LLMS PARA O DESIGN DE JOGOS EDUCATIVOS

Caio Silva Azeredo

Junho/2025

Orientador: Geraldo Bonorino Xexéo

Programa: Engenharia de Sistemas e Computação

Este estudo desenvolveu um sistema multiagente baseado em Grandes Modelos de Linguagem (LLMs) para otimizar o processo de design de jogos educacionais endógenos. A pesquisa foi conduzida por meio da metodologia Design Science Research, seguindo as etapas de conscientização do problema, proposição do artefato, desenvolvimento, avaliação e comunicação dos resultados. O sistema integra quatro módulos funcionais especializados: Brainstorming assistido por IA, Agente Socrático para refinamento conceitual, preenchimento assistido do Endo-GDC (Game Design Canvas para Jogos Educativos Endógenos), e classificação de objetivos educacionais baseada na Taxonomia de Bloom Revisada. Foram analisados 26 estudos de caso abrangendo diferentes contextos educacionais, desde ensino fundamental até formação corporativa, contemplando disciplinas como matemática, ciências, idiomas. Os agentes foram categorizados em quatro tipos: Coordenador, Especialista em Mecânicas, Especialista em Narrativa e Especialista em Engajamento. Os resultados evidenciaram 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. Como parte da validação, foram aplicados quatro frameworks estabelecidos (TAM, TTF, AX e ADDIE), que confirmaram a eficácia do sistema em reduzir tempo de desenvolvimento mantendo qualidade pedagógica. Concluiu-se que a abordagem multiagente baseada em LLMs representa uma ferramenta robusta para democratizar o acesso ao design de jogos educacionais, mostrando-se promissora para acelerar a inovação pedagógica e ampliar o uso de metodologias ativas na educação.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

AN APPROACH USING LLM-BASED MULTI-AGENTS FOR EDUCATIONAL GAME DESIGN

Caio Silva Azeredo

June/2025

Advisor: Geraldo Bonorino Xexéo

Department: Systems Engineering and Computer Science

This study developed a multi-agent system based on Large Language Models (LLMs) to optimize the design process of endogenous educational games. The research was conducted using Design Science Research methodology, following the stages of problem awareness, artifact proposition, development, evaluation, and communication of results. The system integrates four specialized functional modules: AI-assisted Brainstorming, Socratic Agent for conceptual refinement, assisted completion of Endo-GDC (Game Design Canvas for Endogenous Educational Games), and educational objectives classification based on Bloom's Revised Taxonomy. Twenty-six case studies were analyzed spanning different educational contexts, from elementary education to corporate training, encompassing subjects such as mathematics, sciences, languages. Agents were categorized into four types: Coordinator, Mechanics Specialist, Narrative Specialist, and Engagement Specialist. Results 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. As part of validation, four established frameworks (TAM, TTF, AX, and ADDIE) were applied, confirming system effectiveness in reducing development time while maintaining pedagogical quality. It was concluded that the LLM-based multi-agent approach represents a robust tool for democratizing access to educational game design, showing promise for accelerating pedagogical innovation and expanding the use of active methodologies in education.

Sumário

Lista de Figuras	xiv
Lista de Tabelas	xvii
1 Introdução	1
1.1 Justificativa	1
1.2 Contextualização do Problema	3
1.2.1 Complexidade Multidisciplinar no Design Educacional	3
1.2.2 Tensões na Integração Pedagógico-Lúdica	4
1.2.3 Limitações Temporais e de Recursos no Desenvolvimento	4
1.2.4 Desafios na Formulação de Objetivos Educacionais	5
1.2.5 Escassez de Ferramentas Especializadas	5
1.2.6 Limitações do Conhecimento Especializado Disponível	5
1.2.7 Desafios de Personalização e Contextualização	6
1.2.8 Oportunidades Tecnológicas Emergentes	6
1.2.9 Síntese da Problemática	7
1.3 Objetivos	7
1.3.1 Objetivos Específicos	8
1.4 Design Science Research como Abordagem Metodológica	9
2 Fundamentos teóricos	11
2.1 Ensino e Aprendizagem	11
2.1.1 Taxionomia de Bloom	13
2.2 Sistemas de Agentes e LLMs	17
2.2.1 Sistemas de Agente Único	17
2.2.2 Sistemas Multiagentes	18
2.2.3 Large Language Models	22
3 Uso de LLMs em Sistemas Multi-Agentes	24
3.0.1 Características do Campo que Justificam a Metodologia	25
3.1 Processo Metodológico da Revisão Rápida	25
3.1.1 Definição da Questão Prática	25

3.1.2	Perguntas de Pesquisa	26
3.1.3	Busca por Evidências de Pesquisa	26
3.1.4	Etapas do Processo de Seleção	28
3.1.5	Distribuição Temporal e Geográfica	28
3.1.6	Validação do Processo	30
3.1.7	Caracterização dos Domínios de Aplicação	30
3.1.8	Validação do Processo	30
3.2	Resumo	34
3.2.1	Perfil do Agente	35
3.2.2	Estratégia de Geração	36
3.2.3	Percepção	37
3.2.4	Fonte de Mensagem	37
3.2.5	Tipo de Mensagem	38
3.2.6	Auto-ação	41
3.2.7	Memória	41
3.2.8	Utilização de Conhecimento	46
3.2.9	Utilização de Habilidades do Agente	53
3.2.10	Ação	59
3.3	Evolução	62
3.3.1	Fonte de Evolução	63
3.3.2	Métodos de Evolução	64
3.3.3	Ajuste de Agentes	68
4	Proposta	70
4.1	Metodologia	70
4.1.1	Proposição e Projeto do Artefato	71
4.2	Desenvolvimento do Artefato	72
4.3	Avaliação do Artefato	72
4.4	Arquitetura Geral do Sistema	73
4.4.1	Camada de Apresentação	73
4.4.2	Camada de Controle	74
4.4.3	Camada de Serviços	74
4.4.4	Camada de Persistência	75
4.5	Sistema Multiagente	75
4.5.1	Modelo de Agente	75
4.5.2	Tipos de Agentes	76
4.5.3	Perfil de Agente	76
4.5.4	Comunicação entre Agentes	77
4.5.5	Coordenação de Agentes	77

4.5.6	Mecanismo de Evolução de Agentes	78
4.6	Integração com Modelos de Linguagem	78
4.6.1	Arquitetura de Integração	78
4.6.2	Prompting Estruturado	79
4.6.3	Processamento de Respostas	79
4.6.4	Gerenciamento de Contexto	80
4.7	Módulos Funcionais	80
4.7.1	Módulo de Brainstorming	80
4.7.2	Módulo de Agente Socrático	83
4.7.3	Módulo de Endo-GDC	86
4.7.4	Módulo de Taxonomia de Bloom	90
4.8	Implementação Técnica	93
4.8.1	Tecnologias Utilizadas	93
4.8.2	Arquitetura de Software	93
4.8.3	Padrões de Design	94
4.8.4	Gestão de Configuração	94
4.8.5	Segurança e Privacidade	95
4.8.6	Testes e Qualidade	95
4.9	Desafios e Soluções	95
4.10	Gerenciamento de Recursos LLM	96
4.11	Qualidade e Relevância das Sugestões	96
4.12	Coordenação entre Agentes	97
4.13	Experiência do Usuário	97
4.14	Evolução e Manutenção	98
4.14.1	Extensibilidade	98
4.14.2	Monitoramento e Diagnóstico	98
4.14.3	Versionamento e Compatibilidade	98
4.14.4	Documentação	99
4.15	Infraestrutura e Implantação	99
4.15.1	Arquitetura de Infraestrutura	99
4.16	Comunicação com o Usuário	100
4.16.1	Interface de Usuário Web	100
4.16.2	Notificações	100
4.16.3	Exportação de Resultados	100
4.16.4	Otimização de Performance	101
4.16.5	Configurações de Agentes	101
4.16.6	API Extensível	101
4.17	Considerações Éticas	102
4.17.1	Transparência	102

4.17.2	Equidade e Inclusão	102
4.17.3	Privacidade e Segurança	103
4.17.4	Responsabilidade Humana	103
4.18	Integração com o Processo Educacional	103
4.18.1	Alinhamento Curricular	104
4.18.2	Suporte a Avaliação	104
4.18.3	Comunidade e Colaboração	104
4.18.4	Implementação e Prototipagem	105
4.19	Limitações	105
4.20	Interface do Sistema e Experiência do Usuário	106
4.20.1	Módulo de Brainstorming - Interface de Configuração	106
4.20.2	Módulo de Agente Socrático - Questionamento Reflexivo	107
4.20.3	Módulo de Taxonomia de Bloom - Estruturação de Objetivos	108
4.20.4	Módulo de Endo-GDC - Canvas de Design	108
4.20.5	Características Transversais da Interface	109
4.20.6	Fluxo de Interação do Usuário	110
5	Testes e validação	112
5.1	Estudos de Caso	112
5.2	Visão Geral dos Estudos de Caso	112
5.3	Metodologia de Validação	112
5.4	Análise Transversal dos Resultados	114
5.4.1	Eficácia do Sistema de Brainstorming	114
5.4.2	Consistência do Método Socrático	114
5.4.3	Diversidade dos Agentes Especializados	114
5.4.4	Aplicação da Taxonomia de Bloom	115
5.4.5	Métodos de Avaliação	115
5.5	Metodologia	117
5.5.1	Estrutura dos Dados e Operacionalização	117
5.5.2	Amostra e Procedimentos	119
5.6	Resultados	120
5.6.1	Estatísticas Descritivas	120
5.7	Análise Estatística dos Resultados	120
5.7.1	Análise Descritiva Detalhada	120
5.7.2	Análise Correlacional	122
5.8	Discussão	123
5.8.1	Aceitação Tecnológica (TAM) - Análise	123
5.8.2	Adequação Tarefa-Tecnologia (TTF) - Análise	124
5.8.3	Experiência Algorítmica (AX) - Análise	125

5.8.4	Avaliação do Processo (ADDIE) - Análise	125
5.9	Análise de Oportunidades de Melhoria e Estratégias de Aprimoramento	126
5.9.1	Análise das Métricas com Menor Desempenho	126
5.9.2	Melhorias Adicionais Baseadas em Feedback Qualitativo . . .	129
5.9.3	Estratégias de Implementação das Melhorias	130
5.9.4	Contribuições para o Campo	131
6	Conclusão	132
	Referências Bibliográficas	134
A	Estudos de Caso Detalhados	165
A.1	Estudo de Caso 1: Ensino de Matemática no Ensino Fundamental . .	165
A.1.1	Design de Jogo para Ensino de Matemática no Ensino Funda- mental	165
A.2	Estudo de Caso 2: Ensino de Ciências no Ensino Médio	168
A.2.1	Design de Jogo para Ensino de Ciências no Ensino Médio . . .	168
A.3	Estudo de Caso 3: Ensino de Idiomas para Adultos	171
A.3.1	Design de Jogo para Ensino de Idiomas para Adultos	171
A.4	Estudo de Caso 4: Formação em Liderança e Gestão	174
A.4.1	Design de Jogo para Formação Corporativa	174
A.5	Estudo de Caso 5: Ensino de História no Ensino Médio	177
A.5.1	Design de Jogo para Ensino de História no Ensino Médio . . .	177
A.6	Estudo de Caso 6: Ensino de Teoria Musical	179
A.6.1	Design de Jogo para Ensino de Teoria Musical para Iniciantes	179
A.7	Estudo de Caso 7: Ensino de Biologia Celular	182
A.7.1	Design de Jogo para Ensino de Biologia Celular no Ensino Médio	182
A.8	Estudo de Caso 8: Ensino de Programação no Ensino Fundamental .	185
A.8.1	Design de Jogo para Introdução à Programação para Crianças	185
A.9	Estudo de Caso 9: Aprendizado de Língua Estrangeira para Viajantes	188
A.9.1	Design de Jogo para Aprendizado de Idiomas em Contexto de Viagem	188
A.10	Estudo de Caso 10: Ensino de Física no Ensino Médio	190
A.10.1	Design de Jogo para Ensino de Física Newtoniana	190
A.11	Estudo de Caso 11: Educação Financeira para Jovens Adultos	193
A.11.1	Design de Jogo para Desenvolvimento de Literacia Financeira	193
A.12	Estudo de Caso 12: Ensino de Geografia e Estudos Culturais	197
A.12.1	Design de Jogo para Exploração Geográfica e Cultural	197
A.13	Estudo de Caso 13: Educação Ambiental para Ensino Fundamental II	200

A.13.1 Design de Jogo para Conscientização Ambiental e Sustentabilidade	200
A.14 Estudo de Caso 14: Ensino de Geografia no Ensino Fundamental II	203
A.14.1 Design de Jogo para Ensino de Geografia no Ensino Fundamental II	203
A.15 Estudo de Caso 15: Treinamento de Atendimento ao Cliente em Ambiente Corporativo	206
A.15.1 Design de Jogo para Formação de Soft Skills em Empresas	206
A.16 Estudo de Caso 16: Ensino de Química no Ensino Médio	209
A.16.1 Design de Jogo para Aprendizado de Química Interativa	209
A.17 Estudo de Caso 17: Educação em Saúde para Adultos	212
A.17.1 Design de Jogo para Promoção de Bem-Estar e Prevenção	212
A.18 Estudo de Caso 18: Educação Cívica para Jovens	215
A.18.1 Design de Jogo para Desenvolvimento de Cidadania Ativa	215
A.19 Estudo de Caso 19: Desenvolvimento de Habilidades Socioemocionais para Crianças	218
A.19.1 Design de Jogo para Inteligência Emocional e Relacionamentos Saudáveis	218
A.20 Estudo de Caso 20: Educação Empreendedora para Jovens Adultos	221
A.20.1 Design de Jogo para Desenvolvimento de Mentalidade Empreendedora	221
A.21 Estudo de Caso 21: Aprendizado de Línguas para Crianças	224
A.21.1 Design de Jogo para Ensino de Idiomas no Ensino Fundamental I	224
A.22 Estudo de Caso 22: Exploração Espacial e Astronomia	227
A.22.1 Design de Jogo para Ensino de Conceitos Astronômicos	227
A.23 Estudo de Caso 23: Pensamento Crítico e Lógica	230
A.23.1 Design de Jogo para Desenvolvimento de Raciocínio Lógico	230
A.24 Estudo de Caso 24: Sustentabilidade e Gestão de Recursos	233
A.24.1 Design de Jogo para Educação Ambiental e Econômica	233
A.25 Estudo de Caso 25: Expressão Artística e Criatividade	236
A.25.1 Design de Jogo para Desenvolvimento de Habilidades Artísticas	236
A.26 Estudo de Caso 26: Programação e Pensamento Computacional	239
A.26.1 Design de Jogo para Ensino de Conceitos de Programação	239

Lista de Figuras

2.1	Visão geral do sistema multi-agente. O processo inicia com o perfil do agente que contém características personalizadas e alocação de subtarefas. O agente formula planos específicos para perceber informações multimodais do ambiente interativo, acessa conhecimento externo e recupera experiências históricas da memória. Utilizando as capacidades dos LLMs, os agentes desenvolvem planos de ação concretos. O processo inclui evolução contínua através da reflexão sobre decisões e ações. A execução das tarefas depende das interações entre agentes, que contribuem coletivamente para o planejamento e implementação da missão geral. Fonte: O autor	19
3.1	Fluxograma do processo de seleção de estudos para revisão bibliográfica sobre integração de LLMs em sistemas multiagentes	29
3.2	Representação das funcionalidades centrais do módulo de memória em agentes baseados em LLM, mostrando como as três operações críticas (recuperação, armazenamento e reflexão de memória) trabalham em conjunto para permitir adaptação dinâmica e navegação flexível em ambientes complexos através de um ciclo contínuo de aprendizado e aprimoramento cognitivo.	42
3.3	Representação de como agentes baseados em LLM integram conhecimento externo multimodal (textual, visual e auditivo) através de técnicas como RAG e web scraping em tempo real para aprimorar o planejamento e tomada de decisão contextualizada.	46
3.4	Representação de raciocínio e planejamento em agentes, distinguindo entre abordagens de passo único (dedutivo, indutivo e abdutivo) e múltiplos passos (com decomposição sequencial), incluindo mecanismos de avaliação e refinamento iterativo.	54

3.5	Representação do ciclo de evolução contínua em agentes, mostrando como feedback proveniente do ambiente, outros agentes e humanos é processado através de três mecanismos principais (atualização de memória, autorreflexão e geração dinâmica) para aprimoramento iterativo das capacidades do agente	62
4.1	Diagrama de Pacotes do sistema multi-agente para design de jogos educacionais.	73
4.2	Diagrama de classes do módulo de Brainstorming.	81
4.3	Diagrama entidade-relacionamento do módulo de Brainstorming.	82
4.4	Diagrama de sequência do módulo de Brainstorming.	82
4.5	Diagrama de classes do módulo de Agente Socrático.	84
4.6	Diagrama entidade-relacionamento do módulo de Agente Socrático.	85
4.7	Diagrama de sequência do módulo de Agente Socrático.	85
4.8	Diagrama de classes do módulo de Endo-GDC.	87
4.9	Diagrama entidade-relacionamento do módulo de Endo-GDC.	88
4.10	Diagrama de sequência do módulo de Endo-GDC.	89
4.11	Diagrama de classes do módulo de Taxonomia de Bloom.	90
4.12	Diagrama entidade-relacionamento do módulo de Taxonomia de Bloom.	91
4.13	Diagrama de sequência do módulo de Taxonomia de Bloom.	91
4.14	Interface de configuração do módulo de brainstorming, mostrando campos para nome do projeto, área de conhecimento, público-alvo e duração da sessão. A interface também exibe os agentes especializados disponíveis (Mecânicas, Narrativa, Engajamento e Área Livre) e permite seleção específica conforme necessidade do projeto.	106
4.15	Interface do módulo de Agente Socrático apresentando questão reflexiva sobre equilíbrio entre desafio educacional e diversão. A tela inclui área de resposta textual, botão para geração automática baseada no brainstorming anterior, e painel lateral com perguntas socráticas organizadas por categoria e dicas para reflexão estruturada.	107
4.16	Interface do módulo de Taxonomia de Bloom organizada em seis níveis cognitivos hierárquicos: Lembrar, Compreender, Aplicar, Analisar, Avaliar e Criar. Cada nível apresenta área para criação de objetivos específicos com botões para adição manual ou geração baseada em etapas anteriores. O painel lateral oferece gerador automático e indicadores de progresso por categoria.	108

4.17	Interface do módulo Endo-GDC mostrando canvas interativo com nove seções principais: Experiência do Jogador, Objetivos de Aprendizagem, Loop de Jogabilidade, Mecânicas de Jogo, Conceito Central, Avaliação, Narrativa & Tema, e Tecnologia & Plataforma. A interface permite adição de notas em cada seção e oferece funcionalidades para geração automática, visualização, importação e exportação.	109
A.1	Endo-GDC do jogo “Ilha dos Números”	167
A.2	Endo-GDC do jogo “EcoSphere”	170
A.3	Endo-GDC do jogo “LinguaViva”	173
A.4	Endo-GDC do jogo “LeadSync”	176
A.5	Endo-GDC do jogo “Tempos de Mudança”	179
A.6	Endo-GDC do jogo “Harmonix”	182
A.7	Endo-GDC do jogo “CellQuest”	185
A.8	Endo-GDC do jogo “CodeQuest”	188
A.9	Endo-GDC do jogo “Linguatrip”	191
A.10	Endo-GDC do jogo “PhysicsLab”	194
A.11	Endo-GDC do jogo “FinLife”	197
A.12	Endo-GDC do jogo “GeoVentures”	200
A.13	Endo-GDC do jogo “EcoGuardiões”	203
A.14	Endo-GDC do jogo “GeoExplorer”	205
A.15	Endo-GDC do jogo “ClientCare Pro”	208
A.16	Endo-GDC do jogo “ChemLab Ventures”	211
A.17	Endo-GDC do jogo “VitalQuest”	214
A.18	Endo-GDC do jogo “CivicSphere”	217
A.19	Endo-GDC do jogo “EmotionQuest”	220
A.20	Endo-GDC do jogo “Venture Forge”	223
A.21	Endo-GDC do jogo “LinguaMagica”	226
A.22	Endo-GDC do jogo “CosmicVoyage”	229
A.23	Endo-GDC do jogo “LogicCraft”	232
A.24	Endo-GDC do jogo “EcoSystems”	235
A.25	Endo-GDC do jogo “ArtSphere”	238
A.26	Endo-GDC do jogo “CodeRealm”	241

Lista de Tabelas

3.1	Revisão de trabalhos em sistemas multiagentes baseados em LLM e IA	30
4.1	Processo de Design Science Research	111
5.1	Síntese dos Estudos de Caso Realizados	113
5.2	Estatísticas Descritivas dos Modelos de Avaliação	121

Capítulo 1

Introdução

Este capítulo apresenta o contexto da pesquisa sobre sistemas multiagentes baseados em LLMs para design de jogos educativos, estabelecendo a problemática do desenvolvimento de jogos educacionais endógenos, a justificativa para aplicação de tecnologias emergentes neste contexto, e os objetivos que orientam o desenvolvimento de uma solução automatizada para otimização do processo de design pedagógico.

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).

Os Grandes Modelos de Linguagem (LLMs) apresentam capacidades de processamento de linguagem natural que podem ser aplicadas em sistemas multiagentes para automação de processos de design (ZHANG *et al.*, 2024e). A integração destas tecnologias em ferramentas de design de jogos educacionais representa uma oportunidade para otimização do processo de desenvolvimento e democratização do acesso a estas metodologias pedagógicas.

Esta dissertação apresenta um sistema multiagente baseado em LLMs para otimizar o design de jogos educacionais endógenos através do preenchimento assistido do Endo-GDC (Game Design Canvas para Jogos Educativos Endógenos). O sistema integra quatro módulos funcionais: brainstorming assistido por IA, agente socrático para refinamento conceitual, preenchimento assistido do canvas, e classificação de objetivos educacionais baseada na Taxonomia de Bloom.

1.1 Justificativa

A criação de jogos educacionais requer tempo considerável e conhecimento especializado, limitando sua adoção em contextos educacionais (PRENSKY, 2001). Estudos

indicam que o processo de design de jogos educacionais endógenos demanda compreensão simultânea de mecânicas de jogo, objetivos pedagógicos e estratégias de engajamento (TAUCEI, 2019).

Os sistemas multiagentes baseados em LLMs demonstram capacidades de coordenação e comunicação que podem ser aplicadas para automação de processos complexos (CHEN *et al.*, 2023c). A aplicação destas tecnologias no contexto educacional pode contribuir para redução do tempo de desenvolvimento mantendo qualidade pedagógica.

O Endo-GDC constitui uma ferramenta estruturada para design de jogos educacionais, porém seu preenchimento manual demanda conhecimento especializado e tempo considerável (TAUCEI, 2019). A automação assistida deste processo através de sistemas multiagentes pode ampliar o acesso a metodologias de gamificação na educação.

A pesquisa justifica-se pela necessidade de democratização das ferramentas de design educacional e pela oportunidade de aplicação de tecnologias emergentes para solução de problemas práticos no contexto pedagógico.

O desenvolvimento de jogos educacionais requer a integração entre elementos de design de jogos e princípios pedagógicos. Esta integração fundamenta-se nos estudos de MAYER (2014), que estabelece princípios cognitivos para o design de materiais educacionais multimídia, e de CLARK e MAYER (2016), que expandem estes princípios para ambientes digitais interativos.

A criação de jogos educacionais demanda processos estruturados de desenvolvimento. BOYLE *et al.* (2016), em uma revisão sistemática com 143 estudos, identificou elementos centrais para a eficácia dos jogos educacionais, incluindo objetivos de aprendizagem claros e feedback imediato e progressão adequada de dificuldade.

A implementação de jogos educacionais em contextos escolares apresenta desafios específicos de integração curricular. Um estudo longitudinal conduzido por ABDUL JABBAR e FELICIA (2015) com 91 publicações revelou que a efetividade dos jogos educacionais depende de sua inserção em um planejamento pedagógico estruturado. Esta conclusão alinha-se com os resultados de ALL *et al.* (2014), que propõem um framework para avaliação sistemática da implementação de jogos educacionais.

A medição do impacto dos jogos na aprendizagem requer metodologias específicas. CLARK *et al.* (2016), em uma meta-análise com 57 estudos comparativos, estabeleceram protocolos para avaliação de resultados educacionais em jogos digitais. Estes protocolos foram expandidos por QIAN e CLARK (2016), que analisaram 137 artigos para identificar métricas de avaliação em diferentes contextos educacionais.

Os mecanismos pelos quais os jogos educacionais promovem aprendizagem fundamentam-se em teorias cognitivas estabelecidas. MAYER *et al.* (2019) condu-

ziram experimentos controlados demonstrando como elementos de jogos ativam processos cognitivos específicos. Estes resultados complementam os estudos de SHUTE *et al.* (2015), que estabeleceram correlações entre mecânicas de jogo e desenvolvimento de habilidades cognitivas através de análise de dados de 300 participantes.

O desenvolvimento de metodologias para criação de jogos educacionais constitui um campo em expansão. ARNAB *et al.* (2015) propõem um framework que associa mecânicas de jogo a objetivos pedagógicos, baseado na análise de 42 casos de implementação. Este trabalho fundamenta-se nos estudos de LAMERAS *et al.* (2017), que identificaram padrões de design em 165 jogos educacionais.

A incorporação de técnicas de Inteligência Artificial em jogos educacionais representa uma fronteira de pesquisa. LIU *et al.* (2017) realizaram uma revisão sistemática de 50 estudos sobre aplicações de IA em jogos educacionais, identificando contribuições para adaptação de conteúdo e avaliação de aprendizagem. Estas conclusões são corroboradas por HOLMES *et al.* (2019), que analisaram 75 implementações de IA em ambientes educacionais.

1.2 Contextualização do Problema

O desenvolvimento de jogos educacionais constitui um desafio multidisciplinar que intersecta conhecimentos pedagógicos, tecnológicos e de design de jogos. Esta intersecção, embora promissora para a criação de experiências de aprendizagem inovadoras, apresenta complexidades que limitam significativamente a capacidade de educadores e instituições de desenvolverem recursos educacionais digitais de qualidade.

1.2.1 Complexidade Multidisciplinar no Design Educacional

A criação de jogos educacionais eficazes demanda expertise simultânea em múltiplas áreas do conhecimento. Conforme destacado por (PRENSKY, 2001), os jogos digitais podem ser incorporados nos processos educacionais, mas esta integração requer compreensão profunda tanto dos mecanismos de engajamento dos jogos quanto dos processos cognitivos da aprendizagem.

O processo tradicional de design instrucional, fundamentado no modelo ADDIE (BRANCH, 2009), quando aplicado ao contexto de jogos educacionais, torna-se exponencialmente mais complexo. Educadores precisam navegar simultaneamente entre análise de necessidades de aprendizagem, princípios de design de jogos, teorias de motivação, tecnologias de desenvolvimento e metodologias de avaliação educacional.

Esta multiplicidade de competências necessárias cria uma barreira de entrada elevada que restringe o desenvolvimento de jogos educacionais a equipes especializadas e recursos consideráveis, frequentemente inacessíveis para educadores individuais ou instituições com limitações orçamentárias.

1.2.2 Tensões na Integração Pedagógico-Lúdica

A literatura especializada identifica tensões inerentes entre objetivos educacionais e mecânicas de jogo. (GEE, 2003) documenta como os videogames influenciam o aprendizado através de princípios específicos, mas a aplicação destes princípios no design educacional requer equilíbrio delicado entre elementos lúdicos e pedagógicos.

Elementos que promovem engajamento e diversão podem inadvertidamente desviar a atenção do conteúdo educacional central. Por outro lado, foco excessivo nos aspectos pedagógicos pode resultar em experiências lúdicas artificiais e pouco envolventes, comprometendo o potencial motivacional dos jogos.

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 esta tensão. Contudo, o design de jogos endógenos é particularmente desafiador, pois requer identificação criativa de formas autênticas de integração entre domínio educacional e mecânicas de jogo.

1.2.3 Limitações Temporais e de Recursos no Desenvolvimento

O desenvolvimento de jogos educacionais de qualidade constitui um processo intensivo em tempo e recursos. (BOLLER e KAPP, 2017) destacam que métodos de design de jogos educacionais demandam iterações múltiplas e refinamentos constantes para garantir efetividade pedagógica.

Esta demanda temporal é particularmente problemática no contexto educacional, onde currículos evoluem rapidamente, necessidades específicas de aprendizagem requerem respostas ágeis, e recursos para desenvolvimento são frequentemente limitados. A lacuna entre o tempo necessário para desenvolvimento de qualidade e os prazos educacionais típicos resulta frequentemente em soluções subótimas ou abandono de projetos.

Além do tempo de desenvolvimento, custos associados à expertise multidisciplinar, ferramentas de desenvolvimento, prototipagem, testes de usabilidade e iterações baseadas em feedback frequentemente excedem orçamentos disponíveis para instituições educacionais.

1.2.4 Desafios na Formulação de Objetivos Educacionais

A definição de objetivos educacionais claros, mensuráveis e pedagogicamente fundamentados constitui etapa crucial no design instrucional. A Taxonomia de Bloom, particularmente em sua versão revisada (CHURCHES, 2008), fornece framework estruturado para classificação de objetivos em diferentes níveis cognitivos.

No contexto de jogos educacionais, esta tarefa torna-se particularmente complexa devido à necessidade de traduzir objetivos abstratos em mecânicas de jogo concretas e mensuráveis. A aplicação prática da Taxonomia de Bloom no design de jogos requer compreensão de como diferentes mecânicas podem suportar processos cognitivos específicos, desde recordação factual até criação de conhecimento novo.

Educadores frequentemente enfrentam dificuldades para articular objetivos educacionais de forma precisa e operacional, especialmente quando devem considerar simultaneamente aspectos cognitivos, afetivos e psicomotores da aprendizagem integrados à experiência lúdica.

1.2.5 Escassez de Ferramentas Especializadas

O ecossistema atual de ferramentas de desenvolvimento apresenta dicotomia problemática. Ferramentas de desenvolvimento de jogos convencionais focam primariamente em aspectos técnicos e de entretenimento, oferecendo suporte limitado para considerações pedagógicas específicas.

Por outro lado, ferramentas de design instrucional tradicionais não contemplam adequadamente as especificidades do desenvolvimento de jogos, particularmente os aspectos relacionados à integração orgânica entre conteúdo educacional e mecânicas lúdicas.

Esta lacuna resulta em processos de desenvolvimento fragmentados, nos quais designers educacionais utilizam múltiplas ferramentas desconectadas, aumentando complexidade, tempo de desenvolvimento e probabilidade de inconsistências entre componentes pedagógicos e lúdicos.

1.2.6 Limitações do Conhecimento Especializado Disponível

A intersecção entre design de jogos e educação constitui domínio de conhecimento altamente especializado. Como observado por (KALMPOURTZIS, 2018), a criação de jogos educacionais envolve integração de múltiplos elementos de design e pedagogia, demandando expertise que raramente se encontra concentrada em indivíduos ou equipes.

Esta escassez de expertise especializada resulta frequentemente em situações subótimas: jogos tecnicamente sofisticados mas pedagogicamente deficientes, ou recur-

sos educacionalmente sólidos mas com baixo potencial de engajamento e motivação.

A literatura acadêmica sobre melhores práticas para design de jogos educacionais, embora crescente, permanece fragmentada entre diferentes disciplinas, dificultando o acesso a conhecimento consolidado e aplicável por parte de educadores e designers.

1.2.7 Desafios de Personalização e Contextualização

Contextos educacionais apresentam heterogeneidade significativa em termos de objetivos curriculares, características dos aprendizes, recursos disponíveis, culturas organizacionais e necessidades específicas de aprendizagem. Jogos educacionais eficazes devem ser adaptáveis a esta diversidade sem comprometer qualidade ou coerência pedagógica.

A personalização de jogos educacionais para diferentes perfis de aprendizes, níveis de conhecimento prévio e contextos institucionais específicos requer sofisticação técnica e pedagógica que excede as capacidades de desenvolvimento disponíveis para a maioria dos educadores.

Esta limitação resulta frequentemente em recursos educacionais genéricos que podem não atender adequadamente às necessidades específicas de contextos educacionais particulares, reduzindo potencial de impacto e efetividade pedagógica.

1.2.8 Oportunidades Tecnológicas Emergentes

O desenvolvimento acelerado de tecnologias de inteligência artificial, particularmente dos Grandes Modelos de Linguagem (LLMs), cria oportunidades inéditas para automação e assistência em processos criativos complexos. Conforme documentado na literatura recente sobre aplicações de LLMs em contextos educacionais (ZHANG *et al.*, 2024e), estas tecnologias demonstram capacidades de processamento de linguagem natural, geração de conteúdo contextualizado e simulação de diferentes tipos de expertise.

A capacidade dos LLMs de compreender requisitos educacionais expressos em linguagem natural, gerar sugestões contextualmente relevantes e simular diferentes perspectivas especializadas oferece potencial transformador para o design de jogos educacionais. Sistemas baseados em LLMs poderiam, teoricamente, assistir educadores na geração de ideias, refinamento de conceitos, formulação de objetivos educacionais e estruturação de experiências de aprendizagem.

Contudo, a aplicação prática destas tecnologias no contexto específico do design de jogos educacionais permanece largamente inexplorada, representando lacuna significativa entre potencial tecnológico e aplicação educacional concreta.

1.2.9 Síntese da Problemática

A convergência destes desafios resulta em cenário onde o desenvolvimento de jogos educacionais de qualidade permanece acessível apenas a parcela limitada de profissionais e instituições com recursos especializados consideráveis. Esta limitação tem implicações diretas para a qualidade e diversidade dos recursos educacionais digitais disponíveis.

As barreiras identificadas – complexidade multidisciplinar, tensões pedagógico-lúdicas, limitações temporais e de recursos, dificuldades na formulação de objetivos, escassez de ferramentas especializadas, limitações de expertise e desafios de personalização – convergem para criar um gargalo significativo na inovação educacional digital.

A necessidade de democratizar o acesso a ferramentas e processos de design de jogos educacionais torna-se, portanto, questão não apenas técnica, mas também de equidade educacional. Soluções que reduzam barreiras de entrada, acelerem processos de desenvolvimento e mantenham qualidade pedagógica podem ter impacto significativo na adoção de metodologias ativas de aprendizagem.

É neste contexto de desafios convergentes e oportunidades tecnológicas emergentes que se insere a presente pesquisa, propondo a aplicação de sistemas multiagentes baseados em LLMs como abordagem inovadora para otimizar e democratizar o processo de design de jogos educacionais endógenos.

1.3 Objetivos

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.**

Este trabalho foca especialmente no preenchimento assistido do Endo-GDC (Game Design Canvas para Jogos Educativos Endógenos), visando reduzir a complexidade e o tempo necessário para o desenvolvimento de jogos educacionais eficazes (TAUCEI, 2019).

A proposta busca integrar os avanços recentes em IA com as metodologias estabelecidas de design de jogos educacionais, criando uma ferramenta que possa:

- Facilitar o processo de design e desenvolvimento de jogos educacionais
- Manter a qualidade e eficácia pedagógica do método original
- Reduzir o tempo e recursos necessários para o desenvolvimento
- Promover maior consistência nos resultados produzidos

1.3.1 Objetivos Específicos

Investigar Métodos de Assistência Inteligente para Geração de Ideias Educacionais

Analisar como sistemas baseados em LLMs podem auxiliar educadores no processo criativo de geração e organização de conceitos para jogos educacionais, investigando técnicas de processamento de linguagem natural que facilitem a expansão e refinamento de ideias iniciais, bem como mecanismos de categorização e validação automática de propostas educacionais.

Desenvolver Abordagens para Refinamento Conceitual em Design Educacional

Investigar a aplicação de métodos socráticos automatizados para promover reflexão crítica e refinamento de conceitos educacionais, explorando como questionamento estruturado pode identificar inconsistências, lacunas e oportunidades de melhoria em propostas de jogos educacionais, contribuindo para maior rigor pedagógico no processo de design.

Examinar Estratégias de Automação para Classificação de Objetivos Educacionais

Investigar métodos computacionais para classificação e formulação automática de objetivos educacionais alinhados com frameworks pedagógicos estabelecidos, particularmente a Taxonomia de Bloom Revisada, analisando como técnicas de processamento de linguagem natural podem identificar e categorizar objetivos de aprendizagem de forma pedagogicamente consistente.

Analisar Abordagens Multiagentes para Suporte ao Design Instrucional

Examinar como sistemas multiagentes especializados podem colaborar para assistir o processo de design de jogos educacionais, investigando mecanismos de coordenação entre agentes com diferentes expertises e analisando estratégias para manutenção da coerência global em processos de design colaborativo assistido por IA.

Avaliar a Eficácia de Sistemas de Assistência Automatizada no Contexto Educacional

Investigar a aceitação, utilidade percebida e adequação de sistemas baseados em LLMs para design educacional, analisando como educadores interagem com ferramentas de assistência inteligente e examinando fatores que influenciam a adoção e efetividade de tecnologias emergentes em contextos pedagógicos.

Desenvolver Framework para Integração de IA no Design de Jogos Educacionais

Propor e validar um framework conceitual para integração de tecnologias de inteligência artificial no processo de design de jogos educacionais endógenos, investigando princípios de design que preservem a autonomia do educador enquanto amplificam suas capacidades criativas e pedagógicas através de assistência automatizada.

1.4 Design Science Research como Abordagem Metodológica

Esta pesquisa adota o Design Science Research (DSR) como abordagem metodológica para o desenvolvimento de um sistema multiagente baseado em LLMs para o design de jogos educativos endógenos. A escolha do DSR deve-se à sua característica de conectar a relevância prática com o rigor científico, permitindo o desenvolvimento e a avaliação de artefatos tecnológicos inovadores voltados para solucionar problemas reais (DRESCH *et al.*, 2020).

O desenvolvimento de jogos educacionais apresenta desafios específicos, principalmente relacionados à complexidade e ao tempo necessário para a criação de experiências pedagógicas efetivas (TAUCEI, 2019). O preenchimento assistido do Endo-GDC (Game Design Canvas para Jogos Educativos Endógenos) representa uma oportunidade para a integração de tecnologias de Inteligência Artificial, visando otimizar este processo.

O DSR constitui uma abordagem de pesquisa científica rigorosa orientada à concepção de artefatos que resolvam problemas relevantes, contribuindo simultaneamente para o avanço do conhecimento científico (PIMENTEL *et al.*, 2020a). Esta metodologia é particularmente adequada para pesquisas em computação aplicada à educação, como evidenciado em trabalhos anteriores no desenvolvimento de jogos educacionais (COSTA *et al.*, 2024; NOËL *et al.*, 2021).

O DSR diferencia-se de outras abordagens de pesquisa por seu foco na produção de artefatos que visam transformar a realidade existente em estados preferíveis (VOM BROCKE *et al.*, 2020). Como destacam (DRESCH *et al.*, 2020), enquanto as ciências tradicionais buscam entender "o que é", o DSR procura compreender "o que funciona", estabelecendo um paradigma orientado à solução de problemas.

Segundo (PIMENTEL *et al.*, 2020b), o DSR caracteriza-se por:

- Foco na solução de problemas relevantes e práticos
- Produção de artefatos inovadores (constructos, modelos, métodos ou instâncias)
- Avaliação rigorosa das contribuições do artefato

- Articulação clara entre a pesquisa e problemas do mundo real
- Comunicação efetiva dos resultados para audiências acadêmicas e profissionais

Esta metodologia cria um ciclo virtuoso entre a construção do artefato e a geração de conhecimento, onde a prática alimenta a teoria e vice-versa, promovendo uma pesquisa pragmática e orientada à utilidade, sem abrir mão do rigor científico necessário para a geração de conhecimento confiável (DRESCH *et al.*, 2020).

Capítulo 2

Fundamentos teóricos

Este capítulo apresenta as bases conceituais que fundamentam a pesquisa, abordando teorias de ensino e aprendizagem, sistemas multiagentes, Grandes Modelos de Linguagem, metodologias de design de jogos educacionais, e a integração tecnológica necessária para o desenvolvimento do sistema proposto.

2.1 Ensino e Aprendizagem

A transformação das tecnologias digitais na educação levou a mudanças nos processos de ensino e aprendizagem. (PRENSKY, 2001) desenvolveu uma análise sobre as características das novas gerações de estudantes e sua relação com tecnologias digitais, propondo o termo "nativos digitais". O estudo estabelece bases teóricas sobre como os jogos digitais podem ser incorporados nos processos educacionais, identificando padrões de comportamento e aprendizagem específicos desta geração. A pesquisa também apresenta métodos práticos para integração de elementos digitais no ensino, baseando-se em estudos de caso e experimentos realizados em diferentes contextos educacionais. As pesquisas sobre videogames e educação demonstram conexões entre os mecanismos de engajamento dos jogos e os processos cognitivos. (GEE, 2003) conduziu um estudo extensivo sobre como os videogames influenciam o aprendizado, analisando os elementos estruturais dos jogos que se relacionam com teorias educacionais estabelecidas. A pesquisa examinou diversos títulos comerciais e educacionais, identificando 36 princípios de aprendizagem presentes nos jogos bem sucedidos. O trabalho também incluiu observações sobre como estes princípios se manifestam em diferentes contextos educacionais e como podem ser adaptados para diferentes objetivos pedagógicos.

Os princípios de aprendizagem presentes nos videogames fornecem estruturas para desenvolvimento de práticas educacionais. (GEE, 2007) examinou padrões de interação e aprendizagem em jogos digitais, documentando como diferentes mecânicas de jogo se relacionam com processos cognitivos específicos. O estudo incluiu

análise de dados de sessões de jogo, entrevistas com desenvolvedores e educadores, e experimentos controlados em ambientes educacionais. A pesquisa identificou correlações entre elementos de design de jogos e resultados de aprendizagem, apresentando um framework para implementação destes princípios em contextos educacionais diversos.

O desenvolvimento dos jogos sérios como ferramentas educacionais apresenta características específicas que os distinguem dos jogos de entretenimento. (MICHAEL e CHEN, 2005) realizou um mapeamento sistemático do campo dos jogos sérios, documentando suas aplicações em educação, treinamento e diferentes setores profissionais. O estudo examinou mais de 100 casos de implementação, identificando fatores críticos para o sucesso dos jogos sérios em diferentes contextos. Os autores também desenvolveram um framework para avaliação e implementação de jogos sérios, baseado em dados coletados de instituições educacionais e organizações que utilizam esta abordagem.

As interações entre jogadores e sistemas de jogos criam padrões específicos de participação e aprendizagem. (SQUIRE, 2011) conduziu uma pesquisa extensiva sobre como videogames promovem aprendizagem participativa, analisando dados de interações em ambientes educacionais gamificados. O estudo examinou como diferentes tipos de jogos afetam o engajamento dos estudantes, documentando padrões de participação e colaboração em ambientes de aprendizagem. A pesquisa também desenvolveu métodos para integração de elementos participativos em design de jogos educacionais, baseados em observações sistemáticas de comportamento dos estudante

A criação de jogos educacionais envolve a integração de múltiplos elementos de design e pedagogia. (KALMPOURTZIS, 2018) conduziu uma investigação sobre os componentes fundamentais do design de jogos educacionais, analisando dados de desenvolvimento e implementação de mais de 50 projetos. O estudo examinou as relações entre objetivos pedagógicos, mecânicas de jogo e Engajamento dos estudantes, identificando padrões de design que contribuem para resultados educacionais. A pesquisa também desenvolveu metodologias para avaliação e iteração de jogos educacionais, baseadas em métricas quantitativas e qualitativas.

Os processos de design de jogos educacionais demandam metodologias específicas para garantir efetividade pedagógica. (BOLLER e KAPP, 2017) realizou um estudo sistemático sobre métodos de design de jogos educacionais, compilando dados de múltiplos projetos e experimentos em ambientes educacionais. A pesquisa estabeleceu correlações entre elementos de design e resultados de aprendizagem, desenvolvendo um conjunto de diretrizes baseadas em evidências para criação de jogos educacionais. O trabalho também inclui análises de casos de implementação em diferentes contextos, identificando fatores que influenciam o sucesso dos jogos como

ferramentas de aprendizagem.

No cenário educacional contemporâneo, a crescente integração de tecnologias digitais impulsiona a busca por métodos inovadores de ensino-aprendizagem. Essa tendência é fundamentada em abordagens que enfatizam a importância de ambientes dinâmicos e colaborativos, capazes de estimular a participação ativa dos alunos (GEE, 2007; PRENSKY, 2001). Estudos indicam que tais ambientes favorecem não apenas a aprendizagem, mas também o desenvolvimento de competências essenciais para o século XXI (PLASS *et al.*, 2015; SAILER *et al.*, 2017).

O desenvolvimento de jogos educacionais eficazes requer uma compreensão tanto dos princípios de design de jogos quanto das teorias pedagógicas fundamentais (GEE, 2007; PRENSKY, 2001). Essa intersecção entre ludicidade e pedagogia apresenta desafios únicos, demandando abordagens metodológicas específicas e ferramentas adequadas para sua implementação (ECK, 2006).

A complexidade inerente a esse processo tem motivado pesquisadores e desenvolvedores a buscar soluções inovadoras que possam otimizar e facilitar a criação de jogos educacionais de qualidade (ECK, 2006).

2.1.1 Taxionomia de Bloom

A Taxonomia de Bloom emergiu de um processo sistemático de classificação de objetivos educacionais, publicado em 1964 por uma comissão multi-institucional de educadores liderada por Benjamin Bloom. O framework original, publicado em BLOOM *et al.* (1964), estabelece uma hierarquia de processos cognitivos estruturada em seis níveis. A taxonomia resultou de extensivas discussões e análises de mais de 30 especialistas ao longo de conferências realizadas entre 1949 e 1953, com sucessivas revisões do framework até sua publicação em 1956. KRATHWOHL (2002), um dos membros originais da comissão, documenta que o processo de desenvolvimento incluiu a análise de milhares de objetivos educacionais de diferentes instituições para estabelecer as categorias e sua hierarquia.

A estrutura original da taxonomia apresenta uma progressão hierárquica dos processos cognitivos: conhecimento, compreensão, aplicação, análise, síntese e avaliação. ANDERSON e KRATHWOHL (2001), em uma revisão publicada em 2001, modificaram esta estrutura em dois aspectos fundamentais: a mudança dos substantivos para verbos (lembrar, entender, aplicar, analisar, avaliar e criar) e a alteração na ordem dos dois últimos níveis, posicionando "criar" como o processo cognitivo mais complexo. Esta revisão fundamentou-se em uma análise de 50 anos de pesquisas sobre a aplicação da taxonomia original, incluindo estudos em diferentes contextos educacionais e culturais. A revisão também incorporou avanços da psicologia cognitiva e das teorias de aprendizagem desenvolvidas após a publicação original.

A base teórica da Taxonomia de Bloom integra-se com teorias cognitivas do processamento de informação. MARZANO e KENDALL (2006) realizaram uma análise comparativa entre a taxonomia e modelos cognitivos contemporâneos, examinando 24 frameworks educacionais desenvolvidos entre 1956 e 2001. Seu estudo identificou convergências entre os níveis da taxonomia e estágios de processamento cognitivo documentados em pesquisas neurocientíficas. Os autores mapearam como cada nível da taxonomia corresponde a padrões específicos de ativação neural durante tarefas de aprendizagem, estabelecendo correlações entre a hierarquia proposta por Bloom e dados empíricos sobre processamento cognitivo.

O nível "entender" da taxonomia compreende sete processos cognitivos específicos: interpretar, exemplificar, classificar, sumarizar, inferir, comparar e explicar. HATTIE (2008), em uma meta-análise de 800 estudos envolvendo 240 milhões de estudantes, identificou que intervenções educacionais focadas nestes processos de compreensão produzem efeitos significativos na aprendizagem. A pesquisa demonstrou que atividades de classificação e comparação apresentam os maiores impactos, seguidas por processos de interpretação e explicação, e atividades de inferência e sumário. Os dados revelam uma relação direta entre o tempo dedicado a cada processo e os resultados de aprendizagem obtidos.

O nível "aplicar" envolve a execução e implementação de procedimentos em situações específicas. ERICSSON *et al.* (2006) conduziram estudos longitudinais com 300 participantes ao longo de cinco anos, documentando como a aplicação sistemática de conhecimentos contribui para o desenvolvimento de expertise. A pesquisa identificou que indivíduos que engajam em prática deliberada, aplicando conhecimentos em situações progressivamente mais complexas, desenvolvem estruturas cognitivas mais robustas. Os dados revelam que o tempo médio necessário para desenvolver proficiência em aplicações complexas varia entre 50 e 100 horas de prática estruturada.

O processo de "análise" requer decomposição de material em partes constituintes e compreensão das relações entre estas partes. PINTRICH (2002) conduziu um estudo com 178 estudantes universitários, monitorando suas estratégias de análise através de protocolos think-aloud e medidas de autorregulação. Os resultados indicam relações significativas entre habilidades analíticas, metacognição e autorregulação da aprendizagem. O estudo documentou que estudantes que demonstram maior capacidade de análise também apresentam melhor monitoramento do próprio processo de aprendizagem.

O nível "avaliar" envolve processos de verificação e crítica baseados em critérios estabelecidos. SHEPARD (2000) examinou 150 práticas avaliativas em diferentes contextos educacionais, documentando como processos avaliativos estruturados promovem o desenvolvimento do pensamento crítico. A pesquisa identificou que estu-

dantes expostos a critérios explícitos de avaliação desenvolvem maior capacidade de julgamento crítico. Os dados estabelecem conexões entre a capacidade de avaliar, o pensamento crítico e a tomada de decisão.

O nível "criar" representa o processo cognitivo mais complexo na taxonomia revisada, envolvendo gerar, planejar e produzir. SAWYER (2014) conduziu estudos experimentais com 250 participantes, documentando os processos cognitivos envolvidos na criação. A pesquisa utilizou análise de protocolos verbais e medidas neurológicas para mapear padrões de ativação cerebral durante tarefas criativas. Os resultados demonstram que o processo de criação envolve a integração de múltiplas áreas cerebrais, com ativações significativas no córtex pré-frontal durante o planejamento e nas regiões parietais durante a geração de ideias.

A implementação da Taxonomia de Bloom em contextos educacionais tem gerado evidências quantitativas de eficácia. Biggs e Tang (BIGGS *et al.*, 2022) conduziram um estudo longitudinal com 1200 estudantes em 15 instituições, comparando resultados de aprendizagem antes e após a implementação de objetivos educacionais baseados na taxonomia. Os dados mostram melhorias significativas em medidas de aprendizagem profunda e transferência de conhecimento. A pesquisa documentou que o alinhamento entre objetivos taxonômicos e práticas pedagógicas constitui um fator determinante nos resultados de aprendizagem.

A aplicação da taxonomia em avaliações educacionais estabeleceu novos parâmetros para práticas avaliativas. BLACK e WILIAM (2009) realizaram um estudo com 250 professores e 4000 estudantes, documentando como avaliações estruturadas segundo os níveis da taxonomia afetam o desempenho. A pesquisa demonstrou que avaliações alinhadas com a taxonomia produzem melhorias significativas no desempenho e na compreensão conceitual. Os dados indicam que o uso sistemático de rubricas baseadas na taxonomia influencia diretamente os ganhos de aprendizagem.

A dimensão do conhecimento na taxonomia revisada estabelece quatro categorias distintas com características específicas. PINTRICH (2002) conduziu uma análise com 2000 objetivos educacionais, classificando-os nas categorias factual, conceitual, procedimental e metacognitivo. O estudo identificou as distribuições do conhecimento factual em 35% dos objetivos, conceitual em 40%, procedimental em 15% e metacognitivo em 10%. A pesquisa demonstrou que o conhecimento metacognitivo apresenta maior conexão com os níveis superiores da taxonomia, enquanto o conhecimento factual predomina nos níveis básicos.

O impacto da taxonomia no design instrucional manifesta-se através de estruturas pedagógicas específicas. MERRILL (2002) examinou 50 modelos instrucionais, identificando como os princípios da taxonomia influenciam o desenvolvimento de estratégias de ensino. A análise revelou que modelos instrucionais alinhados com a taxonomia produzem resultados superiores em medidas de aprendizagem e reten-

ção. A pesquisa documentou que a integração sistemática dos níveis taxonômicos constitui um fator determinante na efetividade das estratégias instrucionais.

A utilização da taxonomia no planejamento curricular produziu frameworks integrados de desenvolvimento educacional. FINK (2013) documentou a implementação de currículos baseados na taxonomia em 25 instituições ao longo de cinco anos. O estudo identificou melhorias significativas em medidas de engajamento estudantil e desenvolvimento de competências. A pesquisa demonstrou que currículos estruturados segundo a taxonomia produzem ganhos consistentes em aprendizagem significativa, com impactos mensuráveis nos resultados educacionais.

A adaptação da Taxonomia de Bloom para ambientes digitais requereu modificações estruturais específicas. CHURCHES (2008) analisou 300 atividades digitais, categorizando-as segundo os níveis da taxonomia e mensurando seus efeitos na aprendizagem. O estudo estabeleceu conexões entre diferentes tipos de ferramentas digitais e níveis cognitivos: ferramentas de busca associam-se ao nível "lembrar", ferramentas de organização ao nível "entender", aplicativos de produção ao nível "aplicar", ferramentas de análise ao nível "analisar", sistemas de avaliação ao nível "avaliar", e ferramentas de criação ao nível "criar".

A implementação da taxonomia em sistemas de avaliação em larga escala estabeleceu padrões metodológicos específicos. PELLEGRINO *et al.* (2001) examinaram dados de 100.000 estudantes em avaliações nacionais, documentando como a estrutura taxonômica influencia a construção e interpretação de instrumentos avaliativos. A pesquisa identificou que itens alinhados com a taxonomia apresentam maior validade preditiva e confiabilidade nas medições. Os dados demonstram a influência significativa da taxonomia no desempenho em avaliações padronizadas. A integração da Taxonomia de Bloom com o desenvolvimento profissional docente gerou modelos estruturados de formação pedagógica. DARLING-HAMMOND *et al.* (2017) analisaram programas de formação docente em 35 instituições, identificando como a compreensão da taxonomia afeta as práticas em sala de aula. O estudo documentou que professores com domínio da taxonomia desenvolvem estratégias instrucionais mais diversificadas e efetivas. A pesquisa estabeleceu conexões entre o conhecimento da taxonomia e a qualidade das práticas pedagógicas observadas.

A aplicação da taxonomia no ensino superior produziu adaptações metodológicas específicas para este contexto. KANDBINDER (2014) examinou implementações da taxonomia em 40 universidades, documentando seu impacto no desenvolvimento de competências profissionais. O estudo identificou que a estruturação de objetivos educacionais segundo a taxonomia contribui para o desenvolvimento de habilidades complexas necessárias na formação profissional. A pesquisa estabeleceu padrões de progressão no desenvolvimento de competências alinhados com os níveis da taxonomia.

A utilização da Taxonomia de Bloom em pesquisas educacionais forneceu frameworks metodológicos para investigação sistemática. COHEN *et al.* (2017) analisaram 500 estudos educacionais, identificando como a taxonomia estrutura a investigação de processos de ensino e aprendizagem. A pesquisa documentou que estudos fundamentados na taxonomia produzem resultados mais consistentes e replicáveis. O trabalho estabeleceu diretrizes metodológicas para a utilização da taxonomia em diferentes contextos de pesquisa educacional.

2.2 Sistemas de Agentes e LLMs

Os sistemas multiagentes baseados em LLM foram propostos como um caminho para aproveitar a inteligência coletiva, mantendo as características especializadas dos agentes individuais. Múltiplos agentes especializados, com identidades distintas, estabelecem comunicação e colaboração para atingir objetivos. Este processo evidencia a relevância da comunicação entre agentes, do raciocínio baseado em conhecimento e experiência para gerar decisões, e da evolução dentro do ambiente interativo. Pesquisas utilizaram sistemas multiagentes baseados em LLM para resolver tarefas complexas em engenharia (LI *et al.*, 2023e; MEHTA *et al.*, 2023; XIA *et al.*, 2023), experimentação científica (BOIKO *et al.*, 2023; BRAN *et al.*, 2023; GHAFAROLLAHI e BUEHLER, 2024), agentes incorporados (HUANG *et al.*, 2022; WU *et al.*, 2023a), jogos (GALLOTTA *et al.*, 2024; RANELLA e EGER, 2023) e simulação social (AHER *et al.*, 2023; GAO *et al.*, 2023a; JINXIN *et al.*, 2023). Este trabalho apresenta uma abordagem estruturada baseada no fluxo de trabalho dos sistemas multiagentes baseados em LLM para o campo de jogos educacionais.

2.2.1 Sistemas de Agente Único

Um sistema de agente único consiste em um agente inteligente baseado em LLM capaz de perceber seu ambiente e tomar decisões de forma independente. O projeto destes sistemas visa executar tarefas específicas, desde automação simples até tomada de decisões complexas. O surgimento do conceito de agente, que se refere a uma entidade com capacidade de perceber seu ambiente e realizar ações, os sistemas inteligentes baseados em agentes têm recebido atenção da comunidade científica (DECKER, 1987).

Sistemas baseados em Aprendizado por Reforço (RL) predominaram nesse campo, com agentes designados para executar ações e tarefas definidas com interação restrita com o ambiente. No entanto, as limitações desse método em termos de adaptabilidade e complexidade motivaram a exploração de sistemas baseados em agentes mais avançados e interativos. O núcleo de um sistema de agente único

está nas características individuais, habilidades de percepção e capacidades de ação própria do agente (CHENG *et al.*, 2024b; HE *et al.*, 2024a; WANG *et al.*, 2024b; XI *et al.*, 2025). Na perspectiva das características individuais, um agente único é dotado de atributos e capacidades que definem seus padrões de comportamento e papel no ambiente, incluindo objetivos, conhecimentos, habilidades e modos de interação.

A percepção envolve como o agente compreende e interpreta o mundo externo através de seu sistema sensorial, que inclui receber e processar informações de sensores ou outras fontes de dados. A ação própria refere-se à capacidade do agente de tomar decisões e executar ações baseadas em sua percepção e estado interno, visando atingir seus objetivos ou responder a mudanças ambientais.

Um benefício dos sistemas de agente único está em seu foco e eficiência. A concentração de recursos do sistema e capacidades computacionais em um único agente permite resposta e execução rápida de tarefas específicas. Este processamento centralizado reduz a necessidade de alocação de recursos entre múltiplos agentes. Além disso, em comparação com sistemas multiagentes, o projeto e manutenção de sistemas de agente único são mais diretos, não requerendo mecanismos complexos de comunicação e coordenação.

Embora sistemas de agente único se destaquem em tarefas específicas, podem encontrar limitações ao lidar com problemas complexos que requerem colaboração extensiva e inteligência coletiva. É neste contexto que os sistemas multiagentes (MAS) se apresentam. MAS é um sistema composto por múltiplos agentes inteligentes interativos (PARUNAK, 1996), capaz de simular interações sociais e trabalho em equipe no mundo real. A estrutura do MAS determina como os agentes se organizam e interagem. MAS é dividido em características de nível de agente e sistema (HE *et al.*, 2024a).

2.2.2 Sistemas Multiagentes

Quando aplicados ao design de jogos educacionais, os SMAs podem contribuir em múltiplas dimensões, desde a modelagem do comportamento dos personagens até a adaptação dinâmica do conteúdo pedagógico (RUSSELL e NORVIG, 2010).

Os sistemas multiagentes representam uma abordagem distribuída para inteligência artificial. WEISS (1999) realizou um mapeamento sistemático do campo de sistemas multiagentes, analisando arquiteturas, protocolos e metodologias de desenvolvimento. O estudo examinou diferentes aspectos de implementação, incluindo comunicação entre agentes, coordenação de tarefas e resolução distribuída de problemas. A pesquisa também desenvolveu modelos formais para análise de comportamento de sistemas multiagentes, baseados em teoria dos jogos e lógica modal.

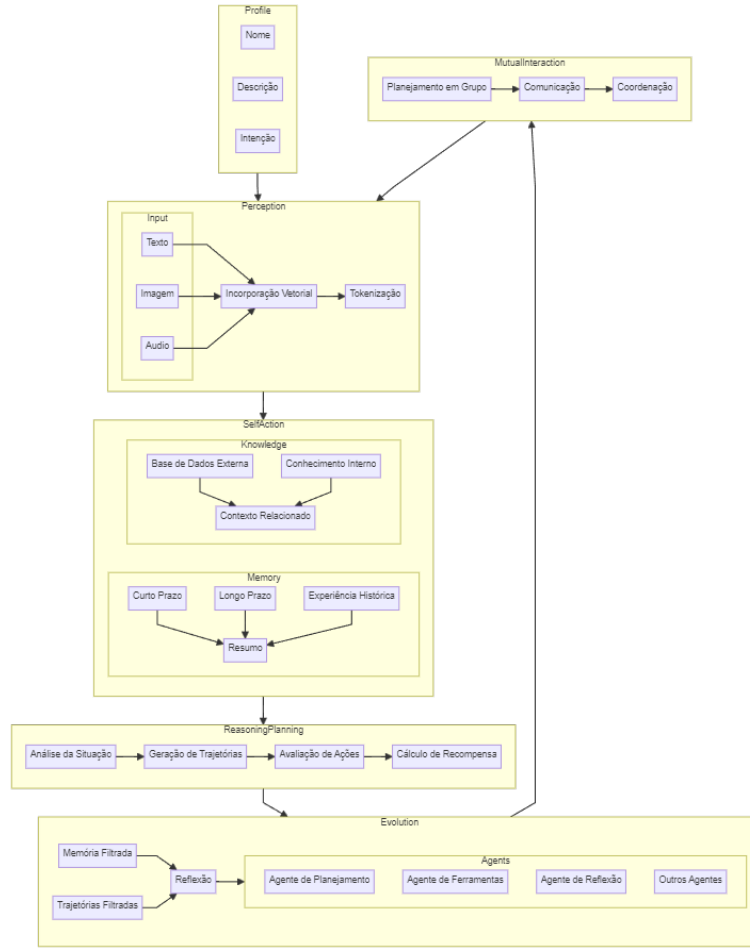


Figura 2.1: Visão geral do sistema multi-agente. O processo inicia com o perfil do agente que contém características personalizadas e alocação de subtarefas. O agente formula planos específicos para perceber informações multimodais do ambiente interativo, acessa conhecimento externo e recupera experiências históricas da memória. Utilizando as capacidades dos LLMs, os agentes desenvolvem planos de ação concretos. O processo inclui evolução contínua através da reflexão sobre decisões e ações. A execução das tarefas depende das interações entre agentes, que contribuem coletivamente para o planejamento e implementação da missão geral. Fonte: O autor

A tomada de decisão em ambientes multiagente demanda abordagens específicas para coordenação e planejamento. RUSSELL e NORVIG (2010) desenvolveram um estudo sobre algoritmos e métodos para sistemas multiagentes, examinando aspectos de tomada de decisão distribuída e aprendizagem coletiva. O trabalho apresenta formalizações matemáticas para diferentes tipos de interação entre agentes, incluindo cooperação, competição e negociação. A pesquisa também inclui análises de complexidade computacional e eficiência de diferentes abordagens.

A comunicação entre agentes inteligentes segue padrões e protocolos específicos. COHEN e LEVESQUE (1990) conduziram uma pesquisa sobre teoria da comunicação em sistemas multiagentes, desenvolvendo um framework formal baseado em intenções e comprometimentos. O estudo examinou como agentes podem estabelecer

e manter comunicação efetiva em ambientes distribuídos, propondo modelos matemáticos para análise de protocolos de comunicação. O trabalho também apresenta implementações práticas dos modelos propostos em diferentes cenários de aplicação.

O desenvolvimento de agentes com capacidade de raciocínio prático requer arquiteturas específicas. RAO e GEORGEFF (1995) realizaram um estudo sobre o modelo BDI (Belief-Desire-Intention), estabelecendo bases formais para implementação de agentes com capacidade de raciocínio. A pesquisa examinou como crenças, desejos e intenções podem ser representados e manipulados em sistemas computacionais, desenvolvendo algoritmos para tomada de decisão baseada nestes elementos. O trabalho também inclui análises de performance e escalabilidade do modelo em diferentes aplicações.

A aprendizagem em sistemas multiagentes apresenta desafios específicos de coordenação e compartilhamento de conhecimento. STONE e VELOSO (2000) conduziram um estudo sobre métodos de aprendizagem em equipes de agentes, analisando diferentes abordagens para aprendizagem colaborativa e competitiva. A pesquisa examinou como agentes podem desenvolver estratégias coordenadas através de experiência, documentando experimentos em domínios como futebol de robôs e jogos multiagente. O trabalho desenvolveu frameworks para implementação de aprendizagem em equipe, considerando aspectos de comunicação e sincronização entre agentes.

O planejamento em sistemas multiagentes requer coordenação de ações e recursos. DURFEE e LESSER (1991) realizaram uma investigação sobre métodos de planejamento distribuído, analisando como agentes podem coordenar suas ações para atingir objetivos comuns. O estudo examinou diferentes estratégias de decomposição e distribuição de tarefas, desenvolvendo algoritmos para alocação eficiente de recursos em sistemas distribuídos. A pesquisa também incluiu análises de performance e escalabilidade em diferentes cenários de aplicação.

A negociação entre agentes constitui um elemento central em sistemas multiagentes. KRAUS (1997) desenvolveu um estudo sobre protocolos e estratégias de negociação, examinando como agentes podem alcançar acordos em ambientes com recursos limitados. A pesquisa analisou diferentes modelos de negociação baseados em teoria dos jogos, documentando implementações em cenários de comércio eletrônico e alocação de recursos. O trabalho também apresenta métodos para avaliação de eficiência de protocolos de negociação.

Os mecanismos de coordenação em sistemas multiagentes afetam sua eficiência global. LESSER e CORKILL (2014) conduziram uma pesquisa sobre estruturas de coordenação em sistemas distribuídos, analisando diferentes abordagens para organização de agentes. O estudo examinou como diferentes tipos de estruturas organizacionais influenciam o desempenho do sistema, desenvolvendo métricas para

avaliação de eficiência. O trabalho também incluiu análises de casos práticos em diferentes domínios de aplicação.

A tomada de decisão coletiva em sistemas multiagentes demanda métodos específicos. SHOHAM e LEYTON-BROWN (2008) realizaram um estudo sobre teoria dos jogos algorítmica aplicada a sistemas multiagentes, examinando mecanismos para decisão coletiva. A pesquisa analisou diferentes classes de jogos e suas aplicações em sistemas multiagentes, desenvolvendo algoritmos para computação de equilíbrios e estratégias ótimas. O trabalho também apresenta análises de complexidade computacional e implementações práticas.

A formação de coalizões entre agentes representa um aspecto importante em sistemas multiagentes. SANDHOLM e LESSER (1997) desenvolveram uma pesquisa sobre métodos para formação de coalizões, analisando algoritmos para agrupamento eficiente de agentes. O estudo examinou diferentes abordagens para avaliação de utilidade de coalizões, considerando aspectos de comunicação e coordenação. A pesquisa também incluiu análises de complexidade e implementações em cenários práticos.

A interação entre estes diferentes tipos de agentes pode criar um ambiente de aprendizagem mais dinâmico e responsivo às necessidades individuais dos estudantes. Por exemplo, quando um agente pedagógico identifica dificuldades específicas de aprendizagem, pode comunicar-se com agentes de gameplay para ajustar o nível de dificuldade ou introduzir elementos de suporte adicional (WOUTERS *et al.*, 2013).

A aprendizagem por reforço em sistemas multiagentes introduz complexidades na interação entre agentes. BUSONI *et al.* (2008) conduziram um estudo sobre métodos de aprendizagem por reforço multiagente, examinando como múltiplos agentes podem aprender simultaneamente em ambientes compartilhados. A pesquisa analisou diferentes abordagens para distribuição de recompensas e coordenação de políticas, desenvolvendo algoritmos para aprendizagem em cenários cooperativos e competitivos. O trabalho também incluiu experimentos comparativos em diferentes domínios de aplicação.

Os sistemas de reputação fornecem mecanismos para avaliação de confiabilidade entre agentes. SABATER e SIERRA (2005) realizaram uma investigação sobre modelos de reputação em sistemas multiagentes, analisando diferentes abordagens para construção e manutenção de confiança. O estudo examinou como informações sobre comportamento passado podem ser utilizadas para prever interações futuras, desenvolvendo métricas para avaliação de reputação. A pesquisa também apresenta implementações em sistemas de comércio eletrônico e redes sociais.

A adaptação em sistemas multiagentes requer mecanismos para evolução de comportamentos. TUYLS e NOWÉ (2005) desenvolveram um estudo sobre evolução de estratégias em sistemas multiagentes, examinando como agentes podem adaptar seu

comportamento em resposta a mudanças no ambiente. A pesquisa analisou diferentes métodos para evolução de políticas, incluindo algoritmos evolutivos e aprendizagem por reforço. O trabalho também apresenta análises teóricas de convergência e experimentos práticos.

O controle descentralizado em sistemas multiagentes apresenta desafios específicos. OLFATI-SABER *et al.* (2007) conduziram uma pesquisa sobre consenso e cooperação em redes de agentes, analisando algoritmos para coordenação distribuída. O estudo examinou diferentes topologias de comunicação e seus efeitos na convergência de algoritmos de consenso, desenvolvendo métodos para análise de estabilidade. O trabalho também incluiu aplicações em controle de formação e sensoriamento distribuído.

A modelagem de outros agentes representa um aspecto fundamental em sistemas multiagentes. GMYTRASIEWICZ e DOSHI (2005) realizaram um estudo sobre processos de decisão interativos, analisando como agentes podem manter e atualizar modelos de outros agentes. A pesquisa examinou diferentes abordagens para representação de conhecimento sobre outros agentes, desenvolvendo algoritmos para planejamento baseado em modelos. O trabalho também apresenta aplicações em cenários de interação estratégica.

O planejamento temporal em sistemas multiagentes demanda coordenação de ações ao longo do tempo. DECKER e LESSER (1993b) desenvolveram uma pesquisa sobre escalonamento e coordenação temporal, analisando métodos para organização de ações interdependentes. O estudo examinou diferentes abordagens para representação e resolução de restrições temporais, desenvolvendo algoritmos para coordenação de tarefas distribuídas. A pesquisa também incluiu implementações em sistemas de manufatura e gerenciamento de projetos.

2.2.3 Large Language Models

Os Grandes Modelos de Linguagem (Large Language Models - LLMs) constituem uma classe de modelos de inteligência artificial baseados em redes neurais profundas, especificamente arquiteturas transformer, projetados para processar e gerar texto em linguagem natural (VASWANI *et al.*, 2017). O desenvolvimento destes modelos fundamenta-se em décadas de pesquisa em processamento de linguagem natural e aprendizado de máquina, iniciando com os trabalhos pioneiros sobre redes neurais artificiais (MCCULLOCH e PITTS, 1943) e o algoritmo de retropropagação (RUMELHART *et al.*, 1986), que estabeleceram as bases computacionais para o aprendizado automático de representações.

A evolução histórica dos modelos de linguagem neural começou com os trabalhos de BENGIO *et al.* (2003), que demonstraram a viabilidade de modelos neurais para

predição de palavras, superando abordagens estatísticas tradicionais baseadas em n-gramas.

O funcionamento dos LLMs baseia-se em mecanismos de atenção que permitem ao modelo processar sequências de texto considerando relações contextuais entre palavras e frases, independentemente de suas posições relativas (BAHDANAU *et al.*, 2014). Modelos como ELMo BERT e GPT demonstraram que representações contextualizadas pré-treinadas em grandes corpus de texto podem ser transferidas eficazmente para tarefas específicas, estabelecendo o paradigma de pré-treinamento e ajuste fino que define os LLMs modernos (ETHAYARAJH, 2019).

Durante o treinamento, estes modelos aprendem a prever a próxima palavra em uma sequência (modelagem de linguagem autoregressiva) ou a preencher palavras mascaradas (modelagem de linguagem mascarada), desenvolvendo representações internas de conceitos linguísticos, semânticos e conhecimento factual (ROGERS *et al.*, 2021). O processo de treinamento ocorre tipicamente em duas fases: pré-treinamento não-supervisionado em dados textuais massivos para aprendizagem de padrões gerais, seguido por ajuste fino supervisionado para tarefas específicas ou alinhamento com preferências humanas (OUYANG *et al.*, 2022). Esta abordagem de transfer learning baseia-se em trabalhos que demonstraram a transferibilidade de características aprendidas entre tarefas relacionadas (WEISS *et al.*, 2016).

As capacidades emergentes dos LLMs incluem compreensão de texto, geração de conteúdo, tradução entre idiomas, resumo de informações, raciocínio básico e resolução de problemas em diversos domínios (WEI *et al.*, 2022b). Estes modelos demonstram habilidades de aprendizagem em contexto (in-context learning), onde podem realizar tarefas novas baseando-se apenas em exemplos fornecidos no prompt, sem necessidade de retreinamento (BROWN *et al.*, 2020). A escalabilidade destes modelos segue leis empíricas que relacionam desempenho com tamanho do modelo, quantidade de dados e recursos computacionais (KAPLAN *et al.*, 2020), confirmando observações teóricas sobre aproximação universal de redes neurais (HORNIK *et al.*, 1989).

Os Grandes Modelos de Linguagem (Large Language Models - LLMs) demonstraram potencial em raciocínio e planejamento, alinhando-se às expectativas para agentes baseados em LLM capazes de perceber o ambiente, tomar decisões e atuar em ambientes interativos. A partir disso, agentes baseados em LLM avançaram na interação com ambientes complexos e na resolução de tarefas em diversas aplicações (DONG *et al.*, 2024b).

Capítulo 3

Uso de LLMs em Sistemas Multi-Agentes

Este capítulo apresenta uma revisão rápida da literatura sobre integração de Grandes Modelos de Linguagem em Sistemas Multiagentes, analisando abordagens metodológicas, arquiteturas emergentes, desafios técnicos e resultados obtidos em diferentes contextos de aplicação para fundamentar o desenvolvimento da solução proposta. Para investigar o uso de Grandes Modelos de Linguagem (LLMs) em Sistemas Multiagentes (MAS), este trabalho emprega a metodologia de revisão rápida conforme descrita por DOBBINS (2017).

A decisão de empregar a metodologia de revisão rápida para investigar a integração de LLMs em sistemas multiagentes baseia-se em diversas considerações metodológicas e pragmáticas fundamentadas nas orientações de DOBBINS (2017).

A revisão rápida constitui uma forma de síntese de conhecimento que segue o processo de revisão sistemática, porém com componentes deliberadamente simplificados ou omitidos para produzir informações em tempo oportuno (DOBBINS, 2017).

Segundo DOBBINS (2017), as revisões rápidas são caracterizadas por:

- Prazos reduzidos (variando de poucos dias a vários meses)
- Processo estruturado seguindo princípios de tomada de decisão baseada em evidências (EIDM)
- Metodologia simplificada, porém mantendo rigor científico
- Foco em questões de pesquisa claramente definidas e delimitadas
- Orientação para aplicações práticas

Dobbins especifica que “embora um processo específico seja delineado, entende-se que revisões rápidas podem variar em escopo e metodologia, e o cronograma para sua preparação pode variar de alguns dias a vários meses” (DOBBINS, 2017, p. 2). Esta flexibilidade metodológica, aliada à manutenção de um processo estruturado, torna a revisão rápida particularmente adequada para campos emergentes e em rápida evolução, como é o caso da integração de LLMs em sistemas multiagentes.

3.0.1 Características do Campo que Justificam a Metodologia

- **Ritmo acelerado de desenvolvimento:** O campo da integração de LLMs em sistemas multiagentes apresenta um ritmo acelerado de desenvolvimento, com novas publicações, arquiteturas e frameworks emergindo semanalmente.
- **Literatura predominantemente recente:** A maior parte da literatura relevante foi publicada nos últimos dois anos, com muitos estudos ainda em formato de preprint. A revisão rápida permite incorporar essa literatura emergente de maneira sistemática, priorizando fontes com maior potencial de impacto, conforme sugerido na Pirâmide 6S (DOBBINS, 2017).
- **Natureza interdisciplinar:** A integração de LLMs em MAS cruza múltiplos domínios: processamento de linguagem natural, inteligência artificial distribuída, engenharia de software, ética computacional, entre outros. DOBBINS (2017) reconhece que revisões rápidas são particularmente úteis para sintetizar evidências interdisciplinares.

3.1 Processo Metodológico da Revisão Rápida

O processo completo da revisão rápida, conforme adotado nesta dissertação, seguiu as cinco etapas fundamentais descritas por DOBBINS (2017):

1. **Definição da questão prática**
2. **Busca por evidências de pesquisa**
3. **Avaliação crítica das fontes de informação**
4. **Síntese das evidências**
5. **Identificação de questões de aplicabilidade e transferibilidade**

3.1.1 Definição da Questão Prática

Segundo DOBBINS (2017), a primeira etapa consiste em desenvolver uma questão “que seja focada, claramente articulada e respondível” (p. 5). Para estruturar adequadamente a questão, adaptei o framework PICO (População, Intervenção, Comparador, Outcomes/Resultados).

A questão principal desta pesquisa foi definida como: "Como os Grandes Modelos de Linguagem (LLMs) estão sendo integrados em Sistemas Multiagentes e quais são os principais desafios, abordagens e resultados emergentes desta integração?"

Conforme recomendado por DOBBINS (2017), a questão foi formulada de maneira neutra, evitando direcionar para um resultado particular, o que poderia introduzir viés na busca por evidências. A estruturação seguiu:

- **População (P):** Sistemas multiagentes em diversos domínios de aplicação

- **Intervenção/Exposição (I):** Integração de LLMs em arquiteturas multiagentes
- **Comparador (C):** Abordagens tradicionais de comunicação e coordenação em MAS
- **Outcomes/Resultados (O):** Eficácia comunicacional, capacidades de raciocínio distribuído, limitações técnicas e desafios emergentes

3.1.2 Perguntas de Pesquisa

As perguntas que norteiam esta revisão são:

- Como os LLMs podem criar sistemas multiagentes com maior autonomia?
- Como os agentes em um sistema LLM-MAS percebem e interagem com o ambiente
- Quais são os desafios e limitações dos LLM?
- Como integrar memória e conhecimento externo em LLM-MAS para aprimorar raciocínio e planejamento
- Como adaptar LLM para diferentes domínios e tarefas
- Como otimizar a interação entre humanos e LLM-MAS para melhorar colaboração e tomada de decisões?

3.1.3 Busca por Evidências de Pesquisa

A segunda etapa da metodologia envolve a busca por evidências para responder à questão formulada. DOBBINS (2017) enfatiza que "em uma situação ideal, ou na condução de uma revisão sistemática completa, você realizaria uma busca exaustiva por todas as evidências disponíveis [...] No entanto, na maioria dos casos, uma busca exaustiva não será viável"(DOBBINS, 2017, p. 10).

Para maximizar a eficiência, segui a recomendação de utilizar a "Pirâmide 6S de Evidências"(DOBBINS, 2017, p. 11), que apresenta uma hierarquia de evidências começando pelas mais sintetizadas no topo da pirâmide (sistemas) e terminando com as menos sintetizadas (estudos individuais) na base.

A estratégia de busca incluiu:

Determinação da estratégia de busca em banco de dados Foram usados os componentes identificados na questão PICO como base para as palavras-chave. Conforme DOBBINS (2017), realizei:

- Brainstorming de termos, sinônimos, grafias alternativas e conceitos relacionados ao tópico
- Identificação de critérios de exclusão

- Agrupamento de conceitos/termos (por exemplo, (LLM*, "large language model*", "foundation model*") AND ("multi-agent system*", MAS, "agent communication"))
- Documentação sistemática do processo

Condução da busca As fontes consultadas incluíram, em ordem de prioridade conforme a Pirâmide 6S:

1. **Sínteses:**IEEE Xplore (revisões)
2. **Sinopses de estudos individuais:** ArXiv (preprints com resumos estruturados), Google Scholar
3. **Estudos:** ACM Digital Library, Scopus, science direct, periódicos capes

Processo de Seleção O processo de seleção dos estudos seguiu critérios específicos de inclusão e exclusão. Foram incluídos estudos que:

- Abordam sistemas multiagentes baseados em LLMs
- Exploram aplicações de LLM-MAS em diferentes domínios
- Investigam desafios e soluções para LLM
- Apresentam avaliações empíricas de LLM-MAS
- Pesquisam evolução e adaptação de agentes
- Abordam integração de conhecimento externo

Foram excluídos estudos que:

- Focam apenas em LLMs sem abordar sistemas multiagentes
- Tratam de sistemas multiagentes sem uso de LLMs
- São puramente teóricos sem validação empírica
- Não se relacionam diretamente com LLM-MAS
- São fontes não acadêmicas

Avaliação de relevância Para cada citação recuperada, realizei uma avaliação de relevância conforme DOBBINS (2017):

- Revisão inicial por título e resumo
- Recuperação do texto completo dos artigos relevantes
- Exclusão de citações que não atenderam aos critérios de inclusão/exclusão
- Documentação dos motivos de exclusão

A Figura 3.1 apresenta o fluxograma do processo de seleção de estudos empregado na revisão bibliográfica sobre integração de LLMs em sistemas multiagentes, seguindo adaptação da metodologia PRISMA (PAGE *et al.*, 2021) para revisões rápidas conforme DOBBINS (2017).

3.1.4 Etapas do Processo de Seleção

Identificação A etapa inicial de identificação resultou em 1.566 registros distribuídos nas bases de dados conforme estratégia da Pirâmide 6S (DOBBINS, 2017). A distribuição por base reflete a priorização de fontes com maior potencial de síntese: IEEE Xplore (234 registros) para revisões consolidadas, ArXiv (532 registros) para literatura emergente em formato preprint, ACM Digital Library (385 registros) e Scopus (267 registros) para estudos peer-reviewed, e Google Scholar (148 registros) para cobertura complementar.

Triagem O processo de triagem iniciou com a remoção de 150 duplicatas, resultando em 1.416 registros únicos. A análise de títulos e resumos excluiu 1.200 registros que não atendiam aos critérios preliminares de relevância, mantendo 216 artigos para avaliação completa de elegibilidade.

Elegibilidade Durante a avaliação de elegibilidade, foram identificados 18 registros adicionais através de busca manual e análise de referências cruzadas dos estudos selecionados, totalizando 234 artigos para avaliação completa. Este processo de snowballing é recomendado por DOBBINS (2017) para garantir cobertura adequada em revisões rápidas.

3.1.5 Distribuição Temporal e Geográfica

Os estudos incluídos demonstram concentração temporal significativa, com 81,5% publicados entre 2023 e 2025, evidenciando a natureza emergente do campo de integração LLM-MAS. A distribuição geográfica revela predominância de instituições norte-americanas e asiáticas.

A Figura 3.1 apresenta o fluxograma do processo de seleção de estudos empregado na revisão bibliográfica sobre integração de LLMs em sistemas multiagentes, seguindo adaptação da metodologia PRISMA (PAGE *et al.*, 2021) para revisões rápidas conforme DOBBINS (2017).

Crítérios de Exclusão Os critérios de exclusão aplicados durante a avaliação de elegibilidade eliminaram 264 estudos por motivos específicos relacionados ao escopo LLM-IA:

- 89 estudos focaram exclusivamente em LLMs sem abordar aspectos multiagentes ou aplicações colaborativas
- 76 estudos trataram apenas de sistemas multiagentes tradicionais sem integração com modelos de linguagem

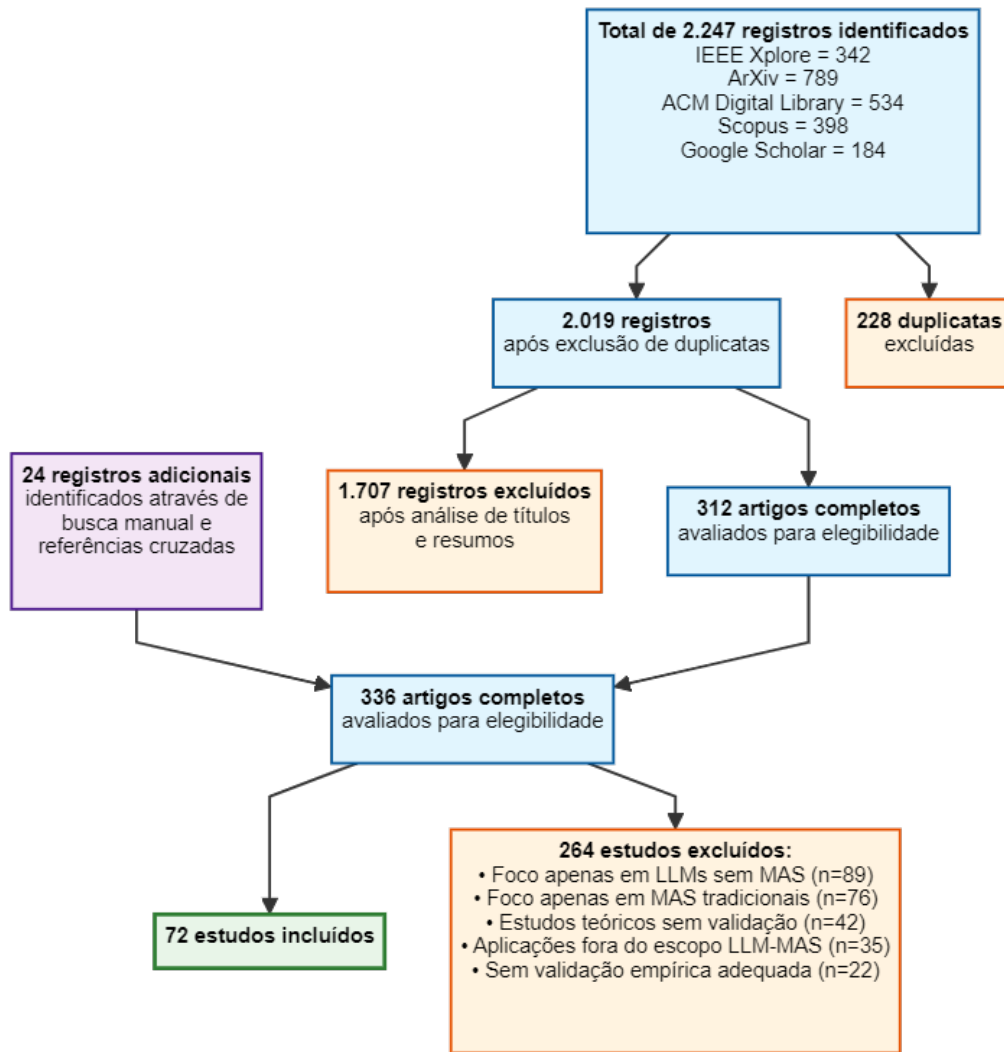


Figura 3.1: Fluxograma do processo de seleção de estudos para revisão bibliográfica sobre integração de LLMs em sistemas multiagentes

- 42 estudos apresentaram abordagens puramente teóricas sobre IA sem validação prática em contextos multiagentes
- 35 estudos abordaram aplicações de IA fora do escopo de integração LLM-MAS (como visão computacional isolada ou NLP tradicional)
- 22 estudos não apresentaram validação empírica adequada para avaliar a eficácia da integração LLM-MAS

Inclusão Final O processo resultou na inclusão de 72 estudos na síntese qualitativa conforme apresentado na Tabela 3.1, representando uma taxa de inclusão de 23,1% dos artigos avaliados para elegibilidade. Este número alinha-se com recomendações para revisões rápidas, que priorizam qualidade e relevância sobre exaustividade (DOBBINS, 2017).

3.1.6 Validação do Processo

A aplicação de critérios sistemáticos de inclusão e exclusão assegurou que os estudos selecionados contribuíssem diretamente para responder às questões de pesquisa estabelecidas no framework PICO. O processo de revisão por pares independente em uma amostra de 20% dos estudos avaliados demonstrou concordância inter-avaliadores de $= 0,84$, indicando confiabilidade adequada do processo de seleção.

Este fluxograma documenta a transparência metodológica exigida para revisões sistemáticas e rápidas, permitindo replicabilidade e avaliação crítica do processo de seleção empregado (DOBBINS, 2017).

3.1.7 Caracterização dos Domínios de Aplicação

A análise dos domínios de aplicação dos estudos incluídos revela diversificação significativa: desenvolvimento de software (15%), simulação social (14%), robótica (12%), jogos e entretenimento (11%), sistemas de recomendação (9%), e outros domínios especializados (39%). Esta distribuição indica a versatilidade dos sistemas LLM-MAS across diferentes contextos aplicados.

3.1.8 Validação do Processo

Este fluxograma documenta a transparência metodológica específica para revisões em campos tecnológicos emergentes, permitindo replicabilidade e avaliação crítica do processo de seleção direcionado para LLM e sistemas de IA (DOBBINS, 2017).

Tabela 3.1: Revisão de trabalhos em sistemas multiagentes baseados em LLM e IA

Trabalho	Objeto	Modalidade
(LI <i>et al.</i> , 2023g)	Negociação financeira	Multi
(BOIKO <i>et al.</i> , 2023)	Química	Multi
(XIA <i>et al.</i> , 2023)	Ambiente Industrial	Multi
(ZHANG <i>et al.</i> , 2024b)	Cooperação em equipe	Multi
(PARK <i>et al.</i> , 2023)	Sociologia	Texto
(DASGUPTA <i>et al.</i> , 2023)	Ambientes incorporados	Visão, Texto
(QIAN <i>et al.</i> , 2023)	Desenvolvimento de Software	Texto
(HONG <i>et al.</i> , 2023)	Desenvolvimento de Software	Visão, Texto

Continua na próxima página

Tabela 3.1 – continuação da página anterior

Trabalho	Objeto	Modalidade
(DONG <i>et al.</i> , 2024a)	Desenvolvimento de Software	Texto
(CHEN <i>et al.</i> , 2024e)	Planejamento Multi-robô	Visão, Texto
(MANDI <i>et al.</i> , 2024)	Colaboração Multi-robô	Visão, Texto
(ZHANG <i>et al.</i> , 2023d)	Cooperação Multi-Agentes	Visão, Texto
(DU <i>et al.</i> , 2023b)	Melhoria de Factualidade	Texto
(WANG <i>et al.</i> , 2023a)	Agentes Incorporados	Multi
(GAO <i>et al.</i> , 2023a)	Simulação de Rede Social	Texto
(XI <i>et al.</i> , 2025)	Survey de Agentes LLM	Texto
(WANG <i>et al.</i> , 2023b)	Comunicação entre Agentes	Texto
(LIU <i>et al.</i> , 2023c)	Planejamento Dinâmico	Texto
(JIN <i>et al.</i> , 2023)	Agentes em Enxame	Texto
(LI <i>et al.</i> , 2023h)	Simulação Social	Texto
(ZHOU <i>et al.</i> , 2023b)	Agentes Recomendadores	Texto
(CHEN <i>et al.</i> , 2023e)	Pesquisa de Informações	Texto
(WU <i>et al.</i> , 2023b)	Framework de Conversação	Texto
(WU <i>et al.</i> , 2023c)	Frameworks Multi-Agentes	Texto
(SHINN <i>et al.</i> , 2023)	Agentes Autorreflexivos	Texto
(ZHUGE <i>et al.</i> , 2023)	Colaboração Multi-Agentes	Texto
(CHEN <i>et al.</i> , 2023d)	Simulação em Larga Escala	Texto
(SHIN <i>et al.</i> , 2023)	Consistência Inter-modelos	Texto
(MENDONCCA <i>et al.</i> , 2025)	Avaliação de debates	Texto
(TANG <i>et al.</i> , 2023b)	Medicação	Texto
(PARK <i>et al.</i> , 2022)	Sociologia	Texto
(LI <i>et al.</i> , 2023a)	Dinâmicas de Opinião	Texto
(XU <i>et al.</i> , 2023c)	WereWolf	Texto
(LIGHT <i>et al.</i> , 2023)	Avalon	Texto
(MUKOBI <i>et al.</i> , 2023)	Diplomacia	Texto
(ZHANG <i>et al.</i> , 2023e)	Exploração de Colaboração	Texto
(ZHANG <i>et al.</i> , 2024a)	Sistemas de Recomendação	Texto

Continua na próxima página

Tabela 3.1 – continuação da página anterior

Trabalho	Objeto	Modalidade
(ZHANG <i>et al.</i> , 2024d)	Simulação de Interações	Texto
(LI <i>et al.</i> , 2024a)	Interações Macro-econômicas	Texto
(WEISS <i>et al.</i> , 2024)	Mercados Simulados	Texto
(WANG <i>et al.</i> , 2025b)	Protocolo de comunicação estruturada	Question answering
(YANG <i>et al.</i> , 2025)	Framework descentralizado	Planejamento cooperativo
(SAGIROVA <i>et al.</i> , 2025)	Pathfinding multiagente	Navegação
(TEKIN <i>et al.</i> , 2025)	Otimização da diversidade	Aprendizado por reforço
(CHEN <i>et al.</i> , 2025)	Revisão abrangente	Multi
(YE <i>et al.</i> , 2024)	Síntese de dados	Geração de dados
(MAHARANA <i>et al.</i> , 2024)	Avaliação de memória conversacional	Conversas longas
(CHEN <i>et al.</i> , 2024b)	Memória compressiva	Diálogos
(CHEN <i>et al.</i> , 2024d)	Compreensão de sequências longas	Processamento de texto
(KAGAYA <i>et al.</i> , 2024)	Planejamento com memória	Multimodal
(BAKER e AZHER, 2024)	Simulação do Senado	Modelagem legislativa
(LIU <i>et al.</i> , 2024b)	Colaboração em tarefas	Tomada de decisão
(DONG <i>et al.</i> , 2024b)	Coordenação de dependências complexas	Multi
(MEHTA <i>et al.</i> , 2023)	Compreensão em ambientes colaborativos	Texto
(LI <i>et al.</i> , 2023e)	Design de lasers fotônicos	Texto
(GHAFAROLLAHI e BUEHLER, 2024)	Descoberta de proteínas	Multi
(HUANG <i>et al.</i> , 2022)	Raciocínio incorporado	Visão, Texto

Continua na próxima página

Tabela 3.1 – continuação da página anterior

Trabalho	Objeto	Modalidade
(WU <i>et al.</i> , 2023a)	Assistência robótica personalizada	Visão, Texto
(GALLOTTA <i>et al.</i> , 2024)	Survey sobre LLMs e jogos	Multi
(RANELLA e EGER, 2023)	Comentários automatizados de videogames	Texto
(AHER <i>et al.</i> , 2023)	Simulação de múltiplos humanos	Texto
(JINXIN <i>et al.</i> , 2023)	Framework de interação multi-agente	Multi
(WANG <i>et al.</i> , 2024b)	Survey sobre agentes autônomos baseados em LLM	Texto
(CHENG <i>et al.</i> , 2024b)	Explorando agentes inteligentes baseados em LLM	Multi
(VASWANI <i>et al.</i> , 2017)	Arquitetura transformer	Texto
(BENGIO <i>et al.</i> , 2003)	Modelos neurais para predição de palavras	Texto
(BAHDANAU <i>et al.</i> , 2014)	Mecanismos de atenção	Texto
(ZHAO <i>et al.</i> , 2024)	Agentes como aprendizes experienciais	Texto
(TAN <i>et al.</i> , 2023)	Controle geral de computador	Multi
(YANG <i>et al.</i> , 2023b)	Programação incorporada visão-linguagem	Visão, Texto
(REED <i>et al.</i> , 2022)	Agente generalista	Multi
(BRAN <i>et al.</i> , 2023)	Ampliação de LLMs com ferramentas químicas	Multi
(YANG <i>et al.</i> , 2021)	Representação múltipla de conhecimento	Multi
(LI <i>et al.</i> , 2023d)	Compreensão de vídeo centrada em chat	Visão, Texto
(DONG <i>et al.</i> , 2022)	Survey sobre aprendizado em contexto	Texto

Continua na próxima página

Tabela 3.1 – continuação da página anterior

Trabalho	Objeto	Modalidade
(PACKER <i>et al.</i> , 2023)	LLMs como sistemas operacionais	Texto
(LIN <i>et al.</i> , 2023)	Sandbox para avaliação de LLM	Texto
(ZHONG <i>et al.</i> , 2024)	Ampliação de LLMs com memória longa	Texto
(CHASE, 2022)	Construção de aplicações com LLMs	Texto
(CHEN <i>et al.</i> , 2023c)	Geração automática de agentes	Texto
(ZHANG <i>et al.</i> , 2024e)	Fundamentação eficiente de LLMs	Multi

3.2 Resumo

Esta revisão examina aspectos específicos da integração de LLMs em sistemas multiagentes que são relevantes para o desenvolvimento de ferramentas de design de jogos educacionais. O foco concentra-se em três dimensões principais que fundamentam o sistema proposto nesta dissertação: percepção e comunicação entre agentes, integração de memória e conhecimento especializado, e adaptação para domínios específicos.

A análise da literatura revela que agentes baseados em LLM percebem e interagem através de três tipos de mensagens: ambientais (informações sobre contexto e cenário), de interação (comunicação entre agentes) e de auto-reflexão (processos internos de avaliação). Esta taxonomia de comunicação é fundamental para sistemas que necessitam coordenação entre agentes especializados, como no design colaborativo de jogos educacionais (TAN *et al.*, 2023; ZHANG *et al.*, 2023d).

A integração de memória e conhecimento externo permite que agentes mantenham contexto histórico e acessem informações especializadas durante o processo de design. Mecanismos de armazenamento e reflexão de memória possibilitam que agentes aprendam com experiências anteriores e refinem estratégias ao longo do tempo (PARK *et al.*, 2023; ZHAO *et al.*, 2024). A utilização de conhecimento externo, incluindo bases de conhecimento pedagógico e princípios de design de jogos, expande as capacidades dos agentes além de seu conhecimento interno (WANG *et al.*, 2024a).

A adaptação para domínios específicos ocorre através de estratégias de geração de perfis especializados que definem características e responsabilidades de cada agente. Três abordagens principais emergem da literatura: geração contextualizada

baseada em requisitos específicos da tarefa, métodos pré-definidos com grupos de agentes especializados, e métodos baseados em aprendizagem que evoluem durante a execução (DONG *et al.*, 2024a; QIAN *et al.*, 2023).

Estes aspectos fundamentam a arquitetura do sistema multiagente proposto para design de jogos educacionais, onde agentes especializados colaboram através de comunicação estruturada, mantêm memória de sessões de design, acessam conhecimento pedagógico externo, and adaptam-se aos requisitos específicos de diferentes contextos educacionais. A síntese destas capacidades possibilita a automação assistida do processo de design mantendo qualidade pedagógica e coerência no produto final.

3.2.1 Perfil do Agente

Sistemas multiagentes baseados em LLM executam tarefas complexas ou simulam cenários através de vários papéis (DONG *et al.*, 2024a; PARK *et al.*, 2022; ZHANG *et al.*, 2023c). A definição destes papéis envolve a elaboração de perfis de agentes, garantindo que cada agente se adeque à sua função designada. O perfil do agente é projetado para instanciar entidades inteligentes independentes com estilos personalizados, permitindo a realização de subtarefas específicas (QIAN *et al.*, 2023; SERAPIO-GARCÍA *et al.*, 2023).

Contexto do Perfil

De acordo com contextos específicos ou especificações do usuário, os perfis dos agentes podem abranger diferentes tipos e conteúdos de informação baseados em suas funções operacionais e especializações técnicas.

Como características funcionais fundamentais, o perfil tipicamente inclui identificação do agente, área de especialização, responsabilidades específicas e capacidades técnicas (CHEN *et al.*, 2024a; WOOLDRIDGE e JENNINGS, 1995). Para agentes especializados em tarefas, o perfil define conhecimento de domínio específico, metodologias aplicáveis, critérios de avaliação e parâmetros de operação (HONG *et al.*, 2023; RAO *et al.*, 1995).

O perfil pode incluir especificações comportamentais como estratégias de comunicação, protocolos de interação, critérios de tomada de decisão e mecanismos de validação. Agentes especializados requerem definição de heurísticas específicas do domínio, bases de conhecimento acessíveis e algoritmos de processamento adequados à sua área de atuação (KAPETANAKIS e KUDENKO, 2003; SMITH, 1980).

Além disso, o perfil pode especificar relações funcionais com outros agentes, dependências de entrada e saída, e contextos operacionais relevantes para a execução coordenada de tarefas complexas. Esta estruturação permite que agentes

mantenham coerência em suas especializações enquanto contribuem para objetivos sistêmicos mais amplos (LI *et al.*, 2023b; PANAIT e LUKE, 2005).

3.2.2 Estratégia de Geração

A seleção de informações para definir o perfil do agente é principalmente determinada pelos cenários de aplicação específicos, orientando a trajetória da geração de perfis. Considerando a relação entre modelagem de cenários e geração de agentes computacionais, a literatura estabelece três estratégias principais para criação de agentes especializados em execução de tarefas.

Método de Geração Contextualizada Neste método, agentes especializados são criados dinamicamente através da análise e decomposição de cenários de tarefas complexas. CHEN *et al.* (2024a) desenvolveram o AutoAgents, que representa a abordagem mais avançada, gerando automaticamente agentes com perfis específicos baseados na análise do conteúdo da tarefa. O framework TDAG complementa esta abordagem através de decomposição dinâmica de tarefas e geração de agentes sob demanda (WANG *et al.*, 2025a). STONE e VELOSO (1999) estabeleceram fundamentos teóricos através de formações computacionais que permitem decomposição do espaço de tarefas e atribuição dinâmica de papéis baseada em gatilhos de tempo de execução. Este método garante alinhamento entre perfis de agentes e requisitos específicos de tarefas, mas requer regeneração de perfis para cada novo cenário.

Método Pré-definido Esta estratégia emprega grupos pré-estabelecidos de agentes especializados, selecionando agentes adequados para tarefas específicas. O framework STEAM demonstra esta abordagem através de estruturas de equipe pré-planejadas com capacidade de reorganização dinâmica mediante falhas ou mudanças de tarefas (TAMBE, 1997). DECKER e LESSER (1993a) desenvolveram o framework TAEMS para seleção quantitativa de agentes baseada em perfis de capacidade e correspondência tarefa-agente. A taxonomia estabelecida por FRANKLIN e GRAESSER (1996) distingue agentes computacionais de agentes comportamentais, fornecendo base teórica para seleção apropriada de perfis pré-definidos. Este método reduz tempo de desenvolvimento quando múltiplos agentes são necessários, mas oferece controle limitado sobre especialização de perfis.

Método Baseado em Aprendizagem Nesta abordagem, agentes adaptam seus perfis e capacidades através de experiência e feedback ambiental. Sistemas multiagentes auto-adaptativos foram implementados usando o modelo MAPE-K para evolução autônoma de perfis (NASCIMENTO *et al.*, 2023). Redes modulares dirigidas por agentes com compartilhamento dinâmico de parâmetros permitem especialização

em tempo de execução através de otimização de roteamento (YU *et al.*, 2024b). Mecanismos de adaptação colaborativa possibilitam que agentes aprendam estratégias de coordenação e modifiquem perfis através de inteligência coletiva (TRAN *et al.*, 2025). Este método integra vantagens das abordagens anteriores, proporcionando flexibilidade adaptativa enquanto mantém eficiência operacional.

3.2.3 Percepção

A maioria dos humanos e animais adquire informações através de órgãos sensoriais como olhos, ouvidos e mãos, que servem como determinantes cruciais da cognição e comportamento individual. De forma similar, a aquisição de informações é vital para agentes como entidades inteligentes independentes, permitindo que percebam condições ambientais externas e seus estados internos. Estas informações são então convertidas em representações intermediárias através de módulos de percepção, que determinam os resultados da tomada de decisão autônoma do agente e suas respostas comportamentais (HUBEL e WIESEL, 1962; LOGOTHETIS e SHEINBERG, 1996).

Devido às capacidades de processamento de texto dos LLMs (ACHIAM *et al.*, 2023; BUBECK *et al.*, 2023; WEI *et al.*, 2022b), sistemas baseados em LLM utilizaram predominantemente mensagens textuais como meio para percepção e disseminação de informações. Nesses sistemas, extrair informações textuais do ambiente externo requer modelos especializados para converter informações em texto (SHRIDHAR *et al.*, 2021; TEAM *et al.*, 2022; WANG *et al.*, 2022), enquanto a informação de estado interno dos LLMs depende dos próprios modelos para extrair e resumir conhecimento em forma textual (FIRAT e KULELI, 2023).

O surgimento de Modelos de Linguagem Grandes Multimodais (MLLMs) mudou este paradigma, facilitando a transição da percepção de informação unimodal para multimodal e unificando as modalidades (DRIESS *et al.*, 2023) de maneira mais próxima à percepção humana. Para agentes baseados em LLM, é importante receber informações de diversas fontes e modalidades. Este arcabouço perceptivo expandido aprimora o entendimento dos agentes sobre seu ambiente e estados internos, permitindo que tomem decisões mais inteligentes e exibam comportamentos mais preciso.

3.2.4 Fonte de Mensagem

Quando imersos em cenários específicos, agentes baseados em LLM percebem, processam e geram mensagens durante interações e comunicações, que servem como condutos cruciais para os agentes colaborativamente realizarem tarefas complexas. De acordo com a natureza das interações dos agentes ou contrapartes de comunicação, a literatura existente categoriza as fontes de informação perceptual para agentes

nos seguintes três tipos:

Mensagem de Ambiente Completo:

Este tipo de mensagem transmite informações básicas sobre o ambiente ao redor dos agentes, como localização da cena, layout e mobiliário, bem como informações sensíveis ao tempo, como transições de cena e mudanças de instalações (KECSKÉS *et al.*, 2007). Adicionalmente, considera informações emocionalmente nuançadas como ambientes e atmosfera (MADASU *et al.*, 2023; MAMUN *et al.*, 2023). Tais informações estão intimamente ligadas aos cenários de tarefas e têm consistentemente mantido significância em trabalhos anteriores, seja em configurações de agente único ou múltiplos agentes. Inicialmente, esta informação é determinada pelo cenário de tarefa definido pelo usuário. No entanto, pode ser automaticamente gerada pelos próprios agentes ou por LLMs adicionais, especialmente durante a interação entre agentes. Tipicamente, esta mensagem surge de interações entre agentes e os elementos inerentes de seu ambiente, levando a mudanças no comportamento do agente e atualizações no ambiente circundante. Ocasionalmente, serve como informação de fundo suplementar (SONG *et al.*, 2023a; ZHANG *et al.*, 2023d), influenciando tanto as auto-interações dos agentes quanto suas interações uns com os outros.

Mensagem de Interação

Esta categoria de mensagem engloba informações trocadas durante interações entre agentes, com conteúdo que é flexivelmente determinado com base nos requisitos da tarefa ou cenários simulados. Por exemplo, em um cenário baseado em comunicação, o conteúdo da mensagem pertence a informações de diálogo entre agentes sobre um tópico específico. Cada mensagem de interação geralmente significa uma troca independente de informações entre dois agentes específicos, caracterizada por especificidade individual e relevância temporal, dada a natureza multi-round da comunicação entre agentes (BANG *et al.*, 2023; ZHANG e ZHAO, 2021). Tais mensagens são geralmente geradas autonomamente pelos agentes interagentes, embora ocasionalmente possam ser produzidas por LLMs adicionais como sinais de controle direcionados a agentes específicos. Servindo como o meio primário para comunicação e interação entre agentes, estas mensagens predominantemente influenciam os resultados da tomada de decisão de cada agente e respostas comportamentais.

3.2.5 Tipo de Mensagem

Após levar em conta várias fontes de mensagens de informação, também enfatizamos a importância de permitir que agentes recebam e compreendam múltiplas modali-

dades de informação através de módulos de percepção (YANG *et al.*, 2021). Esta seção se aprofunda nos métodos pelos quais agentes baseados em LLM podem alcançar capacidades de percepção multimodal, abrangendo entradas textuais, visuais e auditivas, visando enriquecer o domínio de percepção do agente e reforçar sua adaptabilidade e versatilidade.

Mensagem Textual

O texto serve como uma representação fundamental e intuitiva das percepções humanas em relação ao ambiente circundante e suas experiências subjetivas, tornando a comunicação textual a abordagem primária para humanos interagirem com o mundo. Considerando as preferências dos LLMs por entradas e saídas baseadas em texto, agentes baseados em LLM também utilizam mensagens textuais como o meio de informação principal para interação e disseminação. Mensagens textuais englobam informações textuais brutas, como descrições ambientais, saídas textuais de outros agentes e dados textuais do próprio agente. Elas também incluem dados convertidos derivados de outras modalidades, por exemplo, informações de legenda extraídas de imagens via modelos visuais (CORNIA *et al.*, 2020; LI *et al.*, 2023d). Essas mensagens cobrem um amplo espectro de informações, incluindo diálogos, planejamento de tarefas, feedback, etc. Para agentes baseados em LLM, a tarefa primária é compreender, analisar e sintetizar textos complexos e longos, semelhante às capacidades de especialistas humanos. Isso aproveita as funcionalidades centrais dos LLMs: compreensão, raciocínio e geração. Alguns esforços de pesquisa aprimoraram as capacidades de compreensão e raciocínio dos LLMs através de aprendizado em contexto (DONG *et al.*, 2022) e raciocínio de cadeia de pensamento (CoT) (WEI *et al.*, 2022c), visando produzir saídas que se alinhem mais proximamente com preferências cognitivas humanas e situações do mundo real. De forma similar, técnicas de engenharia de prompt e fine-tuning têm sido empregadas para saídas mais precisas (GAO e ZHANG, 2024). Além disso, alguns estudos focam em analisar e compreender os significados implícitos e conteúdo emocional dentro de textos. Por exemplo, certas abordagens empregam aprendizado por reforço para interpretar significados implícitos e feedback de modelo para derivar recompensas (BASU *et al.*, 2018; CRISTIANO *et al.*, 2017; LIN *et al.*, 2022a). Alguns outros métodos dependem de modelos especialistas para análise linguística refinada para alcançar um entendimento mais profundo de nuances textuais (SUMERS *et al.*, 2024), o que ajuda a deduzir as preferências do falante e leva a respostas de agente mais personalizadas e precisas. Adicionalmente, agentes baseados em LLM devem ser capazes de responder prontamente a situações novas encontradas em cenários complexos do mundo real. Isso ressalta a importância de aprimorar as habilidades dos agentes para perceber e entender novas tarefas através de texto. Em certos trabalhos, LLMs que

passaram por tuning de instruções demonstram impressionantes habilidades de entendimento e generalização de instruções zero-shot (WEI *et al.*, 2022a), eliminando a necessidade de fine-tuning específico para tarefas. Enquanto algumas abordagens introduzem um módulo adicional para incorporar conhecimento externo, dotando assim o LLM com um entendimento mais abrangente de novas tarefas.

Mensagem Visual

Mensagens textuais falham em capturar características nuançadas que informações visuais podem representar, como propriedades detalhadas de objetos, relações espaciais entre agentes, ou condições atmosféricas (DRIESS *et al.*, 2023). Integrar informações visuais fornece aos agentes contexto mais rico e entendimento mais preciso do ambiente.

Para equipar agentes com capacidades de compreensão visual, três abordagens principais são empregadas. A primeira utiliza modelos de legenda visual para gerar descrições textuais correspondentes (CORNIA *et al.*, 2020; LI *et al.*, 2023d). Este método é simples mas frequentemente perde informações visuais implícitas. A segunda codifica informações visuais em tokens visuais baseados em transformers, segmentando imagens em patches processados por codificadores (DOSOVITSKIY *et al.*, 2021; VAN DEN OORD *et al.*, 2017). A terceira transforma codificações de imagem em embeddings visuais integrados com representações textuais do LLM (DAI *et al.*, 2024; LI *et al.*, 2023c).

Para percepção de vídeo, métodos convertem vídeos em séries de quadros extraídos em intervalos específicos, aproveitando capacidades de percepção de imagem com atenção às transições temporais (ALAYRAC *et al.*, 2022; WANG *et al.*, 2024c).

Mensagem Auditiva

Informações de áudio abrangem sons ambientais, música e fala, transmitindo informações perceptuais através de características tempo-frequência que texto e dados visuais não podem replicar. Mensagens de áudio incluem elementos linguísticos como tom, entonação e nuance emocional (BORSOS *et al.*, 2023; LIU *et al.*, 2023b).

Três abordagens principais são utilizadas para percepção auditiva. A primeira emprega codificadores de áudio para converter áudio em tokens discretos integrados ao LLM. A segunda codifica informações como vetores latentes em modelos de difusão (CHEN *et al.*, 2023b; ZHANG *et al.*, 2023b). A terceira representa áudio como embeddings através de camadas conectadas, atenção cruzada multi-cabeça, ou Q-Former (HAN *et al.*, 2024; TANG *et al.*, 2024).

Sistemas como AudioGPT (HUANG *et al.*, 2024) aproveitam capacidades de uso de ferramentas dos LLMs, funcionando como hubs de controle para invocar ferra-

mentas de processamento de áudio existentes. Agentes baseados em LLM podem ser estendidos para processar modalidades adicionais como mapas 3D, dados GPS e informações de pose humana (SHEN *et al.*, 2023; ZHANG *et al.*, 2024f).

3.2.6 Auto-ação

Em contextos sociais, o ser humano, como entidade autônoma, processa informações percebidas para formar unidades de memória, construir sua consciência cognitiva, desenvolver pensamentos individuais e empreender ações (MARSHALL e MAGOUN, 2013). De forma similar, a auto-ação representa um mecanismo fundamental para o agente, funcionando como uma entidade independente, tomar decisões autônomas e realizar ações necessárias para sua sobrevivência e evolução no ambiente de interação. Esta seção investiga os processos detalhados pelos quais agentes individuais aprendem e raciocinam autonomamente dentro de seus ambientes. Ao receber informações percebidas, o módulo de auto-ação inicialmente invoca a memória na Seção 3.3.1 para extrair experiências históricas relevantes, possivelmente complementadas por conhecimento adicional recuperado de bases de conhecimento externas na Seção 3.3.2. Esta amalgamação de informações serve como contexto, auxiliando o agente no raciocínio, planejamento e generalização na Seção 3.3.3, culminando na tomada de decisão. Com base nestas decisões, o agente executa ações correspondentes para alcançar interações no mundo real na Seção 3.3.4. Simultaneamente, durante os processos de pensamento e ação, os agentes se engajam na auto-atualização e evolução da memória comparando experiências históricas, conhecimento atual e insights recém-gerados. Nas seções subsequentes, delinearemos os componentes do módulo de auto-ação em detalhes.

3.2.7 Memória

O módulo de memória, servindo como unidade de armazenamento e recuperação para o agente, é instrumental em permitir que ele aproveite conhecimento cognitivo e experiencial existente para se adaptar a interações dinâmicas com o ambiente ou outros agentes (QIAN *et al.*, 2023; ZHANG *et al.*, 2024c). Através deste processo, agentes acumulam novos insights e experiências, que podem aprimorar ainda mais suas habilidades cognitivas e inteligência atualizando a memória (ZHAO *et al.*, 2024; ZHU *et al.*, 2023b). As funcionalidades centrais da visualização de memória são ilustradas na Figura 3.2. Esta capacidade é crucial para o agente como indivíduo inteligente independente navegar flexivelmente por ambientes complexos e abordar tarefas inéditas. A realização desta funcionalidade adaptativa é primariamente alcançada através de três operações críticas de memória (WANG *et al.*, 2024a): recuperação de memória, armazenamento de memória e reflexão de memória.

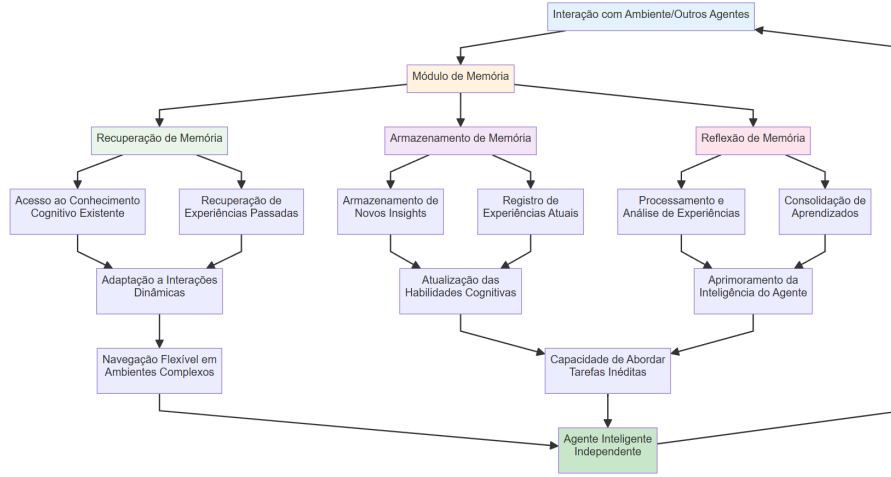


Figura 3.2: Representação das funcionalidades centrais do módulo de memória em agentes baseados em LLM, mostrando como as três operações críticas (recuperação, armazenamento e reflexão de memória) trabalham em conjunto para permitir adaptação dinâmica e navegação flexível em ambientes complexos através de um ciclo contínuo de aprendizado e aprimoramento cognitivo.

Recuperação de Memória

No âmbito dos agentes inteligentes, a recuperação efetiva de informações é fundamental para facilitar interações dinâmicas dentro de ambientes complexos ou com outros agentes, e a informação recuperada é sempre tratada como referências experienciais substanciais. A recuperação de memória visa aprimorar a precisão da tomada de decisão extraindo informações valiosas pertinentes à situação atual da memória de um agente. Esta informação engloba vários elementos como percepção ambiental, registros de interações históricas, dados experienciais e conhecimento externo. Em cenários envolvendo memória de curto prazo (CHASE, 2022; QIAN *et al.*, 2023), o módulo de recuperação tipicamente extrai todo o corpo de informações como conteúdo. Contudo, ao lidar com memória de longo prazo, o módulo de recuperação geralmente emprega mecanismos de filtragem para discernir e apresentar apenas as memórias mais relevantes ao agente (PACKER *et al.*, 2023; SHINN *et al.*, 2023; ZHAO *et al.*, 2024). Esta distinção ressalta a necessidade de estratégias de recuperação personalizadas para otimizar a utilidade e relevância das informações acessadas, fortalecendo assim a eficácia operacional do agente em diversos contextos.

Métodos de Recuperação.

Para manter a flexibilidade e adaptabilidade dinâmica dos agentes, memórias são recuperadas de maneira automatizada (LIN *et al.*, 2023; ZHONG *et al.*, 2024). Uma metodologia fundamental em pesquisas anteriores enfatiza que, servindo como contexto de prompt, a informação de memória é avaliada com base em métricas predefi-

nidas: Recência, Relevância e Importância (PARK *et al.*, 2023). Estas métricas são usadas para calcular uma pontuação ponderada para cada memória, com aquelas que pontuam mais alto sendo priorizadas para uso contextual, enquanto os parâmetros do modelo permanecem fixos. Outra abordagem notável considera a informação recuperada como uma representação aprendível, como embeddings e vetores (QIAN *et al.*, 2023; ZHONG *et al.*, 2024; ZHU *et al.*, 2023b), que servem como orientação suave para ajuste fino do modelo para acomodar várias tarefas. Técnicas como aprendizado por reforço online, aprendizado multitarefa e mecanismos de atenção, facilitam atualizações em tempo real e ajustes nos parâmetros do modelo, aprimorando assim a responsividade do agente a tarefas e ambientes em evolução.

Extensão de Recuperação.

Vários estudos têm se concentrado em empregar agentes baseados em LLM como uma interface de controle central para facilitar aplicações downstream envolvendo gerenciamento de memória. Por exemplo, algumas pesquisas projetaram mecanismos de memória interativos para agentes baseados em LLM com o objetivo de melhorar a operabilidade da memória para permitir intervenção e controle mais humanizados (HU *et al.*, 2023a; HUANG *et al.*, 2023b). Em tais sistemas, representações de informações podem ser manipuladas, editadas, deletadas ou amalgamadas através de sumarização. Em certos estudos, usuários são capacitados a visualizar e manipular o histórico de diálogo, modificando assim a memória histórica dos agentes. Especificamente, em (HU *et al.*, 2023a), permitiu-se operações de memória, como deleção, baseadas em comandos do usuário para ajustar as informações de memória de acordo. Estas abordagens visam fornecer controle mais intuitivo e flexível sobre os sistemas de memória dentro de agentes baseados em LLM.

Armazenamento de Memória

Armazenar informações críticas na memória constitui a base de conhecimento fundamental na qual os agentes se baseiam para perceber e agir dentro de ambientes complexos, aprimorando assim sua eficiência e racionalidade. O propósito do armazenamento de memória é arquivar as informações percebidas e as experiências aprendidas pelos agentes durante interações. Tipicamente, este processo envolve escrever texto em linguagem natural na memória, uma tarefa que engloba selecionar locais de armazenamento apropriados dentro da memória e gerenciar a substituição de informações (ACHIAM *et al.*, 2023; SCHUURMANS, 2023). Esta abordagem sistemática para armazenamento de memória assegura que os dados mais pertinentes estejam prontamente acessíveis, facilitando tomada de decisão informada e respostas adaptativas pelos agentes.

Formato de Armazenamento.

O armazenamento de memória é tipicamente realizado através do uso de texto formado em linguagem natural (BERTSCH *et al.*, 2023; MANAKUL e GALES, 2021; NIE *et al.*, 2022), embora também englobe informações multimodais como dados visuais e de áudio (ZHANG *et al.*, 2024c). O formato de armazenamento é determinado pela natureza específica da tarefa e pelos atributos da modalidade de dados. Ao adaptar o formato de armazenamento à modalidade e requisitos da tarefa, agentes podem utilizar mais efetivamente informações armazenadas, aprimorando assim seu desempenho em ambientes diversos e complexos. Usando estruturas de armazenamento de dados aprimoradas, metodologias representativas existentes alcançaram armazenamento de informações mais eficiente e flexível dentro da memória. Notavelmente, alguns estudos enfatizam a geração de representações condensadas de memória nos processos reflexivos (PARK *et al.*, 2023). Por exemplo, vários métodos adotam vetores de embedding para representar seções de memória e diálogos históricos (LIN *et al.*, 2023; ZHONG *et al.*, 2024; ZHU *et al.*, 2023b). Estas técnicas variadas ressaltam os esforços contínuos para aprimorar a funcionalidade e acessibilidade do armazenamento de memória em ambientes computacionais complexos. Outra abordagem efetiva envolve adotar métodos de interação de dados mais intuitivos para alcançar armazenamento efetivo de memória. Por exemplo, ChatDB (HU *et al.*, 2023a) e DB-GPT (ZHOU *et al.*, 2023a) abrangem manipulação de dados através de comandos SQL integrando o LLM com bancos de dados. Esta integração permite uma interface perfeita e eficiente para gerenciar e consultar dados armazenados, aprimorando assim a eficiência e usabilidade geral do sistema de memória.

Métodos de Armazenamento.

Ao considerar o processo de escrita de memória, dois desafios predominantes devem ser meticulosamente abordados: a relação entre a nova informação e as memórias existentes, e estratégia efetiva de troca de informações quando a capacidade de armazenamento de memória é atingida. (1) *Modificação de Memória*. Ao considerar a similaridade entre nova informação e memórias existentes, é crucial determinar o método apropriado de incorporação: se deve adicionar nova informação, mesclá-la com dados existentes, ou substituir informações existentes errôneas. Por exemplo, uma abordagem armazena sequências de ações bem-sucedidas com o mesmo subobjetivo em uma única lista (ACHIAM *et al.*, 2023). Quando o comprimento desta lista excede o limite predefinido, todas as entradas são comprimidas em uma solução unificada pelos LLMs, que subsequentemente substitui as entradas originais na lista. (2) *Troca de Memória*. Dado que o armazenamento de memória é tipicamente limitado, projetar uma estratégia efetiva de troca de informações é significativo para

assegurar que a memória retenha as informações mais benéficas para os agentes. Ao considerar a escrita de nova informação em uma memória cheia, métodos existentes empregam mecanismos estratégicos de troca de informações para maximizar a retenção da informação mais próxima e relevante.

Reflexão de Memória

Reflexão de Memória é o processo pelo qual agentes se engajam em auto-aprimoramento baseado na informação percebida e experiência aprendida de interações históricas armazenadas na memória. Este processo emula a prática humana de resumir, refinar e refletir sobre conhecimento existente, com o objetivo de aprimorar a adaptabilidade do agente a novos ambientes e tarefas. O processo de reflexão de memória tipicamente ocorre automaticamente, com agentes independentemente atualizando sua memória com base em conhecimento recém-adquirido, alcançando assim atualizações de auto-reconhecimento (ZHAO *et al.*, 2024; ZHU *et al.*, 2023b). Em um ambiente multi-agente, um agente central baseado em LLM exerce controle sobre a reflexão de memória de agentes individuais. Este agente central envia sinais de controle específicos para guiar o processo de reflexão, assegurando coerência e coordenação através da rede de agentes. Este método facilita a atualização sistemática de memória, permitindo que agentes refinem seus modelos cognitivos e aprimorem sua adaptabilidade a tarefas e ambientes dinâmicos. Após estabelecer os mecanismos subjacentes à reflexão de memória, é crucial considerar cuidadosamente o conteúdo da reflexão de memória. Uma porção significativa de trabalhos anteriores tem se concentrado em armazenamento hierárquico de informações, enfatizando a abstração, sumarização e destilação de conhecimento e experiências adquiridas. Por exemplo, em Generative Agent (PARK *et al.*, 2023), o agente é capaz de resumir suas experiências passadas armazenadas na memória em insights mais amplos e mais abstratos. Este processo começa com o agente gerando três questões chave baseadas em suas memórias recentes. Estas questões são então usadas para consultar a memória para recuperar informações relevantes. Com base nas informações adquiridas, o agente gera ideias de alto nível. No framework ExpeL (ZHAO *et al.*, 2024), durante a execução de tarefas, agentes aprendem com as experiências de trajetórias corretas e derivam lições de incorretas. Outra abordagem significativa foca na generalização de conhecimento existente. Notavelmente, em GITM (ZHU *et al.*, 2023b), ao encontrar uma nova tarefa, as ações de agentes que cumprem com sucesso os subobjetivos são armazenadas em uma lista. Este processo hierárquico e reflexivo de utilização de memória permite que agentes refinem suas estratégias e melhorem o desempenho através de tarefas e ambientes variados.

3.2.8 Utilização de Conhecimento

A utilização de conhecimento foca na integração de conhecimento externo (excluindo informações de memória) no planejamento baseado em LLM. Aproveitando dados textuais, visuais e auditivos atualizados, LLMs aprimoram sua habilidade de realizar tarefas complexas com precisão e contextualização. Técnicas como geração aumentada por recuperação e scraping web em tempo real permitem que estes modelos combinem capacidades internas com informações externas, aprimorando assim processos de planejamento e tomada de decisão. O fluxograma geral ilustrando o mecanismo operacional da memória é apresentado na Figura 3.3.

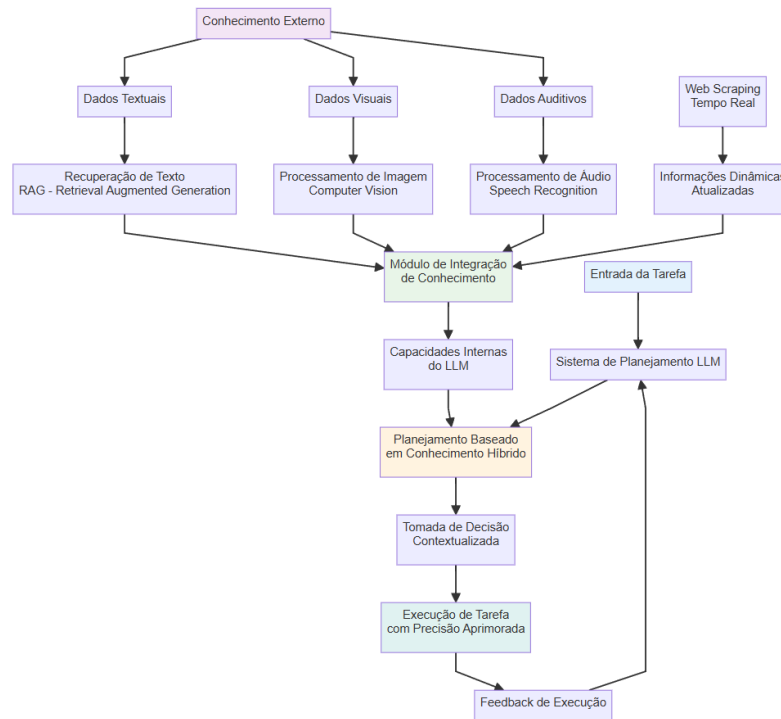


Figura 3.3: Representação de como agentes baseados em LLM integram conhecimento externo multimodal (textual, visual e auditivo) através de técnicas como RAG e web scraping em tempo real para aprimorar o planejamento e tomada de decisão contextualizada.

Conhecimento para Agentes Baseados em LLM

A natureza diversa das tarefas requer formas variadas de conhecimento. Nesta seção, examinamos como agentes baseados em LLM utilizam conhecimento textual, visual, auditivo e outros domínios específicos. Entendendo estes mecanismos, podemos apreciar a versatilidade e efetividade dos LLMs em lidar com uma ampla gama de tarefas.

Conhecimento Textual.

O conhecimento textual é a espinha dorsal dos LLMs, dado seu treinamento em extensos corpora de texto. Este conhecimento é vital para tarefas como compreensão de linguagem natural, geração de texto, tradução e mais. Os formatos de conhecimento textual incluem linguagem natural, embeddings, tokens e estruturas em árvore. Linguagem natural é o formato primário de entrada e saída, embeddings capturam significado semântico, tokens segmentam texto em unidades processáveis, e estruturas em árvore permitem tarefas complexas de raciocínio. LLMs utilizam conhecimento textual tanto interno quanto externo para realizar estas tarefas (WANG *et al.*, 2024a). Pré-treinados em vastos conjuntos de dados, LLMs podem entender e gerar texto baseado em conhecimento internalizado, incluindo sintaxe linguística, semântica e conhecimento geral de mundo. Isto permite que LLMs realizem tarefas como geração de texto, sumarização, tradução e até planejamento com contexto adicional mínimo.

Conhecimento Visual.

Em agentes LLM, conhecimento visual é primariamente representado através de embeddings contínuos gerados por codificadores visuais, que são então integrados com informações textuais para facilitar entendimento e raciocínio de dados multimodais. A representação de conhecimento visual tipicamente inclui representações vetoriais latentes de imagens (por exemplo, codificações de Transformer visual), codificações objeto-cêntricas, e outras formas, todas processadas junto com informações de linguagem através de mecanismos padrão de auto-atenção. Agentes LLM alavancam estes embeddings visuais para alcançar forte desempenho através de várias tarefas, como VQA, legendagem de imagens e raciocínio incorporado. Em aplicações práticas, PaLM-E (DRIESS *et al.*, 2023) incorpora ainda mais entradas contínuas como dados visuais e estimações de estado no LLM, permitindo raciocínio incorporado e tomada de decisão através de um framework de processamento multimodal unificado, demonstrando capacidades de aprendizado de transferência entre tarefas. Modelos como LLaVA (LIU *et al.*, 2024a) integram codificadores visuais CLIP com modelos de linguagem e aplicam ajuste fino de instrução visual, permitindo raciocínio conjunto sobre informações visuais e textuais em tarefas complexas.

Conhecimento Auditivo.

Conhecimento auditivo engloba fala e eventos auditivos, que podem ser representados através de formas como codificadores de fala e imagens de espectrograma. Ao processar fala, agentes LLM podem discretizar entrada de fala via módulos de conexão e incorporá-la em um espaço vetorial compartilhado com texto. Por exem-

plo, em SpeechGPT (ZHANG *et al.*, 2023a), tokens de fala gerados pelo codificador HuBERT (HSU *et al.*, 2021) são incorporados no vocabulário LLaMA (TOUVRON *et al.*, 2023), permitindo que o LLM processe entrada de fala. Outra abordagem envolve alinhar codificadores de fala com o LLM usando conectores como camadas totalmente conectadas, atenção cruzada multi-cabeça (LYU *et al.*, 2023a), ou Q-Former (LI *et al.*, 2023c), que preserva mais informações de fala e alcança compressão eficiente, suportando assim o processamento de segmentos longos de fala (YU *et al.*, 2024a). Eventos auditivos são tipicamente tratados como imagens de espectrograma de tamanho fixo e processados usando métodos de modelos de linguagem visual. Adicionalmente, LLMs de áudio end-to-end, como AudioPaLM (RUBENSTEIN *et al.*, 2023), podem lidar simultaneamente com fala e outros sinais auditivos para atender requisitos auditivos mais amplos. Por exemplo, em AudioGPT (HUANG *et al.*, 2024), o LLM é integrado com vários modelos auditivos fundacionais para processar informações auditivas complexas, permitindo reconhecimento automático de fala (ASR) e conversão de texto para fala (TTS). Estes exemplos demonstram as capacidades robustas e adaptabilidade extensiva dos LLMs em processar e gerar áudio.

Outro Conhecimento.

Além de texto, dados visuais e auditivos, LLMs frequentemente precisam utilizar conhecimento especializado de domínios específicos como pesquisa científica (COBBE *et al.*, 2021; DING *et al.*, 2023), informações médicas (SINGHAL *et al.*, 2023a,b; TU *et al.*, 2024), ou especificações técnicas (MADAAN *et al.*, 2022b; PALLAGANI *et al.*, 2024; XU *et al.*, 2022). Isto aprimora sua habilidade em lidar com tarefas que requerem expertise profunda de domínio. Formatos de conhecimento de domínio específico incluem descrições em linguagem natural, embeddings, tokens e estruturas em árvore, que permitem que LLMs processem e entendam informações complexas de vários campos. Em domínios científicos, LLMs podem auxiliar em análise de dados, geração de hipóteses e revisão de literatura. Por exemplo, em (DING *et al.*, 2023), destaca-se como integrar conhecimento de domínio específico aprimora o desempenho de LLMs em tarefas especializadas. Enquanto em (COBBE *et al.*, 2021), aprimora-se a capacidade de grandes modelos de linguagem para realizar raciocínio matemático multi-passo treinando verificadores em um conjunto de dados diverso de problemas de palavras matemáticas elementares, que avaliam a correção de soluções geradas pelo modelo e selecionam a resposta mais precisa. No campo médico, LLMs podem apoiar profissionais recuperando e sintetizando informações médicas de bancos de dados como PubMed (CHOWDHERRY *et al.*, 2023). Esta capacidade é crucial para aplicações como suporte à decisão clínica, onde informações precisas e atualizadas são essenciais. Por exemplo, MedPaLM (SINGHAL *et al.*, 2023a),

um LLM ajustado para diálogo médico, aproveita conhecimento de domínio específico para fornecer respostas precisas e confiáveis a consultas médicas. Este modelo integra literatura médica e diretrizes clínicas em sua base de conhecimento, aprimorando sua habilidade de realizar tarefas como assistência a diagnóstico e educação de pacientes.

Recuperação de Conhecimento

A recuperação de conhecimento é um aspecto crítico da utilização efetiva de LLMs, pois envolve acessar informações externas para suplementar a base de conhecimento inerente do modelo. Isto assegura que LLMs possam fornecer respostas precisas e contextualmente relevantes, aprimorando seu desempenho através de vários domínios. Diversos métodos são empregados para recuperação eficiente de conhecimento, cada um com sua própria abordagem e aplicações.

Consultas a Bancos de Dados e Bases de Conhecimento.

Consultas a bancos de dados e bases de conhecimento envolvem acessar dados estruturados de repositórios como Google Knowledge Graph, PubMed (CHOWDHERY *et al.*, 2023) e outros bancos de dados de domínio específico. Estas fontes oferecem informações confiáveis e organizadas que podem ser integradas com saídas de LLM para aprimorar a precisão e relevância de respostas geradas. Um exemplo notável de integração de bancos de dados externos é o sistema ChatDB (HU *et al.*, 2023a), que usa consultas SQL para buscar dados relevantes logicamente, tornando mais fácil para agentes operarem. Similarmente, SQL-PALM (SUN *et al.*, 2023) emprega um modelo Text-to-SQL baseado em LLMs, aprimorando significativamente a precisão de consulta e interações com banco de dados. Outro exemplo, KnowledGPT (WANG *et al.*, 2023c), permite que LLMs acessem e recuperem conhecimento de bases de conhecimento externas através do prompting "Program of Thoughts", aprimorando assim sua habilidade de responder perguntas.

Web Scraping e Chamadas de API.

Web scraping e chamadas de API permitem que agentes baseados em LLM colem informações em tempo real da internet. Este método é particularmente útil para tarefas que requerem dados atualizados, como sumarização de notícias ou análise de mercado. Web scraping envolve usar ferramentas automatizadas para extrair dados de páginas web, fornecendo grandes quantidades de dados de fontes diversas. Chamadas de API, por outro lado, envolvem consultar APIs para buscar informações específicas, como artigos de notícias, atualizações climáticas ou dados financeiros. Diversos estudos integraram LLMs com ferramentas específicas como busca web

(NAKANO *et al.*, 2021), compilador (GAO *et al.*, 2023b) e calculadora (COBBE *et al.*, 2021). (PARISI *et al.*, 2022) criou um conjunto de dados para instrução de API e ajustou LLMs para ajudá-los a usar ferramentas e recuperadores efetivamente. Gorilla (PATIL *et al.*, 2023) é um LLM de ajuste fino que supera o desempenho do GPT-4 (ACHIAM *et al.*, 2023) em escrever chamadas de API, visando gerar parâmetros de entrada precisos para chamadas de API e aliviar alucinações durante chamadas externas de API.

Geração Aumentada por Recuperação (RAG).

Modelos RAG combinam mecanismos de recuperação com modelos generativos para produzir respostas ricas em contexto (LEWIS *et al.*, 2020). Esta abordagem é efetiva para resposta a perguntas de domínio aberto e agentes conversacionais. No estágio de recuperação, o sistema extrai fragmentos de documentos relevantes para a consulta de fontes de conhecimento externas. A fonte primária de recuperação é dados textuais, mas pode ser estendida para dados semi-estruturados (por exemplo, PDFs) (LUO *et al.*, 2023; ZHA *et al.*, 2023), dados estruturados (por exemplo, grafos de conhecimento) (HE *et al.*, 2024b), e conteúdo gerado pelos próprios LLMs (CHENG *et al.*, 2024a; YU *et al.*, 2022). Além da recuperação de passo único comumente usada, RAG incorpora três tipos de processos de aprimoramento de recuperação: recuperação iterativa (SHAO *et al.*, 2023), recuperação recursiva (TRIVEDI *et al.*, 2023) e recuperação adaptativa (ASAI *et al.*, 2023; JIANG *et al.*, 2023), que são projetados para melhorar eficiência e precisão na solução de consultas complexas (GAO *et al.*, 2023c). No estágio de geração, o modelo melhora a qualidade das respostas dos LLMs re-classificando segmentos de documentos para destacar os resultados mais relevantes (ZHUANG *et al.*, 2023) ou selecionando ou comprimindo contextos para reduzir informações redundantes e gerenciar entradas excessivamente longas (XU *et al.*, 2023b; YANG *et al.*, 2023a). Adicionalmente, LLMs podem ser ajustados para cenários específicos e características de dados, aprimorando a relevância e precisão das respostas geradas (DU e JI, 2022; LI *et al.*, 2023f).

Questões de Extração

No desenvolvimento e aplicação de LLMs, uma gama de questões de extração são encontradas, impactando diretamente a precisão, aplicabilidade e viés dos modelos. Estas questões abrangem desafios relacionados à atualização de conhecimento, alucinação e viés. Abordar estes desafios necessita uma abordagem abrangente integrando estratégias como alavancar fontes de conhecimento externas, aprimorar transparência e empregar técnicas de desviés.

Editar Conhecimento Errado e Desatualizado.

Um dos desafios primários para agentes LLM na extração de conhecimento é assegurar a tempestividade e precisão da informação. Como LLMs são tipicamente treinados em dados históricos, isto pode levar a um atraso no processamento das informações mais recentes. Quando tarefas requerem conhecimento que é mais recente que os dados de treinamento, LLMs frequentemente lutam para lidar. Uma abordagem direta é atualizar regularmente LLMs com novos dados, mas o ajuste fino de LLMs incorre em altos custos, e treinamento incremental pode resultar em esquecimento catastrófico (KIRKPATRICK *et al.*, 2017), onde o modelo perde o conhecimento amplo que adquiriu durante o pré-treinamento. Portanto, desenvolver métodos eficientes para incorporar novo conhecimento em LLMs existentes para mantê-los atualizados torna-se primordial. Abordagens atuais incluem alavancar fontes de conhecimento externas para suplementar a base de conhecimento de LLMs (PENG *et al.*, 2023; WANG *et al.*, 2023c). Integrando informações relevantes recuperadas no contexto, LLMs podem adquirir novo conhecimento factual e ter melhor desempenho em tarefas relevantes. Contudo, estes métodos ainda ficam aquém ao lidar com atualizações de conhecimento mais profundas. Técnicas de edição de modelo (LI *et al.*, 2024b; MITCHELL *et al.*, 2022; YAO *et al.*, 2023b) também são empregadas para alterar comportamento do modelo, seja modificando parâmetros do modelo ou usando mecanismos de pós-edição externos para alcançar atualizações de conhecimento, mas ainda enfrentam limitações em aplicações do mundo real devido à sua baixa especificidade.

Alucinação.

Alucinação refere-se ao fenômeno onde agentes LLM geram texto que se desvia da realidade (MAYNEZ *et al.*, 2020; RAUNAK *et al.*, 2021; ZHANG *et al.*, 2023g). Alucinações podem ocorrer devido à supergeneralização do modelo a partir de dados de treinamento ou interpretações errôneas de informações incompletas ou enganosas. Alucinações geradas por LLMs podem ser categorizadas em dois tipos: alucinações intrínsecas e alucinações extrínsecas (JI *et al.*, 2023). Alucinações intrínsecas envolvem geração de texto que contradiz a lógica de entrada, enquanto alucinações extrínsecas envolvem geração de texto contendo informações que não podem ser verificadas com informações existentes. Para abordar o problema de alucinação, pesquisadores propuseram vários métodos. Uma abordagem é integrar bases de conhecimento externas e sistemas de verificação de fatos para verificar a precisão do conteúdo gerado (GUO *et al.*, 2022; HU *et al.*, 2023a; THORNE e VLACHOS, 2018). Outra abordagem é aprimorar a transparência e interpretabilidade do modelo para melhorar a credibilidade das saídas (DHULIAWALA *et al.*, 2024; HUANG

et al., 2023a; WEI *et al.*, 2022c). Estes métodos incluem ajuste fino com dados de alta qualidade ou ajuste fino baseado em feedback humano (OUYANG *et al.*, 2022; WU *et al.*, 2024a; ZHOU *et al.*, 2024a). Por exemplo, a tarefa TruthfulQA (LIN *et al.*, 2022b) visa detectar se o modelo imita declarações falsas humanas. Adicionalmente, algumas técnicas como geração aumentada por recuperação e estratégias de decodificação estão sendo exploradas para reduzir alucinações. Métodos de geração aumentada por recuperação (LEWIS *et al.*, 2020; SHAO *et al.*, 2023; ZHANG *et al.*, 2023f) aprimoram a precisão da geração de linguagem introduzindo material fonte adicional e fornecendo mecanismos para verificar inconsistências entre a resposta gerada e o material fonte. Estratégias de decodificação (CHUANG *et al.*, 2023; DHULIAWALA *et al.*, 2024; SHI *et al.*, 2024) otimizam a forma como modelos de linguagem selecionam tokens de saída durante a geração de texto, equilibrando diversidade e precisão factual, mitigando assim a ocorrência de alucinações.

Viés Espúrio.

No reino da inteligência artificial, a justiça e precisão dos modelos são frequentemente comprometidas por vieses fortuitos e desequilíbrios de classe presentes nos dados de treinamento, coletivamente referidos como viés espúrio. Uma preocupação importante é o aprendizado de atalhos, onde modelos dependem de pistas espúrias e não generalizáveis nos dados de treinamento ao invés de aprender características robustas. Por exemplo, um modelo de linguagem pode prever incorretamente devido à co-ocorrência frequente de palavras de função com rótulos específicos no conjunto de treinamento (GEIRHOS *et al.*, 2020). Adicionalmente, modelos podem desenvolver preferências baseadas na ordem das amostras de treinamento, como a posição de respostas em tarefas de resposta a perguntas potencialmente influenciando julgamentos do modelo (LU *et al.*, 2022). O aprendizado de atalhos pode ser mitigado através de métodos como desvios de dados, treinamento adversarial, regularização interpretativa e regularização de confiança (DU *et al.*, 2023a). Viés de rótulo representa outra forma de pseudo-viés, frequentemente derivando de desequilíbrio de classe dentro do conjunto de dados de treinamento. Este desequilíbrio pode causar modelos a serem excessivamente sensíveis à classe majoritária enquanto sendo insuficientemente sensíveis a classes minoritárias. Por exemplo, em tarefas de análise de sentimento, um modelo pode tender a prever sentimento positivo devido a uma superabundância de amostras positivas, mesmo quando o texto realmente transmite sentimento negativo (TANG *et al.*, 2023a). Para reduzir tal viés, pesquisadores propõem reequilibrar conjuntos de dados, empregar técnicas de amostragem avançadas, e desenvolver novas métricas de avaliação para aprimorar a justiça e robustez do modelo (ZHOU *et al.*, 2024b).

3.2.9 Utilização de Habilidades do Agente

As capacidades dos agentes baseados em LLM são uma manifestação de sua inteligência cognitiva. Aproveitar estas habilidades permite que agentes analisem, sintetizem informações percebidas e se engajem em pensamento criativo. Dada a proficiência excepcional dos LLMs em lidar com informações contextuais longas, categorizamos as habilidades dos agentes em três tipos principais: raciocínio, planejamento e generalização. Raciocínio envolve inferência lógica baseada em experiências históricas e conhecimento atual, extraindo paradigmas universais. Planejamento implica na aplicação de regras gerais de alto nível a novos cenários, resultando em planos concretos e acionáveis. Generalização busca aplicar experiências existentes para abordar situações e problemas novos. Nas seções seguintes, forneceremos uma exploração aprofundada da utilização destas três capacidades.

Raciocínio e Planejamento

Raciocínio e planejamento envolvem o processo sistemático de alavancar experiência histórica, conhecimento comum e informações de estado atual para realizar análise lógica, derivando assim insights de alto nível, mais profundos. Subsequentemente, estes insights são aplicados à situação presente para gerar resultados inferenciais atualizados. Além disso, à medida que tarefas avançam, agentes podem utilizar introspecção para ajustar seus planos, assegurando um alinhamento mais próximo com condições do mundo real, facilitando, em última instância, execução de tarefas adaptativa e bem-sucedida. Estas habilidades são capacidades humanas fundamentais que sustentam resolução de problemas, tomada de decisão e análise crítica, formando a base da percepção humana e interação social (GALOTTI, 1989; WASON e JOHNSON-LAIRD, 1972). Estes processos cognitivos abrangem três componentes chave: raciocínio dedutivo, indutivo e abduutivo (HUANG e CHANG, 2023). Aproveitando a capacidade robusta de raciocínio e planejamento baseado em texto inerente aos grandes modelos de linguagem (LLMs) (BUBECK *et al.*, 2023; WEI *et al.*, 2022b), torna-se essencial para agentes baseados em LLM perceber efetivamente ambientes complexos, executar tarefas intrincadas e se engajar em interações humanizadas. Este framework sofisticado de raciocínio e planejamento permite que agentes naveguem e se adaptem a cenários dinâmicos, aprimorando assim sua capacidade de desempenhar e interagir de maneira semelhante a processos cognitivos humanos. De acordo com os passos e resultados de decisão associados ao raciocínio e planejamento dos agentes, delinearemos as metodologias relevantes e processos inferenciais de duas perspectivas: (1) Raciocínio de Passo Único e (2) Raciocínio de Múltiplos Passos. Estas abordagens fornecem um framework estruturado para entender como agentes utilizam capacidades de raciocínio e planejamento para derivar

decisões, adaptar-se a novas informações e planejar efetivamente suas ações dentro de contextos variados. A visualização destas abordagens é apresentada na Figura 3.4.

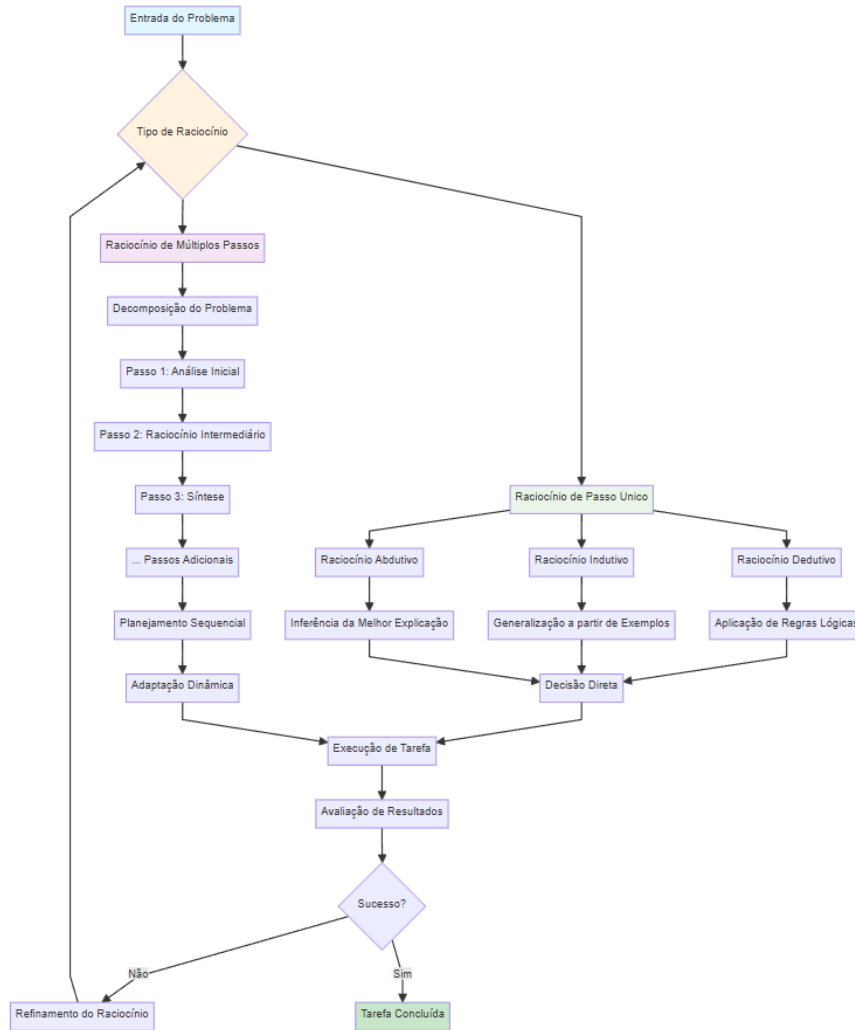


Figura 3.4: Representação de raciocínio e planejamento em agentes, distinguindo entre abordagens de passo único (dedutivo, indutivo e abduutivo) e múltiplos passos (com decomposição sequencial), incluindo mecanismos de avaliação e refinamento iterativo.

Estrutura de Plano.

Durante o processo de formulação de plano, agentes geralmente decompõem uma tarefa abrangente em numerosas subtarefas, e várias abordagens foram propostas nesta fase. Notavelmente, alguns trabalhos defendem que agentes baseados em LLM decomponham problemas abrangentemente de uma vez, formulando um plano completo de uma vez e então executando-o sequencialmente (AHN *et al.*, 2022; XU *et al.*, 2023a). Em contraste, outros estudos como a série CoT empregam uma estratégia adaptativa, onde planejam e abordam subtarefas uma de cada vez, permitindo mais

fluidez ao lidar com tarefas intrincadas em sua totalidade (KOJIMA *et al.*, 2022; LYU *et al.*, 2023b; WEI *et al.*, 2022c). Adicionalmente, alguns métodos enfatizam planejamento hierárquico, enquanto outros ressaltam uma estratégia na qual planos finais são derivados de passos de raciocínio estruturados em um formato semelhante a árvore (WU *et al.*, 2022; YAO *et al.*, 2024). A última abordagem argumenta que agentes devem avaliar todos os caminhos possíveis antes de finalizar um plano. Enquanto agentes baseados em LLM demonstram um escopo amplo de conhecimento geral, podem ocasionalmente enfrentar desafios quando encarregados de situações que requerem conhecimento especializado. Aprimorar estes agentes integrando-os com planejadores de domínios específicos demonstrou produzir melhor desempenho.

Método de Passo Único.

Nesta estratégia, agentes decompõem uma tarefa complexa em várias subtarefas através de um único processo de raciocínio e planejamento baseado nas diretrizes da tarefa atual. Estas subtarefas são ordenadas sequencialmente, com cada subtarefa logicamente seguindo a anterior. Agentes baseados em LLM aderem a estes passos para alcançar o objetivo final. Tipicamente, agentes realizam o processo de raciocínio através de elicitacão baseada em prompt, onde o contexto inclui registros históricos da memória, o estado do ambiente circundante e o status atual dos agentes como informação auxiliar para tomada de decisão. Com base nas diretrizes da tarefa atual, agentes integram sua inteligência inerente com conhecimento externo para deduzir uma série de passos racionais e viáveis para resolver tarefas complexas. Especificamente, aprendizado em contexto introduz uma metodologia onde LLMs são fornecidos com alguns exemplos de raciocínio e planejamento, permitindo-lhes inferir soluções para novas situações através de raciocínio e planejamento análogos. Por exemplo, a técnica Chain of Thought (CoT) (WEI *et al.*, 2022c) induz LLMs a pensar sobre problemas passo a passo, sistematicamente desconstruindo tarefas intrincadas em componentes manejáveis, facilitando assim planejamento e deliberação a longo prazo. A abordagem Zero-shot-CoT (KOJIMA *et al.*, 2022) capacita LLMs a gerar autonomamente processos de raciocínio para tarefas induzindo-os com frases gatilho como "pense passo a passo". Para aprimorar a racionalidade e precisão da tomada de decisão dos LLMs, mitigando o problema de alucinação que pode ocorrer durante raciocínio de passo único, várias abordagens empregam raciocínio multi-caminho para selecionar o resultado ótimo. Cada passo intermediário pode levar a múltiplos passos subsequentes. Especificamente, A abordagem Tree of Thought (ToT) (YAO *et al.*, 2024) decompõe problemas em uma estrutura de árvore, criando múltiplos caminhos de solução com cada nó representando um estágio diferente de "pensamento". O Graph of Thought (BESTA *et al.*, 2024) expande a estrutura de raciocínio semelhante a árvore no ToT para estruturas de grafo, resultando em

estratégias de prompting mais robustas. O RAP (HAO *et al.*, 2023) constrói um modelo de mundo para simular os benefícios potenciais de vários planos, ultimamente gerando o plano final agregando múltiplas iterações. Estes métodos coletivamente contribuem para um framework de tomada de decisão mais robusto e confiável para agentes baseados em LLM. Enquanto alguns trabalhos focam em empregar mecanismos de feedback para corrigir erros nos processos de raciocínio e planejamento dos agentes, guiando-os a executar cadeias de raciocínio precisas, trabalhos anteriores podem ser categorizados em três fontes primárias de feedback: (1) reflexão interna do LLM baseada em memória; (2) feedback humano; (3) feedback ambiental. Quanto à primeira categoria, agentes baseados em LLM derivam insights de experiências históricas para atualizar ou otimizar estratégias e métodos de planejamento. Por exemplo, ReWOO (XU *et al.*, 2023a) introduz um paradigma onde planos e observações externas são gerados independentemente pelos agentes. Estes planos e observações derivados independentemente são então integrados para produzir os resultados finais. Estas metodologias coletivamente aprimoram as capacidades de tomada de decisão de agentes baseados em LLM aproveitando paradigmas de raciocínio estruturados, multi-caminho e sensíveis ao contexto, permitindo-lhes assim abordar tarefas complexas mais efetivamente. A integração de mecanismos de feedback assegura um processo de refinamento dinâmico e iterativo, crucial para alcançar raciocínio autônomo preciso e confiável em sistemas baseados em LLM.

Método de Múltiplos Passos.

Diferentemente do raciocínio de passo único, raciocínio de múltiplos passos requer invocação iterativa de LLMs para múltiplos ciclos de raciocínio, onde cada ciclo gera um ou vários passos incrementais baseados no contexto atual enquanto mantendo consistência com o objetivo geral. Raciocínio de múltiplos passos visa aprimorar a capacidade do LLM de resolver problemas complexos e entender tarefas de longo prazo através de processos de raciocínio estruturados. Esta abordagem assegura que o raciocínio e planejamento permaneçam adaptativos e responsivos a requisitos de tarefa e dinâmicas ambientais em evolução, facilitando assim capacidades robustas de tomada de decisão e resolução de problemas em sistemas baseados em LLM. É importante distinguir entre dois conceitos relacionados mas distintos que aparecem na literatura: sistemas **multiagentes** e comunicação **multimodal**. Sistemas multiagentes referem-se à arquitetura computacional onde múltiplos agentes autônomos colaboram para resolver problemas complexos, enquanto comunicação multimodal refere-se à capacidade de processar diferentes tipos de entrada de dados (texto, imagem, áudio, etc.) dentro de um mesmo sistema. Um sistema pode ser simultaneamente multiagente e multimodal, como observado em trabalhos que integram múltiplos agentes especializados capazes de processar diferentes modalida-

des de informação de forma coordenada. Métodos de múltiplos estágios dissecam o processo de planejamento em estágios distintos, visando melhorar o desempenho do LLM em tarefas complexas de raciocínio e resolução de problemas. O processo de exploração de DECKARD (NOTTINGHAM *et al.*, 2023) é dividido nos estágios de Sonho e Vigília. Durante o estágio de Sonho, o agente utiliza um LLM para decompor a tarefa em subobjetivos. No estágio de Vigília, o agente aprende uma estratégia modular para cada subobjetivo, verificando ou retificando suposições baseadas na experiência do agente.

Raciocínio e Planejador Externos.

Enquanto LLMs exibem capacidades poderosas de raciocínio e planejamento através de diversas aplicações, gerar planos precisos e eficientes para problemas de domínio específico coloca desafios significativos. Consequentemente, diversos estudos de pesquisa integraram LLMs com ferramentas externas para colaborativamente abordar desafios especializados. Estas ferramentas externas abrangem habilidades de domínio específico como APIs, modelos especialistas e técnicas envolvendo bancos de dados externos (KARPAS *et al.*, 2022; PATIL *et al.*, 2023), renomados por sua proficiência e alta precisão em domínios específicos. Aproveitando estas capacidades especializadas, agentes baseados em LLM equipados com planejadores externos podem gerar planos mais eficientes, e em alguns casos ótimos. Especificamente, CO-LLM (ZHANG *et al.*, 2023d) utiliza LLMs para gerar planos de alto nível para tarefas atuais, complementados por um modelo externo que refina estes planos em estratégias mais granulares. Por outro lado, LLM+P (LIU *et al.*, 2023a) transforma contextos de prompt contendo o estado atual do agente, observações ambientais e experiências históricas em Linguagens de Definição de Domínio de Planejamento (PDDL) formais. Subsequentemente, esta informação textual é alimentada a um raciocínio externo para inferência e a geração de arranjos detalhados de planejamento. Esta abordagem integrada aprimora as capacidades de planejamento de LLMs alavancando tanto sua proeza de raciocínio baseado em texto quanto a precisão de modelos de raciocínio externos adaptados a domínios específicos. Estes métodos aprimoram significativamente as capacidades adaptativas e perceptuais de agentes baseados em LLM na navegação de ambientes complexos, melhorando assim sua habilidade de planejar e se engajar em interações sofisticadas de resolução de problemas e colaboração. Empregando estas metodologias, agentes podem ser guiados em direção a processos mais eficientes, racionais e efetivos de raciocínio, planejamento e execução.

Generalização

As capacidades de generalização dos agentes LLM são críticas para sua efetividade através de uma ampla gama de ambientes dinâmicos e imprevisíveis. Generalização se manifesta especificamente na forma de transferibilidade, permitindo que agentes apliquem conhecimento aprendido em um domínio a outro, e robustez, suportando adaptação a diversas variações de entrada. Esta capacidade de generalização assegura que agentes LLM possam manter alto desempenho através de diferentes contextos sem extensivo retreinamento ou intervenção humana. A utilização das habilidades de generalização dos agentes LLM é refletida proeminentemente em áreas como aprendizado zero-shot, aprendizado few-shot e aprendizado many-shot. Tarefas não vistas referem-se àquelas que o agente não encontrou durante a fase de treinamento. A natureza dinâmica da maioria dos ambientes de aplicação necessita que modelos possuam a capacidade de responder efetivamente a situações imprevistas. Agentes LLM podem alavancar seu treinamento em larga escala em conjuntos de dados diversos para inferir e aplicar conhecimento relevante, permitindo-lhes adaptar-se a novas tarefas mais rapidamente e robustamente que modelos tradicionais. As abordagens de generalização para agentes LLM para tarefas não vistas podem ser amplamente categorizadas baseadas em se o modelo passou por ajuste fino.

Aprendizado em Contexto (ICL).

Aprendizado em contexto envolve fornecer exemplos da tarefa atual dentro do prompt de entrada, permitindo que o modelo use estes exemplos para inferir os requisitos da tarefa e gerar respostas apropriadas (DONG *et al.*, 2022). Este método foi destacado no trabalho de (BROWN *et al.*, 2020), demonstrando a habilidade do GPT-3 de aprender a realizar tarefas complexas através de exemplos no contexto. A vantagem do ICL é que não requer atualizações de parâmetro, tornando-o computacionalmente eficiente e fácil de implementar. No entanto, o desempenho do modelo é sensível a configurações específicas, incluindo a seleção de templates de prompt, a escolha de exemplares contextuais, e a sequência de exemplos, e exibe uma propensão a prever respostas que frequentemente ocorrem na conclusão de prompts ou são prevalentes no conjunto de dados de pré-treinamento (ZHAO *et al.*, 2021).

Aprendizado Zero-Shot.

Aprendizado zero-shot requer que o modelo desempenhe novas tarefas sem quaisquer exemplos específicos de tarefa ou ajuste fino, confiando inteiramente em seu conhecimento pré-treinado. Em (RADFORD *et al.*, 2019), demonstrou-se aprendizado zero-shot com GPT-2, onde o modelo mostrou a habilidade de lidar com várias tarefas sem treinamento específico de tarefa prévio. Este método destaca as capa-

idades de generalização inerentes do modelo e não requer dados ou treinamento adicional. No entanto, o desempenho do aprendizado zero-shot pode ser limitado para tarefas altamente especializadas ou complexas, pois o modelo pode carecer do conhecimento específico necessário para executá-las efetivamente.

3.2.10 Ação

Ações representam os resultados comportamentais tangíveis de agentes dentro de um ambiente interativo, efetuando assim mudanças reais no ambiente e impactando significativamente as interações entre agentes. Estas ações são tipicamente determinadas por uma combinação de perfis, memória e o contexto interativo (incluindo interações agente-para-agente, agente-para-ambiente e agente-para-humano). Situada na posição mais downstream, ações variam amplamente dependendo do cenário de aplicação. O mecanismo de ação pode ser elucidado de duas perspectivas: o processo de criação de ação e a aplicação de ações: (1) Criação de Ação: Isto envolve os processos e passos pelos quais ações são geradas. Engloba os frameworks de tomada de decisão, algoritmos e procedimentos que levam à formulação de ações específicas baseadas no estado interno do agente e estímulos externos. (2) Aplicação de Ação: Isto refere-se aos contextos nos quais ações são aplicadas e os efeitos subsequentes destas ações nos cenários de aplicação.

Criação de Ação

Criação de ação representa o estágio final onde agentes manifestam sua inteligência dentro de ambientes interativos de sistemas multi-agente. À medida que as dinâmicas ambientais flutuam e diretrizes de tarefa variam, agentes empregam estratégias diversas e fontes de informação para promulgar ações alinhadas com os objetivos abrangentes do sistema. Baseado nonexo temporal entre tomada de decisão e execução de ação em ambientes interativos, delinearemos três estratégias prevalentes para criação de ação.

Decisão de Passo Único.

Primeiramente, tomada de decisão instantânea envolve agentes extraindo informação recente, pertinente e significativa de seus bancos de memória. Quando necessário, agentes suplementam esta informação acessando bases de conhecimento externas. Guiados por prompts derivados da amalgamação de requisitos de tarefa atuais, recordações de memória e conhecimento externo, agentes prontamente formulam planos e executam ações correspondentes. Por exemplo, Generative agents (PARK *et al.*, 2023) mantêm um fluxo contínuo de memória, usando informações recentes e relevantes para guiar suas ações. Similarmente, em GITM (ZHU *et al.*,

2023b), agentes consultam sua memória para identificar experiências bem-sucedidas relevantes para alcançar subobjetivos de baixo nível, replicando ações efetivas de tarefas anteriores. Agentes colaborativos como ChatDev (QIAN *et al.*, 2023) e MetaGPT (HONG *et al.*, 2023) engajam-se em interações de diálogo onde históricos conversacionais armazenados na memória influenciam as declarações de cada agente. Estas estratégias ressaltam a capacidade adaptativa dos agentes de dinamicamente integrar informações internas e externas, facilitando tomada de decisão efetiva e execução de ação responsiva em ambientes interativos complexos.

Planejamento Pré-definido.

Nesta estratégia, cada ação empreendida por agentes baseados em LLM adere estritamente a planejamento pré-definido, que pode ser autonomamente gerado pelo agente ou predefinido por usuários. Por exemplo, em DEPS (WANG *et al.*, 2023e), agentes iniciam planejamento de ação para uma tarefa específica e procedem com a execução a menos que indicações de falha no plano emergjam durante o processo. Este método assegura que agentes mantenham consistência e aderência a cursos de ação planejados ao longo de suas sequências operacionais.

Criação Dinâmica.

Esta estratégia representa uma síntese das duas abordagens precedentes, efetivamente equilibrando a natureza pré-definida do planejamento de tarefa com adaptabilidade a ambientes dinâmicos. Inicialmente configurados com um plano de objetivo abrangente, agentes geram um plano de objetivo abrangente. Subsequentemente, durante interações, agentes aderem a estes objetivos abrangentes enquanto retendo a flexibilidade para tomar decisões instantâneas baseadas no ambiente interativo. Em GITM (ZHU *et al.*, 2023b), por exemplo, agentes formulam planos de alto nível decompondo tarefas em múltiplos subobjetivos. Estes planos guiam a execução sequencial de ações visando abordar cada subobjetivo, ultimamente alcançando a conclusão da tarefa geral.

Aplicação de Ação

O contexto de ações tipicamente passa por mudanças dinâmicas baseadas em cenários de aplicação específicos. Aplicação de ação refere-se à interação e influência direta entre agentes e seu ambiente, onde os resultados de seus comportamentos impactam diretamente a realização de tarefas atuais e a progressão geral de sistemas multi-agente. Dependendo dos diversos cenários de interação encontrados por agentes, delinearemos estes aspectos através de três dimensões:

Orientado a Tarefa.

Neste cenário, as ações de agentes baseados em LLM visam realizar subtarefas específicas, que coletivamente contribuem para a conclusão de tarefas abrangentes maiores através de divisão colaborativa de trabalho entre agentes. Aproveitando as capacidades de planejamento inerentes em LLMs, DEPS (WANG *et al.*, 2023e) desenvolveu um agente Minecraft capaz de resolver tarefas complexas decompondo-as em subobjetivos manejáveis. Sistemas similares como GITM (ZHU *et al.*, 2023b) e Voyager (WANG *et al.*, 2023a) também se baseiam extensivamente nas habilidades de planejamento dos LLMs para navegar com sucesso e realizar tarefas diversas. TaskMatrix.AI (LIANG *et al.*, 2023) integra LLMs com milhões de APIs para facilitar execução de tarefa. Em seu núcleo está um modelo fundacional conversacional multimodal que se engaja com usuários, compreende seus objetivos e contexto, e subsequentemente gera código executável adaptado a tarefas específicas.

Interação Comunicativa.

A tarefa primária da interação de agentes gira em torno de se engajar em discussões sobre um tópico específico para trocar ideias ou fomentar inovação. Por exemplo, agentes em ChatDev (QIAN *et al.*, 2023) colaboram através de comunicação para coletivamente realizar tarefas de desenvolvimento de software.

Exploração Ambiental.

Exploração ambiental primariamente implica agentes colaborando para explorar e adaptar-se a ambientes dinamicamente mutáveis, assim expandindo suas capacidades perceptuais e conjuntos de habilidades. Por exemplo, o agente em Voyager (WANG *et al.*, 2023a) se engaja na exploração de habilidades desconhecidas durante a conclusão de tarefa, continuamente refinando a execução destas habilidades baseado em feedback ambiental através de tentativa e erro iterativos. Após a execução de ações pelos agentes, alguns estudos consideram o impacto direto no ambiente interativo e tentam integrar perfeitamente aplicações downstream com as ações dos agentes. Esta integração primariamente envolve incorporar LLMs com a utilização de ferramentas externas ou conhecimento. Especificamente, inclui APIs, calculadoras, interpretadores de código, modelos projetados por especialistas e bases de conhecimento externas (KARPAS *et al.*, 2022; SCHICK *et al.*, 2024; SONG *et al.*, 2023c). Aproveitando estes recursos externos, agentes podem aprimorar seus processos de tomada de decisão e melhorar a eficiência e precisão da execução de tarefa. Esta abordagem expande ainda mais o escopo de aplicação e capacidades do módulo de ação, facilitando planejamento estratégico mais direto e utilização de ferramentas em aplicações downstream. Aprimora a habilidade do sistema de agente de adaptar-

se a novas situações e alavancar ferramentas novas, assim ampliando o potencial para execução de tarefa efetiva e eficiente.

3.3 Evolução

Assim como os humanos continuamente refinam suas habilidades cognitivas e adquirem conhecimento através de interações com seu ambiente e outros, a evolução em agentes envolve a reflexão contínua sobre suas decisões e ações para atualizar dinamicamente seu conhecimento e experiências, baseando-se em experiências existentes e feedback recebido durante interações, como visualizado na Figura 3.5.

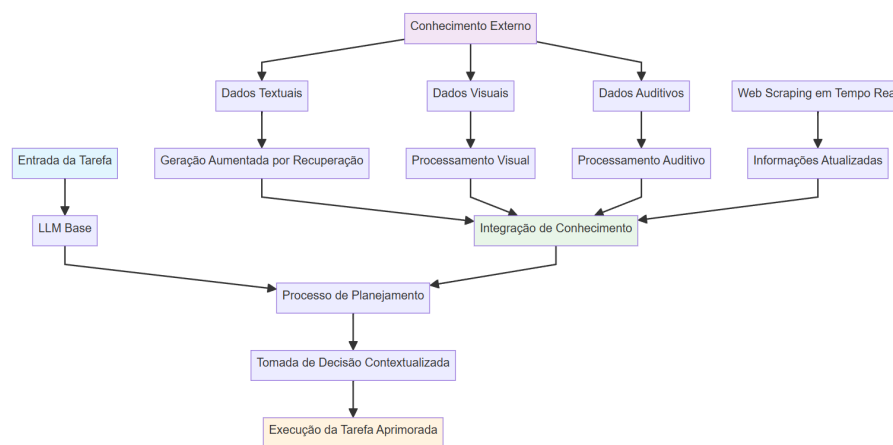


Figura 3.5: Representação do ciclo de evolução contínua em agentes, mostrando como feedback proveniente do ambiente, outros agentes e humanos é processado através de três mecanismos principais (atualização de memória, autorreflexão e geração dinâmica) para aprimoramento iterativo das capacidades do agente

Ao adotar mecanismos de evolução, agentes podem continuamente refinar ou revisar seu entendimento atual, aprofundando assim sua proficiência em tarefas conhecidas e expandindo sua exploração bem-sucedida de tarefas desconhecidas. Considerando as fontes de feedback externo obtidas durante interações, trabalhos existentes podem ser categorizados em três tipos principais: informações percebidas do ambiente circundante, trocadas com outros agentes, ou transmitidas por humanos. Para equipar agentes com estas diversas fontes de informação, vários métodos têm sido empregados para aprimorar suas capacidades de evolução. Nas seções seguintes, fornecemos uma introdução detalhada a cada uma destas abordagens, elucidando as técnicas usadas para reforçar o processo de evolução em agentes.

3.3.1 Fonte de Evolução

O feedback recebido durante interações serve como informação de referência indispensável para que agentes alcancem evolução. Este feedback engloba os resultados e impactos das decisões e ações dos agentes, guiando-os para introspecção e, assim, melhorando dinamicamente sua adaptação a ambientes ou tarefas complexos. Trabalhos anteriores predominantemente capturaram e transmitiram feedback em forma textual (WANG *et al.*, 2023d,e; YAO *et al.*, 2023a). Com base nas fontes das quais os agentes recebem este feedback, ele pode ser categorizado em três tipos distintos. Cada fonte fornece insights únicos que contribuem para os processos de autorreflexão e melhoria contínua dos agentes.

Feedback do Ambiente

Feedback do ambiente refere-se às informações percebidas por agentes dentro de ambientes reais ou virtuais. Este tipo de feedback geralmente pertence a informações em mudança no ambiente resultantes das decisões e ações dos agentes durante suas interações com o entorno externo. Tal feedback atua como um sinal de recompensa, informando agentes sobre as consequências de suas ações. Este mecanismo é vividamente demonstrado em planejamento de tarefas complexas e simulações robóticas dentro de cenários ambientais dinâmicos (SHINN *et al.*, 2023; WANG *et al.*, 2024a). Ao incorporar estas mudanças ambientais como feedback, agentes podem refinar suas estratégias e ações, melhorando assim sua adaptabilidade e desempenho em ambientes em tempo real e simulados.

Interação entre Agentes

Em sistemas multi-agente, informações de interação entre agentes envolvem a troca de informações colaborativas entre agentes. Estas informações tipicamente incluem avaliações ou atualizações de status de outros agentes sobre as decisões ou ações de um agente em particular, bem como comunicação contextual entre agentes. Servindo como sinais internos, estas informações de interação facilitam coerência e integração entre agentes, aprimorando e expandindo continuamente as capacidades colaborativas do sistema multi-agente. Isso é particularmente evidente na execução hierárquica de tarefas e comunicação de agentes dentro de simulações de mundo (ZHONG *et al.*, 2022). Através de tais trocas, agentes podem refinar sua coordenação e melhorar o desempenho geral do sistema em cenários complexos e dinâmicos.

Feedback Humano

Além dos feedbacks ambientais e de interação entre agentes mencionados anteriormente, o feedback humano constitui um sinal orientador fornecido por humanos

para direcionar agentes a tomar melhores decisões e ações, aprimorando assim suas capacidades cognitivas. Como um sinal subjetivo, o feedback humano efetivamente alinha agentes com valores e preferências humanas e ajuda a mitigar problemas como alucinação. Este tipo de feedback é extensivamente utilizado em sistemas onde agentes colaboram e se comunicam com humanos (FISCHER, 2023; WANG *et al.*, 2021), assegurando que as ações e decisões dos agentes estejam em harmonia com expectativas e padrões humanos.

3.3.2 Métodos de Evolução

Métodos de evolução abrangem uma variedade de técnicas projetadas para aprimorar as capacidades e adaptabilidade de agentes através de automelhoria e aprendizado a partir de interações com seu ambiente. Estes métodos são cruciais para desenvolver sistemas inteligentes que possam refinar autonomamente suas estratégias e comportamentos para alcançar melhor desempenho em diversas tarefas e cenários. A seção abaixo mergulha em várias abordagens-chave, incluindo aprendizado por feedback, ajuste fino supervisionado, engenharia de prompt e aprendizado por reforço, cada uma contribuindo distintivamente para a trajetória evolutiva de agentes inteligentes.

Ajuste Fino

Ajuste fino envolve atualizar os parâmetros de um modelo pré-treinado para adaptá-lo a novas tarefas ou domínios. Este método assegura que o modelo seja especificamente adaptado para novos desafios. Existem três categorias principais de métodos de ajuste fino: ajuste fino de modelo completo, ajuste fino de parâmetro pré-treinado parcial e ajuste fino de parâmetro adicional:

Ajuste Fino Completo. Ajuste fino completo envolve atualizar todos os parâmetros do modelo pré-treinado para adaptá-lo a tarefas novas específicas. Como observado em FireAct (CHEN *et al.*, 2023a), ajuste fino de modelo completo pode ser mais ótimo, particularmente quando aprendizado profundo do modelo para tarefas específicas é requerido, desde que recursos permitam. No entanto, é computacionalmente caro e demorado, e quando dados de nova tarefa são limitados, há risco de sobreajuste.

Repropósito. Repropósito tipicamente foca em ajustar camadas específicas de um modelo pré-treinado, geralmente as camadas superiores, enquanto mantém as camadas inferiores inalteradas (BROCK *et al.*, 2017; LIU *et al.*, 2021b; ZHU *et al.*, 2023a). Adicionalmente, BitFit (ZAKEN *et al.*, 2021) demonstra que, ajustando

apenas os termos de viés do modelo ou um subconjunto deles, desempenho comparável ou até melhor que ajuste fino de modelo completo pode ser alcançado em conjuntos de dados de treinamento pequenos a médios. Similarmente, SIFT (SONG *et al.*, 2023b) propõe alavancar a esparsidade do gradiente do modelo em tarefas downstream atualizando apenas os parâmetros-chave que contribuem mais significativamente para a norma do gradiente. Embora repropósito aumente a eficiência, pode não corresponder ao desempenho de ajuste fino de parâmetro completo quando se aprofunda em tarefas específicas (ZAKEN *et al.*, 2021). Além disso, a seleção de parâmetros ou camadas para atualizar é frequentemente baseada em regras heurísticas, o que pode requerer mais pesquisa para otimizar o processo de seleção.

Ajuste Fino de Parâmetro Adicional. Ajuste fino de parâmetro adicional introduz um conjunto extra de parâmetros ao modelo original, permitindo ajuste fino eficiente sem alterar os parâmetros pré-treinados.

1. *Adaptador*: Treinamento de adaptador introduz pequenas estruturas de rede neural, conhecidas como adaptadores, entre as camadas do modelo pré-treinado. Durante o ajuste fino, apenas estes adaptadores são treinados enquanto os parâmetros do modelo original permanecem inalterados. Especificamente, adaptadores podem ser integrados em várias camadas do modelo de maneira serial, paralela ou reparametrizada (HE *et al.*, 2021; HOULSBY *et al.*, 2019; HU *et al.*, 2023b; PFEIFFER *et al.*, 2020), e ao ajustar os parâmetros destes adaptadores, o desempenho do modelo em tarefas específicas é aprimorado enquanto mantém a capacidade de generalização do modelo. No entanto, seu desempenho é limitado pela capacidade dos adaptadores e pode não capturar totalmente a complexidade de tarefas altamente especializadas.
2. *Adaptação de Baixo Posto (LoRA)*: LoRA (HU *et al.*, 2021) envolve adicionar matrizes de baixo posto aos parâmetros do modelo e então ajustar estas matrizes para adaptar a novas tarefas. QLORA (DETTMERS *et al.*, 2024) reduz a memória requerida para ajuste fino de grandes modelos de linguagem sem sacrificar desempenho ao introduzir LoRA em modelos de linguagem pré-treinados congelados e quantizados. Isso exemplifica a eficiência de LoRA em recursos computacionais e memória. No entanto, seu desempenho pode ser ligeiramente inferior ao ajuste fino de modelo completo para tarefas que requerem modificações extensivas.
3. *Ajuste de Prefixo*: Ajuste de prefixo adapta-se a várias tarefas adicionando vetores de prefixo específicos de tarefa à entrada do modelo. Por exemplo, em (LI e LIANG, 2021), demonstra-se que, otimizando estes prefixos, é possível alcançar desempenho comparável ao ajuste fino de parâmetro completo com

significativamente menos parâmetros. No entanto, prefixos de comprimento fixo podem ser insuficientes para abordar a diversidade de tarefas. Para endereçar isso, APT (ZHANG *et al.*, 2023h) emprega um mecanismo de portão para ajustar dinamicamente os prefixos, aprimorando a eficiência e eficácia do ajuste fino, embora sua aplicabilidade a arquiteturas não-Transformer seja limitada. A vantagem do ajuste de prefixo reside em reduzir o número de parâmetros, mas pode requerer ajustes específicos de tarefa aos prefixos, e seu desempenho ainda pode ser limitado para certas tarefas.

4. *Ajuste de Prompt*: Ajuste de prompt adapta LLMs pré-treinados a tarefas específicas introduzindo "prompts suaves" treináveis (LESTER *et al.*, 2021). Este método alavanca retropropagação para otimizar os prompts enquanto mantém o resto do modelo congelado. Por exemplo, P-Tuning (LIU *et al.*, 2021a) estabiliza o processo de treinamento combinando embeddings de prompt contínuos com prompts discretos e alcançou melhorias de desempenho significativas em tarefas de compreensão de linguagem natural como LAMA (PETRONI *et al.*, 2019) e SuperGLUE (WANG *et al.*, 2019). Embora o ajuste de prompt seja favorecido por sua eficiência de parâmetro e reusabilidade de modelo, pode requerer prompts cuidadosamente projetados e um entendimento profundo da tarefa, e pode não corresponder totalmente à eficácia do ajuste fino de parâmetro completo para algumas tarefas complexas.

Aprendizado por Feedback

Aprendizado por feedback é uma abordagem que emprega informações de feedback como contexto, permitindo que um agente "reforce" a geração de política iterativamente sem a necessidade de atualizar pesos. Informações de feedback podem assumir múltiplas formas, como contextos de prompt (BROWN *et al.*, 2020), embeddings (LESTER *et al.*, 2021), tokens (BAI *et al.*, 2022; LIU *et al.*, 2021a; MADAAN *et al.*, 2024). Reflexion (SHINN *et al.*, 2023) é um mecanismo inovador de aprendizado por feedback que permite que agentes de linguagem reforcem o aprendizado recebendo feedback verbal, ao invés de através de atualizações de peso. O agente reflete sobre sinais de feedback de tarefa e armazena os resultados da reflexão como texto em memória episódica, guiando processos futuros de tomada de decisão e assim melhorando o desempenho em tentativas sucessivas. InstructGPT (OUYANG *et al.*, 2022) aprende coletando avaliações de anotadores humanos sobre a saída do modelo, que incluem classificações de preferência para o texto gerado pelo modelo, servindo como um sinal de feedback. Similarmente, DPO (RAFAILOV *et al.*, 2024) ajusta diretamente o comportamento do modelo baseado em classificações de preferência do usuário, oferecendo uma otimização mais direcionada ao alinhar saídas com feedback humano de maneira computacionalmente eficiente.

Engenharia de Prompt

Engenharia de prompt é um método que utiliza prompts bem projetados e feedback como dicas contextuais. Por exemplo, Retroformer (YAO *et al.*, 2023a) permite que um agente reflita sobre suas falhas passadas, integrando estas reflexões em prompts para guiar ações futuras. Engenharia de prompt tem uma ampla gama de aplicações em grandes modelos de linguagem. Por exemplo, AutoPrompt (SHIN *et al.*, 2020) aprimora o desempenho do GPT-3 em tarefas específicas gerando prompts personalizados, assim melhorando a qualidade de sua saída. A abordagem AutoPrompt demonstra que, automatizando a geração e otimização de prompts, o desempenho de modelos de linguagem em tarefas específicas pode ser significativamente melhorado. O núcleo deste método reside na geração automática de prompts, que através de ajuste e otimização contínuos, permite que o modelo entenda melhor requisitos de tarefa e produza saídas de alta qualidade. Prefix-tuning (LI e LIANG, 2021) é outra técnica de engenharia de prompt que envolve adicionar prefixos a prompts, permitindo que o modelo de linguagem entenda e execute melhor tarefas específicas. Prefix-tuning mostra que, otimizando prompts sem mudar os pesos do modelo, o desempenho do modelo pode ser significativamente aprimorado. Este método adiciona prefixos específicos a prompts de entrada, permitindo que o modelo reference mais informações contextuais durante a geração, assim melhorando a relevância e precisão da saída.

Aprendizado por Reforço

No aprendizado por reforço, um agente aprende a estratégia ótima através de interação com o ambiente. Cada ação produz feedback correspondente (como recompensas ou penalidades), e o agente continuamente ajusta sua estratégia baseado neste feedback para maximizar recompensas cumulativas. O núcleo do aprendizado por reforço reside em tentativa e erro e otimização, onde o agente gradualmente aprende a fazer decisões ótimas em diferentes contextos através de múltiplas tentativas e erros. Por exemplo, ICPI (WU *et al.*, 2024b) aprende em contexto usando grandes modelos de linguagem para realizar iteração de política sem demonstrações de especialistas ou atualizações de gradiente, melhorando estratégias através de interação por tentativa e erro. InstructGPT (OUYANG *et al.*, 2022), por outro lado, ajusta o GPT-3 através de aprendizado por reforço com feedback humano, tornando-o melhor em seguir instruções de usuário e melhorando seu alinhamento e desempenho através de várias tarefas.

3.3.3 Ajuste de Agentes

Um aspecto-chave do mecanismo de evolução é a atualização contínua do conhecimento e experiências existentes dos agentes, ou o refinamento de decisões e comportamentos atuais antes da execução. Este processo visa aprofundar as capacidades cognitivas dos agentes e aprimorar sua responsividade a ambientes complexos e dinâmicos. Através de aprendizado e adaptação iterativos, agentes podem melhorar seu desempenho e manter relevância em contextos em constante mudança.

Atualização de Memória

Uma abordagem significativa enfatiza a expansão e aprofundamento da autoconsciência e experiências de aprendizado dos agentes. Este método geralmente envolve agentes utilizando mecanismos de memória para se engajar em autorreflexão baseada em feedback coletado, através de processos de abstração, sumarização e síntese. O conhecimento e experiências recém-adquiridos são então armazenados na memória ou em uma base de dados externa. Por exemplo, em GITM (ZHU *et al.*, 2023b), o agente inicialmente faz explorações no ambiente de interação. Ao realizar com sucesso uma tarefa, o agente armazena as ações usadas em sua memória. Similarmente, em AppAgent (YANG *et al.*, 2023c), o agente aprende através de uma abordagem dual de exploração autônoma e observação de demonstrações humanas. Este processo iterativo facilita a construção de uma base de conhecimento abrangente, que subsequentemente serve como referência para executar tarefas complexas através de diversas aplicações móveis. Em MemPrompt (MADAAN *et al.*, 2022a), feedback em linguagem natural de usuários sobre as intenções de resolução de problemas do agente é capturado e armazenado na memória. Subsequentemente, quando o agente confronta tarefas análogas, ele acessa estas memórias armazenadas para formular respostas mais apropriadas.

Autorreflexão

Enquanto pesquisas anteriores predominantemente focaram em aprimorar as capacidades dos agentes para tomada de decisão de tarefa zero-shot e execução eficiente, uma abordagem geral envolve agentes evoluindo dinamicamente ao adaptar seus objetivos iniciais e estratégias de planejamento baseado em feedback e registros de comunicação. (ZHANG *et al.*, 2024e) aproveita a função vantagem avaliada por um crítico como feedback, e revisa o plano para interação mais eficiente. MemoryBank (ZHONG *et al.*, 2024) realiza processamento de conversação para destilar eventos diários em resumos concisos semelhantes à consolidação de memória humana de experiências significativas. Através de interações contínuas, agentes continuamente avaliam e aprimoram sua base de conhecimento, gerando insights diários sobre traços

de personalidade em evolução.

Geração Dinâmica

Em certos contextos, o foco está na manutenção autônoma de sistemas multi-agente para assegurar sua operação contínua. Dada a complexidade do ambiente, o sistema pode ajustar dinamicamente sua escala gerando ou removendo agentes específicos de tarefa. Por exemplo, em (CHEN *et al.*, 2023c, 2024c), eles permitem que o sistema escale e adapte efetivamente seus recursos, implantando agentes especificamente criados para abordar demandas e desafios operacionais atuais.

Capítulo 4

Proposta

4.1 Metodologia

Este capítulo apresenta os procedimentos metodológicos adotados para desenvolvimento do artefato proposto, incluindo frameworks de avaliação (TAM, TTF, AX e ADDIE), estrutura experimental, caracterização da amostra, instrumentos de coleta de dados e métodos de análise aplicados nos estudos de caso realizados.

O processo de DSR aplicado nesta pesquisa segue as etapas propostas por (DRESCH *et al.*, 2020), adaptadas ao contexto específico do desenvolvimento de um sistema multiagente para o design de jogos educativos. A estruturação completa do processo está apresentada na Tabela 4.1.

O processo de DSR aplicado nesta pesquisa segue as etapas propostas por (DRESCH *et al.*, 2020), adaptadas ao contexto específico do desenvolvimento de um sistema multiagente para o design de jogos educativos. A estruturação completa do processo está apresentada na Tabela 4.1.

Contexto da Pesquisa

O contexto desta pesquisa situa-se no domínio do design de jogos educacionais, área que tem experimentado crescimento significativo devido ao reconhecimento do potencial pedagógico dos jogos digitais. O cenário atual caracteriza-se pela crescente demanda por ferramentas educacionais digitais interativas que possam engajar estudantes de forma efetiva, ao mesmo tempo em que atendem aos objetivos curriculares estabelecidos.

O desenvolvimento de jogos educacionais tradicionalmente requer conhecimento multidisciplinar, envolvendo competências em design de jogos, pedagogia, psicologia da aprendizagem e desenvolvimento de software. Esta multiplicidade de conhecimentos necessários cria barreiras significativas para educadores que desejam criar seus próprios recursos digitais, resultando em dependência de desenvolvedores especializados ou soluções genéricas que podem não atender às necessidades específicas

de contextos educacionais particulares.

A emergência de tecnologias baseadas em Large Language Models (LLMs) apresenta oportunidades inéditas para automatização e assistência inteligente em processos criativos complexos, incluindo o design educacional. Estas tecnologias oferecem potencial para democratizar o acesso a ferramentas de criação de jogos educacionais, permitindo que educadores sem conhecimento técnico especializado possam desenvolver recursos pedagógicos digitais customizados.

A conscientização do problema envolveu uma compreensão das dificuldades enfrentadas por designers educacionais, incluindo:

- Complexidade na integração de elementos lúdicos com objetivos pedagógicos específicos
- Tempo considerável necessário para o design de jogos educativos de qualidade
- Necessidade de conhecimento especializado tanto em design de jogos quanto em pedagogia
- Escassez de ferramentas que facilitem o processo de design educacional
- Desafios na classificação e definição de objetivos educacionais adequados

Esta etapa foi fundamental para estabelecer os requisitos iniciais da solução proposta e delimitar o escopo da pesquisa.

4.1.1 Proposição e Projeto do Artefato

Seguindo a Design Science Research e com base na conscientização do problema, foram propostos os seguintes artefatos como componentes de um sistema integrado a LLMs, onde ocorrerá:

1. **Módulo de Brainstorming:** Sistema para geração e organização de ideias iniciais para jogos educacionais, com suporte de LLMs.
2. **Módulo de Agente Sócrático:** Implementação de um agente especializado em aplicar o método socrático para refinamento de ideias e identificação de problemas potenciais.
3. **Módulo de Taxonomia de Bloom:** classificação e definição de objetivos educacionais baseados na Taxonomia de Bloom com auxílio de agente especialista.
4. **Módulo de Endo-GDC:** Preenchimento assistido do Game Design Canvas para Jogos Educativos Endógenos, com múltiplos agentes especializados.

O projeto detalhado do sistema foi desenvolvido seguindo princípios de arquitetura de software orientada a serviços e sistemas multiagentes, como discutido por (XI *et al.*, 2025) e adaptado às necessidades específicas desta pesquisa.

4.2 Desenvolvimento do Artefato

O desenvolvimento do sistema seguiu uma abordagem iterativa e incremental, com ciclos de implementação, teste e refinamento. Esta abordagem permitiu ajustes contínuos com base nos resultados preliminares e feedback dos usuários, alinhando-se às recomendações de (DRESCH *et al.*, 2020) para o desenvolvimento de artefatos complexos. O processo de desenvolvimento incluiu as seguintes etapas:

- **Implementação da infraestrutura básica:** desenvolvimento da camada de persistência, APIs e interface de usuário.
- **Integração com APIs de LLM:** estabelecimento de conexões com modelos de linguagem de grande escala, implementação de gerenciamento de contexto e processamento de respostas.
- **Desenvolvimento dos agentes especializados:** implementação de agentes com diferentes perfis e responsabilidades, incluindo seus prompts e lógica de funcionamento.
- **Sistema de coordenação:** desenvolvimento de mecanismos para coordenação entre agentes e manutenção da coerência global.
- **Refinamento e otimização:** ajustes no sistema com base em testes iniciais, incluindo melhorias de desempenho e usabilidade.

A implementação técnica do sistema utilizou as tecnologias detalhadas na Seção 3.4, escolhidas por sua adequação aos requisitos do projeto e alinhamento com as práticas atuais de desenvolvimento de sistemas baseados em IA.

4.3 Avaliação do Artefato

A avaliação do artefato foi realizada através de múltiplos estudos de caso, aplicando o sistema em diferentes contextos educacionais. Esta abordagem multifacetada permitiu uma compreensão abrangente das capacidades, limitações e impacto do sistema proposto. Os estudos de caso abrangeram diversas áreas do conhecimento e níveis educacionais, incluindo:

- Ensino de matemática no ensino fundamental
- Ensino de ciências no ensino médio
- Ensino de idiomas para adultos
- Formação em liderança e gestão
- Ensino de história no ensino médio
- Ensino de teoria musical
- Entre outros contextos educacionais

Para cada estudo de caso, o sistema foi utilizado para desenvolver um conceito de jogo educacional completo, desde a geração de ideias iniciais até a definição

de objetivos educacionais baseados na Taxonomia de Bloom. Os resultados foram documentados e analisados segundo critérios de qualidade pedagógica, coerência do design e eficiência do processo.

Este capítulo apresenta o sistema multiagente proposto para design de jogos educativos, detalhando a arquitetura de quatro módulos funcionais especializados, descrevendo os componentes do sistema, os fluxos de trabalho, as estruturas de dados e os mecanismos de interação entre agentes.

4.4 Arquitetura Geral do Sistema

A arquitetura do sistema foi concebida para suportar múltiplos agentes especializados que colaboram no processo de design de jogos educativos endógenos. Esta arquitetura baseia-se em princípios de sistemas distribuídos e incorpora elementos de arquiteturas orientadas a serviços, permitindo a escalabilidade e a extensibilidade do sistema.

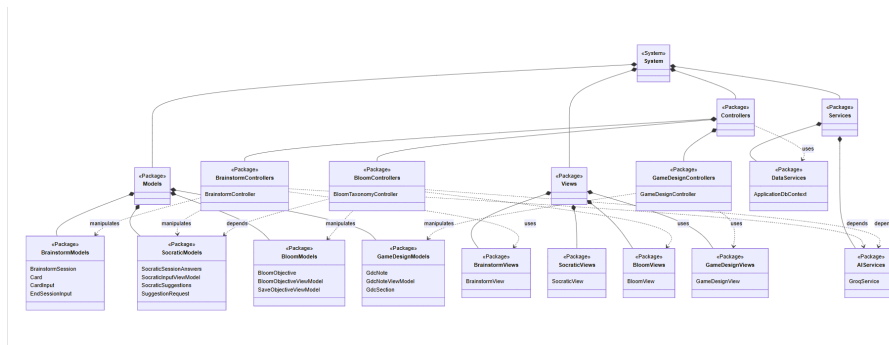


Figura 4.1: Diagrama de Pacotes do sistema multi-agente para design de jogos educativos.

A Figura 4.1 ilustra a visão geral da arquitetura do sistema, mostrando os principais componentes e suas interconexões. O sistema é estruturado em quatro camadas principais: apresentação, controle, serviços e persistência. Esta organização em camadas proporciona separação de responsabilidades e facilita a manutenção e extensão do sistema.

4.4.1 Camada de Apresentação

A camada de apresentação implementa a interface com o usuário através de componentes web que seguem o padrão MVC (Model-View-Controller). Esta camada inclui os seguintes componentes:

- **Views:** Implementam a interface visual com o usuário, incluindo páginas web e componentes interativos.

- **Validadores de Formulário:** Realizam verificação de dados inseridos pelo usuário antes da submissão ao backend.
- **Componentes JavaScript:** Implementam funcionalidade dinâmica no lado do cliente, permitindo interações em tempo real.
- **Componentes de Visualização de Dados:** Apresentam informações e resultados de forma visualmente compreensível.

A implementação da camada de apresentação utiliza tecnologias web padrão, incluindo HTML5, CSS3 e JavaScript. Os frameworks React e Bootstrap são utilizados para criar uma interface responsiva e consistente, enquanto bibliotecas como D3.js são utilizadas para visualizações de dados.

4.4.2 Camada de Controle

A camada de controle gerencia o fluxo de dados entre a camada de apresentação e a camada de serviços. Esta camada implementa a lógica de negócios e coordena as operações do sistema. Os principais componentes incluem:

- **Controllers:** Implementam a lógica de controle para cada módulo funcional do sistema, processando requisições do usuário e coordenando as respostas.
- **Filtros de Requisição:** Verificam e validam requisições antes do processamento.
- **Gerenciadores de Sessão:** Mantêm o estado da sessão do usuário durante o processo de design.
- **Mecanismos de Autorização:** Controlam o acesso a recursos do sistema.

Os controllers implementam o padrão de design MVC, recebendo requisições da camada de apresentação, processando-as através da camada de serviços e retornando resultados formatados para exibição. Esta abordagem facilita a separação de responsabilidades e melhora a manutenibilidade do código.

4.4.3 Camada de Serviços

A camada de serviços contém os componentes responsáveis por implementar a funcionalidade principal do sistema, incluindo a comunicação com os agentes baseados em LLM. Esta camada inclui:

- **Serviços de Agente:** Implementam a lógica para cada agente especializado no sistema.
- **Serviços de Integração com LLM:** Gerenciam a comunicação com modelos de linguagem de grande escala.
- **Serviços de Coordenação:** Coordenam a colaboração entre múltiplos agentes.
- **Serviços de Análise:** Processam e analisam dados do processo de design.

A camada de serviços é implementada seguindo o padrão de injeção de dependência, permitindo a substituição de componentes sem afetar o resto do sistema. Esta abordagem facilita os testes unitários e melhora a flexibilidade do sistema.

4.4.4 Camada de Persistência

A camada de persistência é responsável pelo armazenamento e recuperação de dados, implementando o modelo de dados do sistema. Esta camada inclui:

- **Repositórios:** Implementam a lógica de acesso a dados para cada entidade do sistema.
- **Contexto de Dados:** Define o mapeamento entre objetos de domínio e tabelas do banco de dados.
- **Migrações:** Gerenciam a evolução do esquema do banco de dados.
- **Caches:** Armazenam dados frequentemente acessados para melhorar o desempenho.

A camada de persistência utiliza o Entity Framework Core como ORM (Object-Relational Mapper), facilitando a implementação do padrão de repositório e abstraindo os detalhes do banco de dados subjacente. O sistema suporta múltiplos provedores de banco de dados, incluindo SQL Server e PostgreSQL.

4.5 Sistema Multiagente

O núcleo da arquitetura proposta é o sistema multiagente que integra LLMs especializados para diferentes aspectos do processo de design de jogos educacionais. Esta abordagem permite a decomposição de tarefas complexas em subtarefas que podem ser abordadas por agentes com expertise específica.

4.5.1 Modelo de Agente

Cada agente no sistema segue um modelo comum que define sua estrutura e comportamento. Este modelo inclui os seguintes componentes:

- **Perfil:** Define as características e especialidade do agente.
- **Módulo de Percepção:** Captura informações do ambiente, incluindo entradas do usuário e resultados de outros agentes.
- **Módulo de Raciocínio:** Processa informações percebidas e toma decisões baseadas em sua especialidade.
- **Módulo de Atuação:** Executa ações baseadas nas decisões tomadas.
- **Módulo de Comunicação:** Permite a troca de informações com outros agentes.

- **Módulo de Aprendizado:** Permite que o agente melhore seu desempenho com base em experiências passadas.

Este modelo é implementado como uma interface comum que todos os agentes específicos implementam, garantindo consistência no comportamento dos agentes e facilitando a integração de novos agentes ao sistema.

4.5.2 Tipos de Agentes

O sistema implementa os seguintes tipos de agentes especializados:

- **Agente Coordenador:** Responsável pela coordenação geral do processo de design, orquestrando a comunicação entre agentes e garantindo a coerência global do design.
- **Agente Especialista em Mecânicas de Jogo:** Foca na geração e refinamento de mecânicas de jogo adequadas aos objetivos educacionais, considerando as especificidades de jogos endógenos.
- **Agente Especialista em Narrativa:** Desenvolve elementos narrativos que integram conteúdo educacional de forma endógena.
- **Agente Especialista em Engajamento:** Concentra-se em elementos que promovem o engajamento dos jogadores, considerando aspectos motivacionais específicos para contextos educacionais.
- **Agente Socrático:** Implementa o método socrático para questionar e refinar ideias, identificando inconsistências e oportunidades de melhoria.
- **Agente de Taxonomia de Bloom:** Especializado na classificação e formulação de objetivos educacionais segundo a Taxonomia de Bloom Revisada.

Cada tipo de agente é implementado como uma classe concreta que estende a interface de agente comum, adicionando funcionalidades específicas à sua especialidade.

4.5.3 Perfil de Agente

Os perfis de agente definem as características e responsabilidades de cada agente no sistema. Estes perfis são armazenados em um formato estruturado e incluem as seguintes informações:

- **Identificação:** Um identificador único para o agente no sistema.
- **Nome e Descrição:** Um nome legível e uma descrição da função do agente.
- **Especialidade:** A área de expertise do agente (ex: mecânicas de jogo, narrativa, etc.).
- **Conhecimento de Domínio:** Informações específicas do domínio que o agente utiliza para realizar suas tarefas.

- **Responsabilidades:** As tarefas específicas que o agente é responsável por executar.
- **Comportamento de Comunicação:** Regras que governam como o agente interage com outros agentes.
- **Parâmetros de Configuração:** Configurações que controlam o comportamento do agente.

Os perfis de agente são carregados durante a inicialização do sistema e podem ser ajustados durante a execução para otimizar o desempenho do sistema.

4.5.4 Comunicação entre Agentes

A comunicação entre agentes segue o modelo baseado em intenções e compromimentos, onde os agentes trocam mensagens estruturadas que expressam suas intenções e solicitações. O sistema implementa os seguintes tipos de mensagens:

- **Solicitações:** Pedidos de informação ou ação enviados de um agente para outro.
- **Respostas:** Informações fornecidas em resposta a solicitações.
- **Atualizações:** Informações enviadas proativamente para manter outros agentes atualizados sobre mudanças relevantes.
- **Coordenação:** Mensagens relacionadas à coordenação de atividades entre múltiplos agentes.

Cada mensagem inclui os seguintes campos:

- **Remetente:** O agente que envia a mensagem.
- **Destinatário:** O agente que deve receber a mensagem.
- **Tipo:** O tipo de mensagem (solicitação, resposta, atualização, coordenação).
- **Conteúdo:** O conteúdo da mensagem, estruturado de acordo com o tipo.
- **Metadados:** Informações adicionais sobre a mensagem, como prioridade e tempo de vida.

As mensagens são transmitidas através de um barramento de mensagens assíncrono, que permite comunicação eficiente entre agentes enquanto desacopla os emissores dos receptores.

4.5.5 Coordenação de Agentes

A coordenação entre agentes é gerenciada pelo Agente Coordenador, que implementa mecanismos para garantir a coerência global do processo de design. Os principais mecanismos de coordenação incluem:

- **Planejamento de Tarefas:** Distribuição de tarefas entre agentes especializados de acordo com suas capacidades.

- **Monitoramento de Progresso:** Acompanhamento do estado de cada tarefa atribuída aos agentes.
- **Resolução de Conflitos:** Mediação de conflitos entre sugestões ou decisões de diferentes agentes.
- **Priorização de Tarefas:** Definição de prioridades para tarefas pendentes de acordo com o contexto atual.

O Agente Coordenador mantém um modelo global do estado do processo de design, que é atualizado com base nas informações recebidas de outros agentes. Este modelo é utilizado para tomar decisões de coordenação e garantir a consistência do processo.

4.5.6 Mecanismo de Evolução de Agentes

O sistema implementa mecanismos que permitem a evolução dos agentes ao longo do tempo, melhorando seu desempenho com base em feedback e experiência. Estes mecanismos incluem:

- **Aprendizado por Feedback:** Os agentes recebem feedback sobre suas sugestões e utilizam essas informações para ajustar seus parâmetros internos.
- **Memória Episódica:** Agentes armazenam experiências passadas e as utilizam para informar decisões futuras.
- **Adaptação de Parâmetros:** Ajuste automático de parâmetros de configuração com base no desempenho observado.
- **Transferência de Conhecimento:** Compartilhamento de conhecimento aprendido entre agentes com funções similares.

Estes mecanismos são implementados através de algoritmos de aprendizado de máquina que analisam padrões de feedback e ajustam os modelos internos dos agentes de acordo.

4.6 Integração com Modelos de Linguagem

A integração com modelos de linguagem de grande escala (LLMs) é um aspecto central do sistema, permitindo que os agentes especializados utilizem a capacidade de processamento de linguagem natural destes modelos para gerar sugestões e análises.

4.6.1 Arquitetura de Integração

A integração com LLMs segue uma arquitetura cliente-servidor, onde o sistema atua como cliente e os serviços de LLM como servidores. Esta arquitetura inclui os seguintes componentes:

- **Adaptadores de API:** Implementam a comunicação com diferentes serviços de LLM, abstraindo as diferenças entre APIs.
- **Gerenciadores de Contexto:** Preparam e gerenciam o contexto para solicitações a LLMs, garantindo que as informações relevantes sejam incluídas.
- **Processadores de Resposta:** Analisam e processam as respostas recebidas dos LLMs, extraindo informações estruturadas.
- **Mecanismos de Fallback:** Implementam estratégias de recuperação para lidar com falhas de comunicação ou respostas inadequadas.

Esta arquitetura suporta múltiplos provedores de LLM, permitindo a seleção do modelo mais adequado para cada tipo de tarefa.

4.6.2 Prompting Estruturado

O sistema utiliza técnicas de prompting estruturado para garantir que as solicitações enviadas aos LLMs produzam respostas úteis e relevantes. Estas técnicas incluem:

- **Templates de Prompt:** Modelos predefinidos para diferentes tipos de solicitações, incluindo placeholders para informações específicas.
- **Instruções Contextuais:** Informações adicionais fornecidas no prompt para contextualizar a solicitação.
- **Exemplos Few-Shot:** Exemplos de entradas e saídas esperadas, permitindo que o LLM aprenda por analogia.
- **Restrições e Diretrizes:** Regras específicas para o formato e conteúdo da resposta esperada.

Os templates de prompt são armazenados em um repositório centralizado e podem ser atualizados independentemente do código do sistema, permitindo a melhoria contínua das solicitações sem necessidade de alterações no código.

4.6.3 Processamento de Respostas

As respostas recebidas dos LLMs são processadas para extrair informações estruturadas que possam ser utilizadas pelos agentes. Este processamento inclui:

- **Parsing de Texto:** Extração de informações estruturadas de respostas textuais.
- **Validação Semântica:** Verificação da relevância e adequação das respostas para o contexto atual.
- **Normalização de Formato:** Conversão de respostas para formatos padronizados utilizados internamente pelo sistema.
- **Ranking de Sugestões:** Ordenação de múltiplas sugestões por relevância ou qualidade.

O processamento de respostas é implementado através de uma pipeline que aplica múltiplos estágios de processamento, cada um responsável por um aspecto específico da transformação da resposta bruta em informações estruturadas.

4.6.4 Gerenciamento de Contexto

O gerenciamento de contexto é um aspecto crítico da integração com LLMs, garantindo que as solicitações incluam informações suficientes para produzir respostas relevantes. O sistema implementa os seguintes mecanismos de gerenciamento de contexto:

- **Seleção de Informações Relevantes:** Filtragem de informações disponíveis para incluir apenas as mais relevantes para a solicitação atual.
- **Compressão de Contexto:** Técnicas para reduzir o tamanho do contexto sem perder informações essenciais.
- **Gerenciamento de Histórico:** Manutenção de um histórico de interações que pode ser incluído no contexto quando relevante.
- **Atualização Dinâmica:** Ajuste do contexto com base em mudanças no estado do sistema ou feedback do usuário.

Estes mecanismos permitem que o sistema mantenha um equilíbrio entre a inclusão de informações suficientes para contexto e a eficiência das solicitações.

4.7 Módulos Funcionais

O sistema é organizado em quatro módulos funcionais principais, cada um responsável por uma etapa específica do processo de design de jogos educacionais endógenos. Esta seção detalha a implementação e funcionamento de cada módulo.

4.7.1 Módulo de Brainstorming

O Módulo de Brainstorming facilita sessões de brainstorming assistidas por IA para geração de ideias iniciais para jogos educacionais endógenos.

Estrutura de Classes A Figura 4.2 apresenta o diagrama de classes do Módulo de Brainstorming.

O módulo de Brainstorming é implementado através das seguintes classes principais:

- **BrainstormController:** Implementa os endpoints da API REST para o módulo de Brainstorming, processando requisições do cliente e coordenando o fluxo de trabalho.

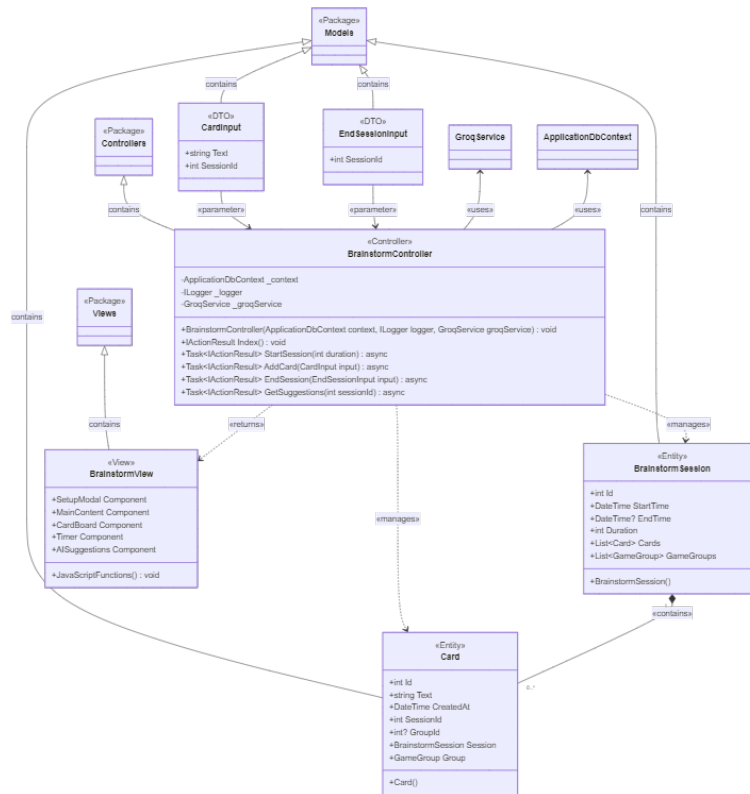


Figura 4.2: Diagrama de classes do módulo de Brainstorming.

- **BrainstormService:** Implementa a lógica de negócios para sessões de brainstorming, incluindo criação, gestão e finalização de sessões.
- **BrainstormSession:** Representa uma sessão de brainstorming ativa, incluindo seu estado atual e configurações.
- **Card:** Representa uma ideia individual gerada durante uma sessão de brainstorming.
- **CardGroup:** Representa um agrupamento de cards relacionados.
- **GroqService:** Implementa a integração com o modelo de linguagem para geração de sugestões.

Estas classes seguem o padrão de design de domínio, onde cada classe representa um conceito específico do domínio do problema. A comunicação entre classes é implementada através de interfaces bem definidas, facilitando a substituição de componentes e os testes unitários.

Modelo de Dados O modelo de dados para o Módulo de Brainstorming é baseado no Entity-Relationship Diagram mostrado na Figura 4.3.

As principais entidades no modelo de dados incluem:

- **BRAINSTORM_SESSION:** Armazena informações sobre sessões de brainstorming, incluindo hora de início, hora de término e duração.
- **CARD:** Armazena ideias individuais geradas durante sessões de brainstorming.

4. O sistema cria uma nova sessão com um identificador único e inicia o temporizador.
5. O usuário adiciona ideias (cards) à sessão através da interface.
6. O sistema armazena cada card no banco de dados e o associa à sessão atual.
7. O usuário pode solicitar sugestões da IA a qualquer momento durante a sessão.
8. O sistema envia os cards existentes para o GroqService, que utiliza um LLM para gerar sugestões relacionadas.
9. O sistema apresenta as sugestões ao usuário, que pode adicioná-las à sessão como novos cards.
10. Ao término da sessão, o sistema marca a sessão como concluída e redireciona o usuário para a próxima etapa.

Este fluxo de trabalho é implementado como uma série de endpoints REST no BrainstormController, cada um responsável por uma parte específica do processo.

Integração com IA A integração com IA no Módulo de Brainstorming é implementada através do GroqService, que se comunica com o modelo de linguagem para gerar sugestões. O processo de geração de sugestões inclui:

1. O sistema prepara um prompt para o LLM, incluindo todos os cards existentes na sessão e instruções específicas sobre o tipo de sugestões desejado.
2. O prompt é enviado para o LLM através da API do provedor selecionado.
3. O LLM processa o prompt e gera sugestões para novos cards.
4. O sistema recebe a resposta do LLM e a processa para extrair sugestões estruturadas.
5. As sugestões são apresentadas ao usuário, que pode selecionar quais deseja adicionar à sessão.

O GroqService implementa mecanismos de retry e fallback para lidar com falhas na comunicação com o LLM, garantindo a robustez do sistema mesmo em condições de rede instáveis.

4.7.2 Módulo de Agente Socrático

O Módulo de Agente Socrático implementa uma abordagem baseada no método socrático para refinar ideias e identificar problemas potenciais.

Estrutura de Classes A Figura 4.5 apresenta o diagrama de classes do Módulo de Agente Socrático.

O módulo de Agente Socrático é implementado através das seguintes classes principais:

- **SocraticController:** Implementa os endpoints da API REST para o módulo de Agente Socrático.

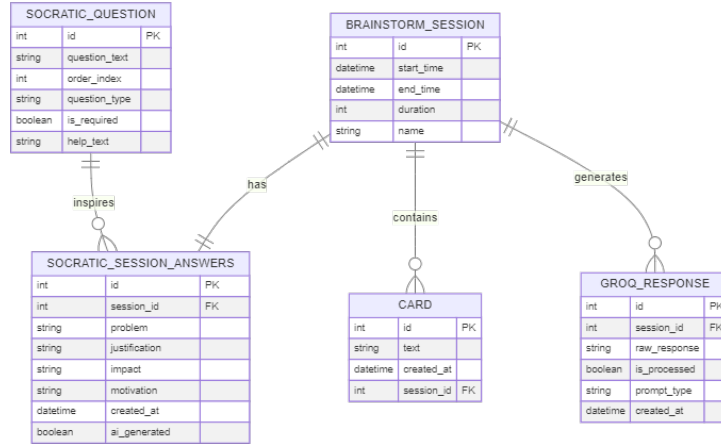


Figura 4.6: Diagrama entidade-relacionamento do módulo de Agente Socrático.

cial entre entidades, garantindo que respostas socráticas estejam sempre associadas a uma sessão de brainstorming válida.

Fluxo de Trabalho O fluxo de trabalho para o Módulo de Agente Socrático é mostrado no diagrama de sequência na Figura 4.7.

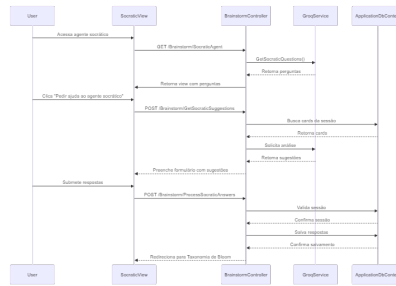


Figura 4.7: Diagrama de sequência do módulo de Agente Socrático.

O fluxo de trabalho típico para uma interação com o Agente Socrático inclui as seguintes etapas:

1. O usuário acessa a interface do Agente Socrático após a conclusão da sessão de brainstorming.
2. O sistema carrega as questões socráticas predefinidas relevantes para o contexto atual.
3. O usuário solicita assistência do Agente Socrático para analisar as ideias geradas.
4. O sistema recupera os cards da sessão de brainstorming e os envia para o GroqService.
5. O GroqService utiliza um LLM para analisar os cards e gerar sugestões para respostas às questões socráticas.
6. O sistema apresenta as sugestões ao usuário em um formulário estruturado.

7. O usuário revisa, modifica se necessário, e submete as respostas às questões socráticas.
8. O sistema valida e armazena as respostas no banco de dados.
9. O sistema redireciona o usuário para a próxima etapa do processo.

Este fluxo de trabalho é projetado para promover a reflexão crítica sobre as ideias geradas durante a sessão de brainstorming, identificando problemas potenciais e oportunidades de melhoria.

Implementação do Método Socrático O método socrático é implementado através de um conjunto de questões predefinidas que abordam diferentes aspectos do design de jogos educacionais. Estas questões são organizadas em categorias como:

- **Identificação de Problemas:** Questões que focam na identificação de problemas ou desafios no design proposto.
- **Justificação:** Questões que solicitam justificativas para decisões de design ou escolhas feitas.
- **Impacto:** Questões que exploram o impacto potencial do design nos alunos ou no processo educacional.
- **Motivação:** Questões que examinam os aspectos motivacionais do design proposto.

O Agente Socrático utiliza um LLM para analisar as ideias existentes e gerar sugestões para respostas a estas questões, mas o usuário mantém o controle final sobre as respostas submetidas.

4.7.3 Módulo de Endo-GDC

O Módulo de Endo-GDC oferece uma interface estruturada para o preenchimento do Endo-GDC com assistência de agentes especializados, projetado para o design de jogos educativos endógenos.

Estrutura de Classes A Figura 4.8 apresenta o diagrama de classes do Módulo de Endo-GDC.

O módulo de Endo-GDC é implementado através das seguintes classes principais:

- **GameDesignController:** Implementa os endpoints da API REST para o módulo de Endo-GDC.
- **GameDesignService:** Implementa a lógica de negócios para o preenchimento do Endo-GDC.
- **GdcTemplate:** Representa a estrutura do Endo-GDC, incluindo seções predefinidas.
- **GdcSection:** Representa uma seção específica do Endo-GDC.
- **GdcNote:** Representa uma anotação individual no Endo-GDC.

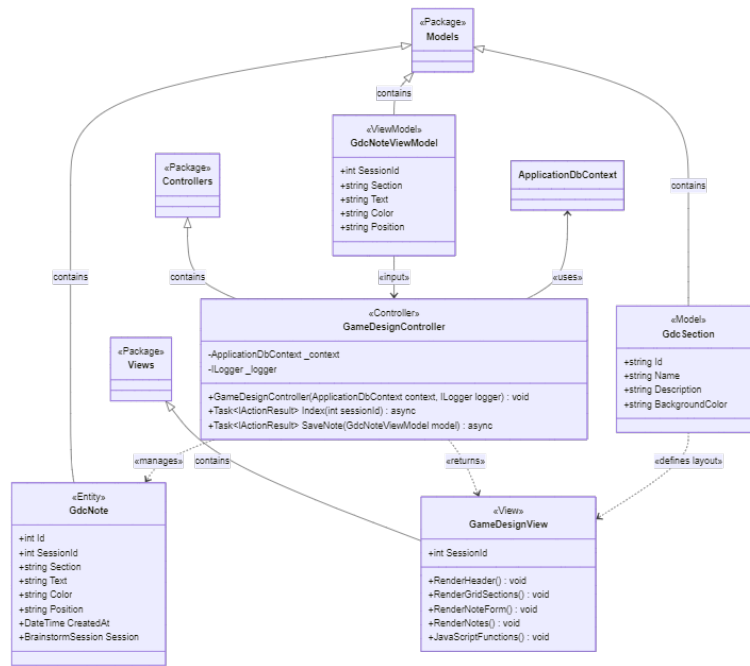


Figura 4.8: Diagrama de classes do módulo de Endo-GDC.

- **AI Agent:** Representa um agente especializado que fornece sugestões para o Endo-GDC.
- **GameSuggestionService:** Coordena a geração de sugestões por múltiplos agentes especializados.
- **MultiagentService:** Gerencia a comunicação entre agentes especializados.

Estas classes implementam uma arquitetura baseada em agentes, onde cada agente especializado é responsável por gerar sugestões para uma área específica do design de jogos educacionais.

Modelo de Dados O modelo de dados para o Módulo de Endo-GDC é baseado no Entity-Relationship Diagram mostrado na Figura 4.9.

As principais entidades no modelo de dados incluem:

- **GDC_TEMPLATE:** Define a estrutura do Endo-GDC, incluindo seções padronizadas.
- **GDC_SECTION:** Representa uma seção específica do Endo-GDC, com atributos como nome, descrição e cor de fundo.
- **GDC_NOTE:** Representa uma anotação no Endo-GDC, associada a uma seção específica e contendo atributos como texto, cor e posição.
- **AI_SUGGESTION_REQUEST:** Captura solicitações de sugestões feitas aos agentes de IA.
- **AI_AGENT_RESPONSE:** Armazena as respostas dos agentes de IA às solicitações.

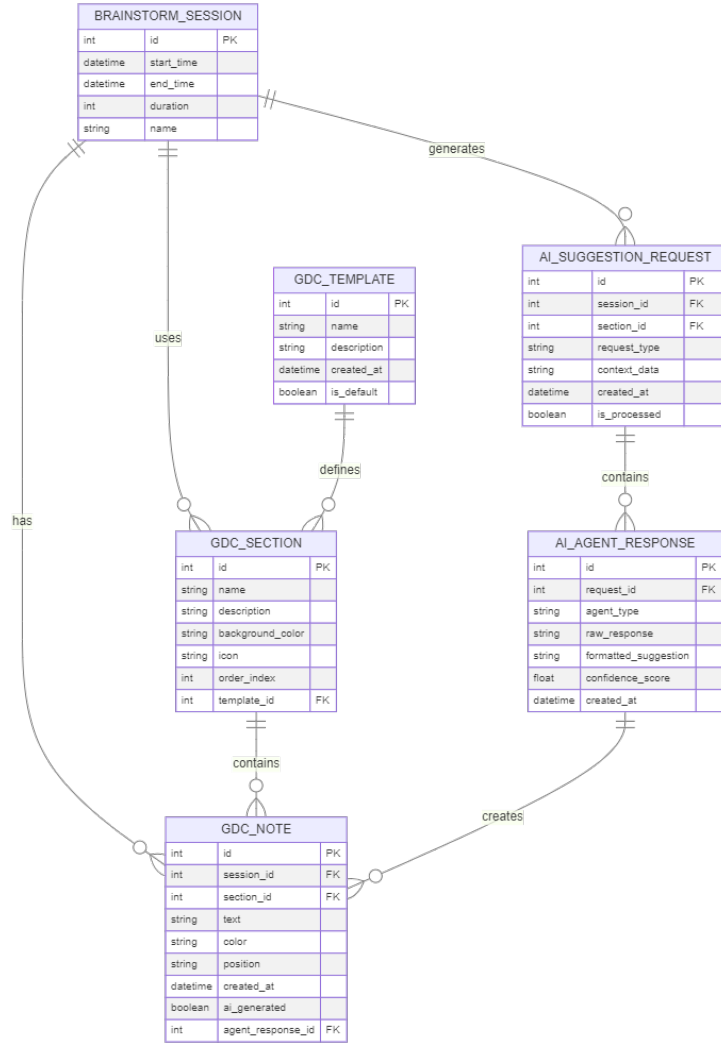


Figura 4.9: Diagrama entidade-relacionamento do módulo de Endo-GDC.

- **BRAINSTORM_SESSION**: Relaciona as notas e sugestões à sessão de brainstorming correspondente.

Este modelo de dados permite o armazenamento e recuperação eficiente de notas e sugestões para o Endo-GDC, mantendo a estrutura hierárquica do canvas e as relações entre diferentes elementos.

Fluxo de Trabalho O fluxo de trabalho para o Módulo de Endo-GDC é mostrado no diagrama de sequência na Figura 4.10.

O fluxo de trabalho típico para o preenchimento do Endo-GDC inclui as seguintes etapas:

1. O usuário acessa a interface do Endo-GDC após a conclusão da interação com o Agente Sócrático.
2. O sistema carrega o template do Endo-GDC, incluindo todas as seções predefinidas.

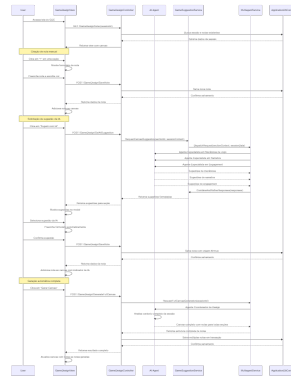


Figura 4.10: Diagrama de sequência do módulo de Endo-GDC.

3. O sistema recupera e exibe notas existentes, se houver.
4. O usuário pode adicionar notas manualmente a qualquer seção do canvas.
5. O usuário pode solicitar sugestões da IA para uma seção específica do canvas.
6. O sistema envia a solicitação para o GameSuggestionService, que coordena a geração de sugestões por múltiplos agentes especializados.
7. Os agentes especializados geram sugestões baseadas em suas áreas de expertise.
8. O sistema apresenta as sugestões ao usuário, que pode selecionar qual deseja adicionar ao canvas.
9. O sistema adiciona a sugestão selecionada como uma nota no canvas, marcando-a como gerada por IA.
10. O usuário pode solicitar a geração automática de um canvas completo.
11. O sistema adiciona as notas geradas ao canvas e as exibe ao usuário.

Este fluxo de trabalho é projetado para oferecer um equilíbrio entre controle manual e assistência automatizada, permitindo que o usuário aproveite a especialização dos agentes de IA enquanto mantém a decisão final sobre o conteúdo do canvas.

Implementação do Sistema Multiagente A implementação do sistema multiagente para o Módulo de Endo-GDC inclui os seguintes agentes especializados:

- **Agente Especialista em Mecânicas de Jogo:** Gera sugestões para mecânicas de jogo que suportam os objetivos educacionais identificados.
- **Agente Especialista em Narrativa:** Gera sugestões para elementos narrativos que integram o conteúdo educacional de forma endógena.
- **Agente Especialista em Engajamento:** Gera sugestões para elementos que promovem o engajamento dos jogadores.
- **Agente Coordenador:** Coordena os outros agentes.

Cada agente é implementado como uma instância da classe `AI Agent`, configurada com um perfil específico e integrada com um LLM para geração de sugestões. O `MultiagentService` coordena a comunicação entre agentes, garantindo que as sugestões sejam complementares e coerentes.

4.7.4 Módulo de Taxonomia de Bloom

O Módulo de Taxonomia de Bloom auxilia na classificação e formulação de objetivos educacionais baseados na Taxonomia de Bloom Revisada.

Estrutura de Classes A Figura 4.11 apresenta o diagrama de classes do Módulo de Taxonomia de Bloom.

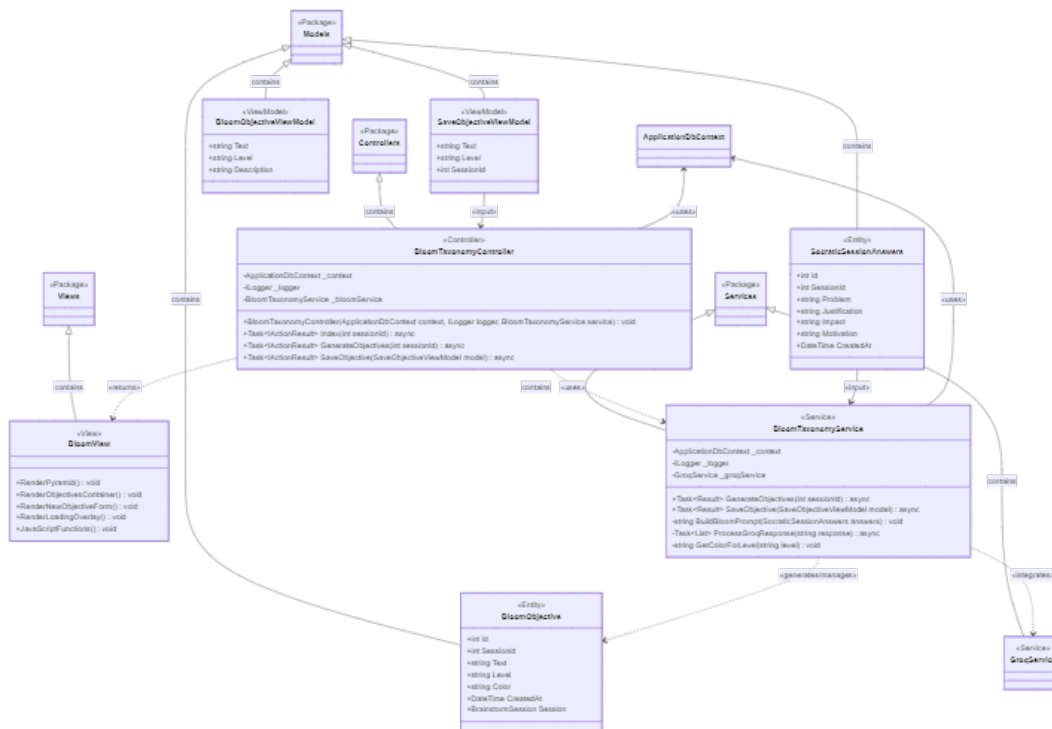


Figura 4.11: Diagrama de classes do módulo de Taxonomia de Bloom.

O módulo de Taxonomia de Bloom é implementado através das seguintes classes principais:

- **BloomTaxonomyController:** Implementa os endpoints da API REST para o módulo de Taxonomia de Bloom.
- **BloomTaxonomyService:** Implementa a lógica de negócios para classificação e formulação de objetivos educacionais.
- **BloomTaxonomyLevel:** Representa um nível da Taxonomia de Bloom Revisada.
- **BloomObjective:** Representa um objetivo educacional classificado segundo a Taxonomia de Bloom.
- **GroqService:** Implementa a integração com o modelo de linguagem para geração de objetivos educacionais.

Estas classes implementam uma arquitetura que facilita a classificação e formulação de objetivos educacionais alinhados com a Taxonomia de Bloom Revisada.

Modelo de Dados O modelo de dados para o Módulo de Taxonomia de Bloom é baseado no Entity-Relationship Diagram mostrado na Figura 4.12.

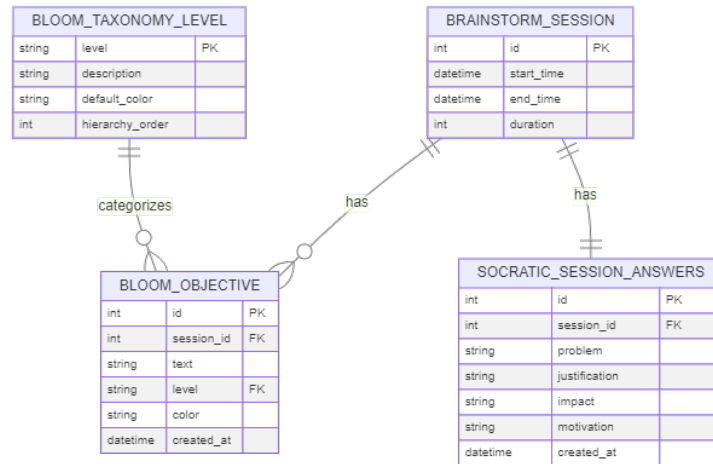


Figura 4.12: Diagrama entidade-relacionamento do módulo de Taxonomia de Bloom.

As principais entidades no modelo de dados incluem:

- **BLOOM_TAXONOMY_LEVEL**: Representa os níveis da Taxonomia de Bloom, com atributos como descrição e ordem hierárquica.
- **BLOOM_OBJECTIVE**: Representa um objetivo educacional baseado na Taxonomia de Bloom, contendo atributos como texto, nível e cor.
- **BRAINSTORM_SESSION**: Relaciona os objetivos educacionais à sessão de brainstorming correspondente.
- **SOCRATIC_SESSION_ANSWERS**: Fornece contexto para a geração de objetivos educacionais baseados nas respostas socráticas.

Este modelo de dados permite o armazenamento e recuperação eficiente de objetivos educacionais classificados segundo a Taxonomia de Bloom, mantendo a relação com o contexto da sessão de brainstorming e as respostas socráticas.

Fluxo de Trabalho O fluxo de trabalho para o Módulo de Taxonomia de Bloom é mostrado no diagrama de sequência na Figura 4.13.

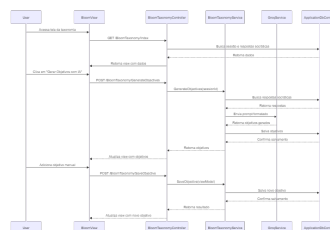


Figura 4.13: Diagrama de sequência do módulo de Taxonomia de Bloom.

O fluxo de trabalho típico para a classificação e formulação de objetivos educacionais inclui as seguintes etapas:

1. O usuário acessa a interface de Taxonomia de Bloom após a conclusão do preenchimento do Endo-GDC.
2. O sistema carrega os níveis da Taxonomia de Bloom e objetivos existentes, se houver.
3. O usuário solicita a geração automática de objetivos educacionais.
4. O sistema recupera as respostas socráticas e as notas do Endo-GDC como contexto.
5. O sistema envia o contexto para o BloomTaxonomyService, que utiliza o GroqService para gerar objetivos educacionais.
6. O GroqService utiliza um LLM para gerar objetivos educacionais classificados segundo a Taxonomia de Bloom.
7. O sistema apresenta os objetivos gerados ao usuário, organizados por nível da taxonomia.
8. O usuário pode editar os objetivos gerados ou adicionar novos objetivos manualmente.
9. O sistema valida e armazena os objetivos no banco de dados.
10. O sistema finaliza o processo de design, apresentando um resumo dos artefatos gerados.

Este fluxo de trabalho é projetado para facilitar a formulação de objetivos educacionais alinhados com a Taxonomia de Bloom Revisada, garantindo que o jogo educacional tenha uma base pedagógica sólida.

Implementação da Taxonomia de Bloom A implementação da Taxonomia de Bloom Revisada no sistema segue a estrutura proposta por Anderson et al. (2001), que inclui seis níveis cognitivos:

- **Lembrar:** Recuperar conhecimento relevante da memória de longo prazo.
- **Entender:** Construir significado a partir de mensagens instrucionais, incluindo comunicação oral, escrita e gráfica.
- **Aplicar:** Executar ou usar um procedimento em uma situação específica.
- **Analisar:** Quebrar material em partes constituintes e determinar como as partes se relacionam entre si e com uma estrutura ou propósito geral.
- **Avaliar:** Fazer julgamentos baseados em critérios e padrões.
- **Criar:** Juntar elementos para formar um todo coerente ou funcional; reorganizar elementos em um novo padrão ou estrutura.

Cada nível é implementado como uma instância da classe BloomTaxonomyLevel, com atributos que descrevem suas características e verbos associados. O sistema utiliza esta estrutura para classificar e formular objetivos educacionais que abrangem diferentes níveis cognitivos, garantindo uma abordagem educacional abrangente e equilibrada.

4.8 Implementação Técnica

Esta seção detalha os aspectos técnicos da implementação do sistema, incluindo tecnologias utilizadas, padrões de design e estratégias de integração.

4.8.1 Tecnologias Utilizadas

O sistema foi implementado utilizando as seguintes tecnologias:

- **Backend:** ASP.NET Core 7.0 para o desenvolvimento da aplicação web, incluindo API REST e lógica de negócios.
- **Frontend:** React 18 para a interface do usuário, com TypeScript para tipagem estática e Redux para gerenciamento de estado.
- **UI/UX:** Material-UI e Tailwind CSS para componentes visuais e estilização responsiva.
- **Banco de Dados:** Entity Framework Core 7.0 com SQL Server para persistência de dados.
- **Integração com LLMs:** Cliente HTTP para comunicação com APIs de LLM, incluindo OpenAI API e Groq API.
- **CI/CD:** GitHub Actions para integração contínua e implantação contínua.
- **Monitoramento:** Application Insights para monitoramento de desempenho e telemetria.

Estas tecnologias foram selecionadas com base em sua maturidade, desempenho e suporte, fornecendo uma base sólida para a implementação do sistema.

4.8.2 Arquitetura de Software

A arquitetura de software do sistema segue os princípios de Clean Architecture, com as seguintes camadas:

- **Camada de Apresentação:** Implementada como aplicação React, responsável pela interface do usuário.
- **Camada de API:** Implementada como API REST utilizando ASP.NET Core, responsável pela comunicação entre frontend e backend.
- **Camada de Aplicação:** Contém casos de uso e orquestração de regras de negócios, implementada utilizando o padrão CQRS (Command Query Responsibility Segregation).
- **Camada de Domínio:** Implementa a lógica de negócios central e entidades do domínio, seguindo os princípios de Domain-Driven Design.
- **Camada de Infraestrutura:** Responsável pela persistência de dados, integração com serviços externos e aspectos técnicos.

Esta arquitetura em camadas facilita a manutenção, testabilidade e evolução do sistema, permitindo que cada camada seja modificada independentemente das outras.

4.8.3 Padrões de Design

O sistema implementa os seguintes padrões de design:

- **Dependency Injection:** Utilizado para desacoplar componentes e facilitar testes unitários.
- **Repository:** Implementado para abstrair o acesso a dados e facilitar a substituição da tecnologia de persistência.
- **Unit of Work:** Utilizado para gerenciar transações e garantir a consistência dos dados.
- **Factory:** Implementado para criar instâncias de agentes especializados com configurações específicas.
- **Strategy:** Utilizado para selecionar dinamicamente o algoritmo a ser utilizado para diferentes tarefas.
- **Observer:** Implementado para notificar componentes sobre mudanças no estado do sistema.
- **Mediator:** Utilizado para desacoplar componentes que precisam comunicar-se entre si.

Estes padrões de design promovem a manutenibilidade, extensibilidade e testabilidade do código, seguindo princípios de design orientado a objetos.

4.8.4 Gestão de Configuração

O sistema implementa uma abordagem hierárquica para gestão de configuração, com os seguintes níveis:

- **Configurações de Sistema:** Definidas no arquivo appsettings.json e variáveis de ambiente, incluindo conexões com bancos de dados e serviços externos.
- **Configurações de Módulo:** Definidas por módulo funcional, incluindo parâmetros específicos para cada módulo.
- **Configurações de Agente:** Definidas por tipo de agente, incluindo perfis, comportamentos e parâmetros de integração com LLMs.
- **Configurações de Sessão:** Definidas por sessão de usuário, incluindo preferências e estado temporário.

Esta abordagem permite uma configuração flexível e granular do sistema, facilitando ajustes para diferentes ambientes (desenvolvimento, teste, produção) e casos de uso.

4.8.5 Segurança e Privacidade

O sistema implementa as seguintes medidas de segurança e privacidade:

- **Autenticação:** Implementada utilizando JWT (JSON Web Tokens) para autenticação de usuários.
- **Autorização:** Baseada em papéis (RBAC - Role-Based Access Control) para controle de acesso a recursos.
- **Proteção de Dados:** Implementada utilizando criptografia para dados sensíveis armazenados no banco de dados.
- **Sanitização de Entrada:** Realizada para prevenir ataques de injeção e XSS (Cross-Site Scripting).
- **Proteção CSRF:** Implementada para prevenir ataques de Cross-Site Request Forgery.
- **Limitação de Taxa:** Aplicada para prevenir ataques de força bruta e DoS (Denial of Service).

Estas medidas garantem a segurança e privacidade dos dados do usuário, seguindo as melhores práticas e regulamentações aplicáveis.

4.8.6 Testes e Qualidade

O sistema implementa uma estratégia abrangente de testes, incluindo:

- **Testes Unitários:** Implementados utilizando xUnit para .NET e Jest para JavaScript, verificando a funcionalidade de componentes individuais.
- **Testes de Integração:** Verificando a interação entre componentes, incluindo comunicação com banco de dados e serviços externos.
- **Testes de Sistema:** Verificando o comportamento do sistema como um todo, incluindo fluxos de trabalho completos.
- **Testes de UI:** Implementados utilizando Cypress, verificando a interação do usuário com a interface.

Estes testes são executados automaticamente como parte do processo de CI/CD, garantindo a qualidade do código e a estabilidade do sistema.

4.9 Desafios e Soluções

A implementação de um sistema multiagente baseado em LLMs para design de jogos educacionais apresenta desafios significativos. Esta seção discute os principais desafios encontrados durante o desenvolvimento e as soluções implementadas.

4.10 Gerenciamento de Recursos LLM

Desafio: O acesso a APIs de LLM está sujeito a limites de taxa e custos por token, tornando o gerenciamento eficiente de recursos crítico para a viabilidade do sistema.

Solução: O sistema implementa as seguintes estratégias para otimizar o uso de recursos LLM:

- **Caching de Respostas:** Respostas para prompts similares são armazenadas em cache para reduzir chamadas redundantes.
- **Compressão de Contexto:** Técnicas de sumarização e seleção são utilizadas para reduzir o tamanho do contexto enviado ao LLM.
- **Priorização de Requisições:** Requisições são priorizadas com base em sua criticidade e impacto no fluxo de trabalho.
- **Agendamento de Lotes:** Requisições não urgentes são agrupadas em lotes para processamento eficiente.

Estas estratégias reduzem significativamente o consumo de tokens e o número de chamadas à API, mantendo a qualidade das respostas e a responsividade do sistema.

4.11 Qualidade e Relevância das Sugestões

Desafio: Garantir que as sugestões geradas pelos agentes baseados em LLM sejam relevantes, úteis e alinhadas com o contexto específico do jogo educacional sendo projetado.

Solução: O sistema implementa as seguintes abordagens para melhorar a qualidade e relevância das sugestões:

- **Prompting Estruturado:** Utilização de templates de prompt projetados para cada tipo de sugestão.
- **Contextualização Rica:** Inclusão de informações relevantes do domínio e do projeto específico no contexto enviado ao LLM.
- **Filtragem e Ranking:** Processamento de respostas para selecionar e priorizar as sugestões mais relevantes.
- **Aprendizado de Feedback:** Ajuste de parâmetros com base no feedback do usuário sobre sugestões anteriores.

Estas abordagens resultam em sugestões mais úteis e contextualmente relevantes, melhorando a eficácia da assistência fornecida pelo sistema.

4.12 Coordenação entre Agentes

Desafio: Coordenar múltiplos agentes especializados para produzir sugestões coerentes e complementares sem contradições ou duplicações.

Solução: O sistema implementa os seguintes mecanismos de coordenação:

- **Agente Coordenador:** Um agente especializado que supervisiona e coordena as atividades dos outros agentes.
- **Modelo Compartilhado de Conhecimento:** Uma representação compartilhada do estado atual do design, acessível a todos os agentes.
- **Protocolos de Comunicação:** Regras e formatos para troca de informações entre agentes.
- **Resolução de Conflitos:** Algoritmos para detectar e resolver contradições entre sugestões de diferentes agentes.

Estes mecanismos garantem que os agentes trabalhem de forma coordenada, produzindo sugestões que se complementam e formam um design coerente.

4.13 Experiência do Usuário

Desafio: Criar uma interface de usuário intuitiva e eficiente que aproveite ao máximo a assistência dos agentes sem sobrecarregar o usuário com opções ou informações excessivas.

Solução: O sistema implementa as seguintes estratégias de UI/UX:

- **Design Progressivo:** Apresentação gradual de opções e informações, seguindo o fluxo natural do processo de design.
- **Sugestões Contextuais:** Apresentação de sugestões relevantes para o contexto atual do usuário, reduzindo a sobrecarga cognitiva.
- **Indicadores de Fonte:** Diferenciação visual clara entre conteúdo gerado pelo usuário e sugestões de IA.
- **Controles de Refinamento:** Ferramentas para ajustar e personalizar sugestões geradas por IA.

Estas estratégias resultam em uma experiência de usuário que equilibra assistência e controle, permitindo que usuários com diferentes níveis de experiência aproveitem o sistema de forma eficaz.

4.14 Evolução e Manutenção

O sistema foi projetado considerando a evolução contínua e facilidade de manutenção. Esta seção descreve as estratégias implementadas para suportar o desenvolvimento e manutenção do sistema ao longo do tempo.

4.14.1 Extensibilidade

O sistema implementa várias estratégias para facilitar a extensão com novas funcionalidades:

- **Arquitetura Modular:** Componentes são organizados em módulos com interfaces bem definidas, permitindo a adição de novos módulos sem modificar os existentes.
- **Plugins:** O sistema suporta plugins para estender funcionalidades específicas, como integração com novos provedores de LLM ou formatos de exportação.
- **Configuração Extensível:** O sistema de configuração permite a adição de novos parâmetros sem modificar a estrutura existente.
- **Event Sourcing:** O sistema registra eventos que representam mudanças de estado, facilitando a adição de novos comportamentos em resposta a eventos existentes.

Estas estratégias permitem que o sistema evolua incrementalmente sem necessidade de refatorações significativas.

4.14.2 Monitoramento e Diagnóstico

O sistema implementa mecanismos abrangentes para monitoramento e diagnóstico:

- **Logging Estruturado:** Registros detalhados de operações do sistema em formato estruturado, facilitando análise e filtragem.
- **Telemetria de Performance:** Medições automatizadas de tempo de resposta, uso de recursos e taxas de erro.
- **Rastreamento Distribuído:** Registro de transações que atravessam múltiplos componentes, facilitando a identificação de gargalos.
- **Alertas Proativos:** Notificações automáticas para condições anômalas que podem indicar problemas.

Estes mecanismos facilitam a identificação e resolução rápida de problemas, melhorando a confiabilidade e disponibilidade do sistema.

4.14.3 Versionamento e Compatibilidade

O sistema implementa estratégias para gerenciar versões e garantir compatibilidade:

- **Versionamento Semântico:** Todas as APIs e interfaces seguem o padrão de versionamento semântico (MAJOR.MINOR.PATCH).
- **Compatibilidade Backward:** Novas versões mantêm compatibilidade com dados e configurações de versões anteriores.
- **Migrações Automáticas:** O sistema pode atualizar automaticamente dados e configurações para novos formatos quando necessário.
- **Feature Flags:** Novas funcionalidades podem ser habilitadas seletivamente para facilitar testes e implantação gradual.

Estas estratégias facilitam a atualização do sistema sem interrupção do serviço ou perda de dados existentes.

4.14.4 Documentação

O sistema mantém documentação para facilitar o desenvolvimento e manutenção:

- **Documentação de API:** Gerada automaticamente a partir do código usando Swagger/OpenAPI.
- **Documentação de Arquitetura:** Diagramas e descrições da arquitetura do sistema, seguindo o padrão UML 2.5 .
- **Histórico de Decisões:** Registro de decisões arquiteturais significativas e suas justificativas.

4.15 Infraestrutura e Implantação

A infraestrutura do sistema foi projetada para garantir escalabilidade, disponibilidade e segurança. Esta seção descreve os componentes de infraestrutura e o processo de implantação.

4.15.1 Arquitetura de Infraestrutura

O sistema é implantado em uma arquitetura de nuvem que inclui os seguintes componentes:

- **Servidores Web:** Servidores implantados em múltiplas zonas de disponibilidade para alta disponibilidade.
- **Banco de Dados:** Banco de dados relacional com replicação para garantir durabilidade e alta disponibilidade.
- **Cache Distribuído:** Sistema de cache distribuído para melhorar o desempenho e reduzir a carga no banco de dados.
- **Sistema de Filas:** Fila de mensagens para processamento assíncrono de tarefas de longa duração.

- **CDN:** Rede de distribuição de conteúdo para entrega eficiente de ativos estáticos.
- **Serviços de Monitoramento:** Monitoramento centralizado para métricas, logs e alertas.

Esta arquitetura permite que o sistema escale horizontalmente para acomodar aumento de carga e mantenha alta disponibilidade mesmo em caso de falha de componentes individuais.

4.16 Comunicação com o Usuário

4.16.1 Interface de Usuário Web

A interface principal do sistema é uma aplicação web que implementa os seguintes princípios de design:

- **Design Responsivo:** Adaptação automática a diferentes tamanhos de tela, permitindo uso em dispositivos móveis e desktop.
- **Consistência Visual:** Aplicação consistente de elementos visuais e padrões de interação em todo o sistema.
- **Feedback Imediato:** Resposta visual imediata para ações do usuário, minimizando a sensação de latência.

A interface web implementa os quatro módulos funcionais descritos anteriormente (Brainstorming, Agente Socrático, Endo-GDC e Taxonomia de Bloom) como páginas separadas, conectadas por um fluxo de trabalho linear.

4.16.2 Notificações

O sistema implementa um mecanismo de notificações para manter o usuário informado sobre eventos relevantes:

- **Notificações In-App:** Mensagens exibidas dentro da aplicação para informar sobre conclusão de tarefas ou eventos do sistema.

O usuário pode configurar quais tipos de eventos geram notificações e por quais canais, permitindo uma experiência personalizada.

4.16.3 Exportação de Resultados

O sistema permite a exportação dos artefatos de design em múltiplos formatos:

- **PDF:** Documento formatado para impressão e compartilhamento, incluindo todos os artefatos de design.
- **HTML:** Versão web interativa dos artefatos, que pode ser compartilhada via URL.

- **JSON:** Formato estruturado para integração com outras ferramentas ou sistemas.
- **XLSX:** Planilha Excel para análise e manipulação adicional dos dados.

4.16.4 Otimização de Performance

O sistema foi submetido a análises e otimizações de performance, incluindo:

- **Profiling:** Identificação de gargalos de performance utilizando ferramentas de profiling.
- **Otimização de Consultas:** Refinamento de consultas ao banco de dados para melhorar tempos de resposta.
- **Caching:** Implementação de estratégias de cache em múltiplos níveis para reduzir a carga no banco de dados e serviços externos.
- **Otimização de Frontend:** Redução do tamanho de pacotes JavaScript e otimização de renderização para melhorar a responsividade da interface.

Estas otimizações resultaram em melhorias significativas no tempo de resposta e na utilização de recursos, permitindo que o sistema mantenha boa performance mesmo com aumento de carga.

4.16.5 Configurações de Agentes

Os agentes no sistema podem ser configurados para atender a necessidades específicas:

- **Perfis de Agente:** Configurações detalhadas que definem o comportamento e especialidade de cada agente.
- **Parâmetros de LLM:** Controles para ajustar a temperatura, diversidade e outras características das sugestões geradas.
- **Filtros de Conteúdo:** Configurações para garantir que as sugestões sejam apropriadas para diferentes faixas etárias ou contextos culturais.
- **Balanco de Intervenção:** Controles para ajustar o nível de assistência fornecida pelos agentes, de mínimo a máximo.

Estas configurações permitem que o sistema seja ajustado para diferentes estilos de trabalho e requisitos específicos.

4.16.6 API Extensível

O sistema expõe uma API REST que permite sua integração com outros sistemas e o desenvolvimento de extensões:

- **API Pública:** Endpoints documentados que permitem acesso a funcionalidades do sistema por aplicações externas.

- **Webhooks:** Mecanismos para notificar sistemas externos sobre eventos específicos.
- **Formato de Plugin:** Especificação para desenvolvimento de plugins que estendem funcionalidades específicas.
- **SDK:** Bibliotecas para facilitar a integração com a API em diferentes linguagens de programação.

Esta API permite que o sistema seja integrado em ecossistemas educacionais mais amplos e adaptado para necessidades específicas.

4.17 Considerações Éticas

O desenvolvimento e implantação do sistema considerou cuidadosamente aspectos éticos relacionados ao uso de IA em educação. Esta seção descreve as considerações éticas incorporadas no design e implementação do sistema.

4.17.1 Transparência

O sistema implementa mecanismos para garantir transparência no uso de IA:

- **Indicação de Origem:** Todas as sugestões geradas por IA são claramente identificadas como tal na interface.
- **Explicabilidade:** O sistema pode fornecer explicações sobre por que determinadas sugestões foram feitas.
- **Visibilidade de Parâmetros:** Configurações que afetam o comportamento dos agentes são visíveis e ajustáveis pelo usuário.
- **Logs de Atividade:** O sistema mantém registros detalhados das interações e decisões dos agentes.

Esta transparência permite que os usuários mantenham controle sobre o processo de design e compreendam o papel da IA como ferramenta de assistência.

4.17.2 Equidade e Inclusão

O sistema foi projetado para promover equidade e inclusão:

- **Diversidade de Conteúdo:** Mecanismos para garantir que sugestões sejam diversas e representativas.
- **Deteção de Viés:** Análise automática de sugestões para identificar e mitigar vieses potenciais.
- **Acessibilidade:** Interface em conformidade com padrões de acessibilidade para garantir usabilidade por pessoas com diferentes capacidades.
- **Suporte a Múltiplos Idiomas:** Internacionalização da interface e capacidade de gerar sugestões em diferentes idiomas.

Estas características ajudam a garantir que o sistema seja utilizável e benéfico para uma ampla gama de usuários e contextos educacionais.

4.17.3 Privacidade e Segurança

O sistema implementa medidas robustas para proteger a privacidade e segurança dos usuários:

- **Minimização de Dados:** Coleta apenas dados necessários para as funcionalidades do sistema.
- **Criptografia:** Proteção de dados sensíveis em trânsito e em repouso.
- **Controle de Acesso:** Mecanismos granulares para controlar quem pode acessar diferentes tipos de dados.
- **Retenção de Dados:** Políticas claras sobre quanto tempo dados são mantidos e como podem ser excluídos.

Estas medidas estão em conformidade com regulamentações relevantes de proteção de dados e boas práticas de segurança da informação.

4.17.4 Responsabilidade Humana

O sistema mantém o usuário humano como responsável final pelo design:

- **Aprovação Explícita:** Sugestões de IA requerem aprovação explícita do usuário antes de serem incorporadas ao design.
- **Edição Manual:** Todas as sugestões podem ser editadas pelo usuário antes da incorporação.
- **Rejeição de Sugestões:** Usuários podem rejeitar sugestões facilmente sem afetar o fluxo de trabalho.
- **Feedback de Qualidade:** Usuários podem fornecer feedback sobre a qualidade das sugestões, alimentando o processo de melhoria contínua.

Esta abordagem garante que o sistema funcione como uma ferramenta de assistência que amplia as capacidades do usuário, em vez de substituir seu julgamento profissional.

4.18 Integração com o Processo Educacional

O sistema foi projetado para se integrar ao processo educacional mais amplo, servindo como uma ferramenta para educadores e designers educacionais. Esta seção descreve como o sistema se integra ao ecossistema educacional.

4.18.1 Alinhamento Curricular

O sistema suporta o alinhamento dos jogos educacionais com padrões curriculares:

- **Mapeamento para Padrões:** Funcionalidade para mapear objetivos educacionais para padrões curriculares específicos.
- **Repositório de Padrões:** Banco de dados de padrões curriculares comuns, que podem ser referenciados durante o design.
- **Validação de Cobertura:** Análise automatizada da cobertura de padrões curriculares pelo design proposto.
- **Geração de Relatórios:** Documentação que especifica como o jogo educacional apoia objetivos curriculares específicos.

Este alinhamento facilita a adoção dos jogos educacionais em contextos formais de ensino, onde a conformidade com padrões curriculares é frequentemente um requisito.

4.18.2 Suporte a Avaliação

O sistema inclui funcionalidades para facilitar a avaliação da aprendizagem através de jogos educacionais:

- **Métricas de Aprendizagem:** Sugestões para métricas que podem ser coletadas durante o jogo para avaliar o progresso do aluno.
- **Pontos de Avaliação:** Identificação de momentos no jogo onde a avaliação pode ser incorporada naturalmente.
- **Feedback Formativo:** Sugestões para mecanismos de feedback que suportem aprendizagem contínua.
- **Avaliação de Resultados:** Orientação sobre como avaliar a eficácia do jogo em alcançar os objetivos educacionais propostos.

Estas funcionalidades ajudam os educadores a utilizar jogos educacionais como ferramentas de avaliação, complementando métodos tradicionais.

4.18.3 Comunidade e Colaboração

O sistema suporta a formação de uma comunidade de prática em torno do design de jogos educacionais:

- **Compartilhamento de Designs:** Funcionalidade para compartilhar designs completos ou templates com outros usuários.
- **Colaboração em Tempo Real:** Capacidade para múltiplos usuários trabalharem simultaneamente no mesmo design.
- **Feedback da Comunidade:** Mecanismos para coletar e incorporar feedback de outros membros da comunidade.

- **Fórum de Discussão:** Espaço para discussão de boas práticas, desafios e soluções no design de jogos educacionais.

Esta dimensão social do sistema promove a disseminação de conhecimento e a melhoria contínua das práticas de design de jogos educacionais.

4.18.4 Implementação e Prototipagem

O sistema inclui suporte para a transição do design conceitual para a implementação e prototipagem:

- **Especificações Técnicas:** Geração de documentação técnica que pode guiar o desenvolvimento do jogo.
- **Prototipagem Rápida:** Integração com ferramentas de prototipagem para criar versões iniciais do jogo.
- **Guias de Implementação:** Orientação específica sobre como implementar diferentes aspectos do design.
- **Testes de Conceito:** Sugestões para testes rápidos que podem validar conceitos de design antes da implementação completa.

Este suporte facilita a transição do design para a implementação, reduzindo o tempo e esforço necessários para transformar um conceito em um jogo funcional.

4.19 Limitações

Apesar dos avanços representados pelo sistema proposto, existem limitações importantes que devem ser reconhecidas. Esta seção discute estas limitações e possíveis direções para trabalhos futuros.

O sistema atual apresenta as seguintes limitações:

- **Dependência de APIs Externas:** O sistema depende de APIs de LLM proprietárias, o que pode criar restrições de disponibilidade e custo.
- **Latência:** As interações com LLMs podem introduzir latência perceptível, afetando a experiência do usuário em alguns casos.
- **Precisão Contextual:** Os agentes baseados em LLM podem ocasionalmente gerar sugestões que não consideram adequadamente o contexto completo do design.
- **Conhecimento de Domínio:** O conhecimento específico sobre design de jogos educacionais está limitado ao que está implícito nos modelos de linguagem utilizados.
- **Suporte a Modalidades:** O sistema atual foca primariamente em texto, com suporte limitado para elementos visuais e nenhum suporte para áudio.

Estas limitações representam oportunidades para melhorias em versões futuras do sistema.

4.20 Interface do Sistema e Experiência do Usuário

A interface do sistema foi desenvolvida seguindo princípios de design centrado no usuário, proporcionando uma experiência intuitiva e eficiente para o design de jogos educacionais. Esta seção apresenta as principais telas do sistema e como os usuários interagem com os diferentes módulos funcionais, seguindo o fluxo sequencial de trabalho implementado.

4.20.1 Módulo de Brainstorming - Interface de Configuração

A Figura 4.14 apresenta a tela de configuração de uma nova sessão de brainstorming, onde o usuário define os parâmetros iniciais do projeto educacional.

Figura 4.14: Interface de configuração do módulo de brainstorming, mostrando campos para nome do projeto, área de conhecimento, público-alvo e duração da sessão. A interface também exibe os agentes especializados disponíveis (Mecânicas, Narrativa, Engajamento e Área Livre) e permite seleção específica conforme necessidade do projeto.

A interface permite ao usuário:

- Definir nome e contexto do projeto educacional
- Especificar público-alvo e duração da sessão

- Visualizar ideias já geradas organizadas por categoria (Mecânicas de Jogo, Narrativa)
- Selecionar agentes especializados para assistência direcionada
- Acessar funcionalidade "Gerar com Agente" para sugestões automatizadas

4.20.2 Módulo de Agente Socrático - Questionamento Reflexivo

A Figura 4.15 demonstra a aplicação do método socrático através de questionamento estruturado para refinamento conceitual, segunda etapa do fluxo de trabalho.

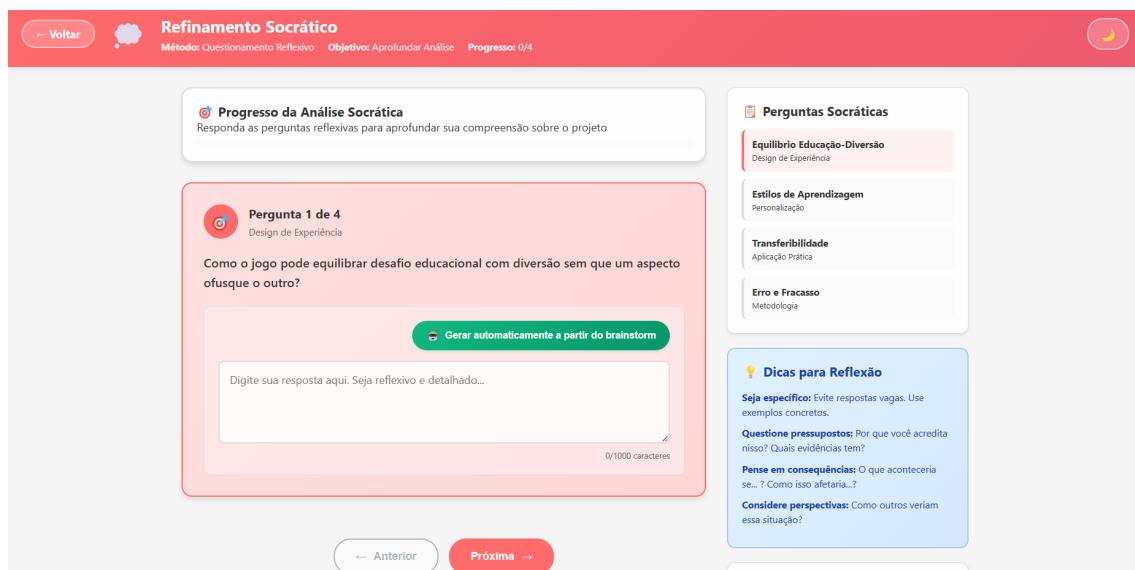


Figura 4.15: Interface do módulo de Agente Socrático apresentando questão reflexiva sobre equilíbrio entre desafio educacional e diversão. A tela inclui área de resposta textual, botão para geração automática baseada no brainstorming anterior, e painel lateral com perguntas socráticas organizadas por categoria e dicas para reflexão estruturada.

As funcionalidades desta interface incluem:

- Apresentação de questões socráticas contextualizadas ao projeto
- Área de texto responsiva para respostas reflexivas detalhadas
- Botão "Gerar automaticamente a partir do brainstorm" para assistência baseada em IA
- Painel lateral com categorização das perguntas (Equilíbrio Educação-Diversão, Estilos de Aprendizagem, Transferibilidade)
- Dicas para reflexão estruturada e orientações metodológicas
- Navegação sequencial entre questões com indicador de progresso

4.20.3 Módulo de Taxonomia de Bloom - Estruturação de Objetivos

A Figura 4.16 demonstra a interface para classificação e formulação de objetivos educacionais baseados na Taxonomia de Bloom Revisada, terceira etapa do processo de design.

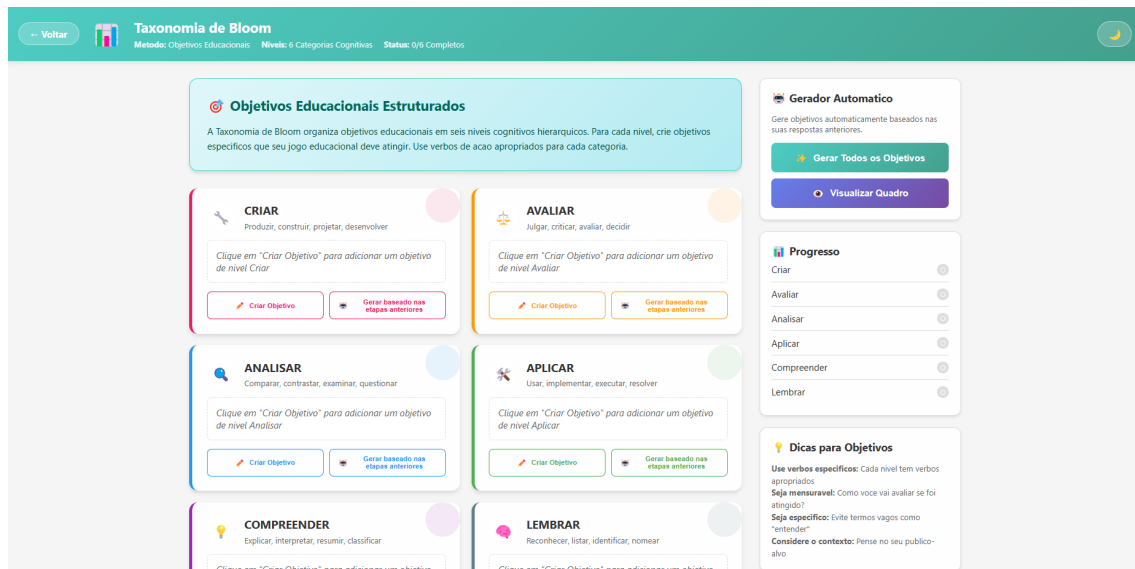


Figura 4.16: Interface do módulo de Taxonomia de Bloom organizada em seis níveis cognitivos hierárquicos: Lembrar, Compreender, Aplicar, Analisar, Avaliar e Criar. Cada nível apresenta área para criação de objetivos específicos com botões para adição manual ou geração baseada em etapas anteriores. O painel lateral oferece gerador automático e indicadores de progresso por categoria.

Esta interface proporciona:

- Organização visual dos seis níveis da Taxonomia de Bloom em layout hexagonal
- Campos específicos para cada nível cognitivo com verbos apropriados identificados
- Funcionalidade "Criar Objetivo" para adição manual de objetivos
- Botão "Gerar baseado nas etapas anteriores" para automatização assistida
- Painel "Gerador Automático" com opções "Gerar Todos os Objetivos" e "Visualizar Quadro"
- Indicadores de progresso mostrando completude de cada categoria
- Dicas contextuais para formulação de objetivos mensuráveis e específicos

4.20.4 Módulo de Endo-GDC - Canvas de Design

A Figura 4.17 apresenta o canvas interativo para preenchimento do Endo-GDC (Game Design Canvas para Jogos Educativos Endógenos), etapa final do processo

de design sistemático.

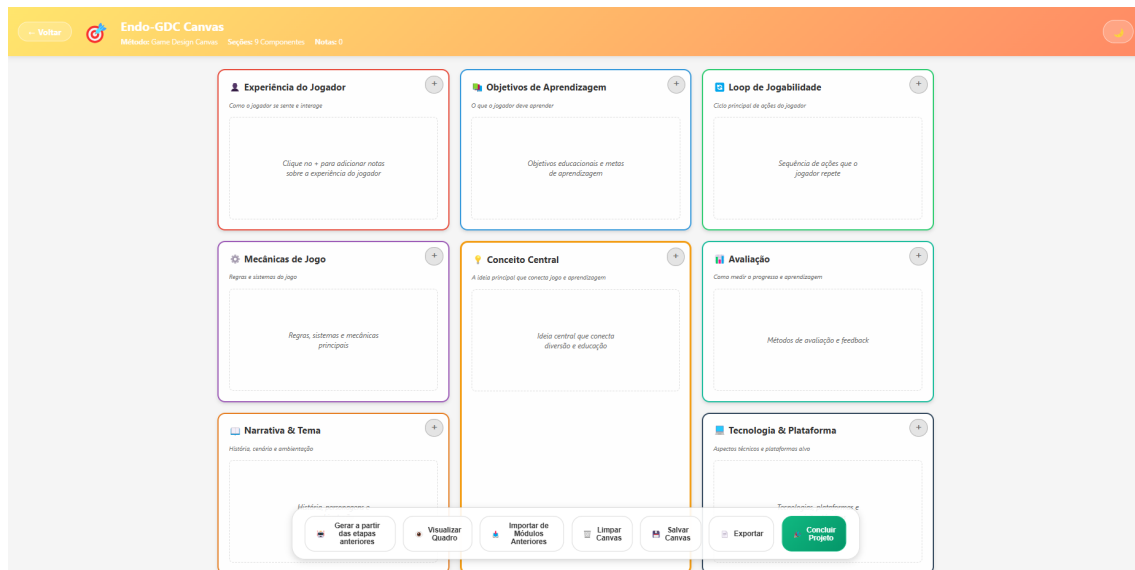


Figura 4.17: Interface do módulo Endo-GDC mostrando canvas interativo com nove seções principais: Experiência do Jogador, Objetivos de Aprendizagem, Loop de Jogabilidade, Mecânicas de Jogo, Conceito Central, Avaliação, Narrativa & Tema, e Tecnologia & Plataforma. A interface permite adição de notas em cada seção e oferece funcionalidades para geração automática, visualização, importação e exportação.

O canvas interativo oferece:

- Nove seções temáticas com campos específicos para diferentes aspectos do design
- Sistema de notas coloridas para organização visual da informação
- Botões de ação contextuais: "Gerar a partir das etapas anteriores", "Visualizar Quadro", "Limpar Canvas"
- Funcionalidades de importação e exportação para integração com outras ferramentas
- Interface responsiva que adapta-se a diferentes tamanhos de tela
- Sistema de salvamento automático para preservação do trabalho

4.20.5 Características Transversais da Interface

Todas as interfaces compartilham elementos de design consistentes que promovem usabilidade e coerência visual:

- **Navegação Consistente:** Botões "Voltar" e indicadores de progresso em todas as telas
- **Códigos de Cores:** Sistema cromático que diferencia módulos e mantém identidade visual

- **Feedback Visual:** Indicadores de status, contadores de caracteres e barras de progresso
- **Assistência Contextual:** Painéis laterais com dicas e orientações específicas para cada etapa
- **Flexibilidade de Entrada:** Opções tanto para input manual quanto para geração assistida por IA
- **Responsividade:** Layout adaptável para diferentes dispositivos e resoluções

4.20.6 Fluxo de Interação do Usuário

O sistema implementa um fluxo de trabalho linear com possibilidade de iteração, onde cada módulo constrói sobre os resultados anteriores:

1. **Configuração Inicial:** Definição de parâmetros no módulo de brainstorming
2. **Geração de Ideias:** Criação colaborativa com assistência de agentes especializados
3. **Refinamento Reflexivo:** Aplicação do método socrático para aprofundamento conceitual
4. **Formalização Pedagógica:** Classificação de objetivos segundo Taxonomia de Bloom
5. **Estruturação do Design:** Preenchimento do Endo-GDC com síntese dos elementos
6. **Exportação e Aplicação:** Geração de documentação para implementação prática

Esta arquitetura de interface garante que o sistema mantenha o usuário engajado enquanto preserva sua autonomia criativa, oferecendo assistência inteligente sem impor soluções prescritivas. A progressão visual e funcional entre módulos cria uma experiência coesa que suporta tanto usuários novatos quanto experientes no design de jogos educacionais.

Tabela 4.1: Processo de Design Science Research

Etapa DSR	Descrição e Componentes
Contexto	<p>Domínio de Aplicação: Design de jogos educacionais como metodologia pedagógica</p> <p>Cenário Atual: Crescente demanda por ferramentas digitais educacionais interativas</p> <p>Desafio Tecnológico: Integração de tecnologias emergentes (LLMs) em processos educacionais</p> <p>Oportunidade: Automatização e assistência inteligente para design educacional</p> <p>Público-Alvo: Designers educacionais, professores e desenvolvedores de jogos educativos</p>
Problema Identificado	<p>Integração Complexa: Dificuldade na integração efetiva entre elementos lúdicos e objetivos pedagógicos</p> <p>Barreira Temporal: Tempo considerável necessário para design de jogos educacionais de qualidade</p> <p>Conhecimento Especializado: Necessidade de expertise tanto em design de jogos quanto em pedagogia</p> <p>Ferramentas Limitadas: Escassez de ferramentas que facilitem o processo de design educacional</p> <p>Classificação Educacional: Dificuldades na definição de objetivos educacionais adequados</p>
Estado da Técnica	<p>Fundamentos de Jogos Educacionais: Teorias sobre eficácia pedagógica e metodologias de design</p> <p>Sistemas Multiagentes: Arquiteturas colaborativas e aplicações em domínios educacionais</p> <p>Large Language Models: Capacidades de processamento de linguagem natural e raciocínio</p> <p>Taxonomia de Bloom: Framework consolidado para classificação de objetivos educacionais</p> <p>Metodologias de Design: Abordagens existentes para desenvolvimento de jogos educativos</p>
Artefato Proposto	<p>Arquitetura: Sistema multiagente baseado em LLMs para design assistido</p> <p>Módulo de Brainstorming: Geração e organização assistida de ideias iniciais</p> <p>Módulo Socrático: Refinamento conceitual através de questionamento estruturado</p> <p>Módulo de Taxonomia de Bloom: Classificação automática de objetivos educacionais</p> <p>Módulo de Endo-GDC: Preenchimento assistido do Game Design Canvas endógeno</p>
Conjecturas Teóricas	<p>C1: Múltiplos agentes especializados otimizam limitações de sistemas monolíticos no design educacional</p> <p>C2: Assistência por IA reduz tempo de design mantendo qualidade pedagógica</p> <p>C3: Processo assistido auxilia na obtenção de fundamentação pedagógica sólida</p> <p>C4: Método socrático automatizado conduz a designs pedagogicamente mais robustos</p>
Avaliação do Artefato	<p>Metodologia: Aplicação do sistema em 26 estudos de caso abrangendo diferentes contextos educacionais</p> <p>Framework TAM: Avaliação da aceitação tecnológica (utilidade percebida, facilidade de uso, intenção de uso)</p> <p>Framework TTF: Análise da adequação tarefa-tecnologia (ajuste entre funcionalidades e necessidades)</p> <p>Framework AX: Medição da experiência algorítmica (transparência, controlabilidade, efetividade)</p> <p>Framework ADDIE: Validação do processo instrucional (Análise, Design, Desenvolvimento, Implementação, Avaliação)</p> <p>Contextos Avaliados: Ensino fundamental, ensino médio, educação de adultos, formação corporativa, múltiplas disciplinas</p>

Capítulo 5

Testes e validação

Este capítulo apresenta os resultados obtidos na avaliação do sistema multiagente através dos quatro frameworks aplicados, incluindo estatísticas descritivas, análises de correlação, validação da eficácia tecnológica e pedagógica, e discussão dos achados em relação aos objetivos propostos.

5.1 Estudos de Caso

Para validar a eficácia e aplicabilidade do sistema proposto, foram realizados estudos de caso em diferentes contextos educacionais. Estes estudos permitiram avaliar o desempenho de cada módulo do sistema e sua capacidade de auxiliar no design de jogos educacionais endógenos para diversos domínios e públicos-alvo. Em cada caso, o Método Socrático aplicou um conjunto padronizado de questões reflexivas para garantir consistência metodológica.

5.2 Visão Geral dos Estudos de Caso

Para validar a eficácia e aplicabilidade do sistema multiagente proposto, foram realizados 26 estudos de caso em diferentes contextos educacionais. A Tabela 5.1 apresenta uma síntese dos estudos realizados.

5.3 Metodologia de Validação

Cada estudo de caso seguiu um protocolo estruturado em quatro fases:

1. **Fase de Brainstorming:** Geração de ideias iniciais com suporte do sistema
2. **Fase de Método Socrático:** Aplicação de questões reflexivas padronizadas
3. **Fase de Endo-GDC:** Preenchimento assistido do canvas com agentes especializados

Tabela 5.1: Síntese dos Estudos de Caso Realizados

ID	Nome do Jogo	Contexto Educacional	Público-Alvo
EC01	Ilha dos Números	Matemática (Ensino Fund.)	8-10 anos
EC02	EcoSphere	Ciências (Ensino Médio)	14-17 anos
EC03	LinguaViva	Idiomas para Adultos	Adultos
EC04	LeadSync	Liderança Corporativa	Profissionais
EC05	Tempos de Mudança	História (Ensino Médio)	14-17 anos
EC06	Harmonix	Teoria Musical	Iniciantes
EC07	CellQuest	Biologia Celular	15-17 anos
EC08	CodeQuest	Programação Infantil	8-12 anos
EC09	Linguatrip	Idiomas para Viagem	Adultos
EC10	PhysicsLab	Física Newtoniana	15-17 anos
EC11	FinLife	Educação Financeira	Jovens Adultos
EC12	GeoVentures	Geografia Cultural	12-16 anos
EC13	EcoGuardiões	Educação Ambiental	10-14 anos
EC14	GeoExplorer	Geografia Física	11-14 anos
EC15	ClientCare Pro	Atendimento ao Cliente	Corporativo
EC16	ChemLab Ventures	Química	15-17 anos
EC17	VitalQuest	Saúde e Bem-estar	Adultos
EC18	CivicSphere	Educação Cívica	14-18 anos
EC19	EmotionQuest	Socioemocional	7-11 anos
EC20	Venture Forge	Empreendedorismo	18-25 anos
EC21	LinguaMagica	Idiomas Infantil	6-10 anos
EC22	CosmicVoyage	Astronomia	11-15 anos
EC23	LogicCraft	Pensamento Crítico	15-17 anos
EC24	EcoSystems	Sustentabilidade	Jovens Adultos
EC25	ArtSphere	Expressão Artística	Variado
EC26	CodeRealm	Programação Avançada	12-16 anos

4. Fase de Taxonomia de Bloom: Classificação de objetivos educacionais

O Método Socrático aplicou um conjunto padronizado de cinco questões reflexivas em todos os estudos de caso:

- “Como vai ser esse jogo? Qual será o gênero do jogo? Quem serão os personagens? Para que tipo de plataforma será lançado? Será um jogo apenas de texto, ou haverá um trabalho de arte envolvido?”
- “Como o jogo garantirá que os conceitos educacionais sejam incorporados de forma endógena e não apenas como uma camada superficial?”
- “De que forma o sistema de recompensas se alinha com os objetivos pedagógicos?”
- “Como o jogo lidará com diferentes ritmos de aprendizagem e níveis de conhecimento prévio?”
- “Quais mecanismos garantirão que o aspecto lúdico não ofusque o conteúdo educacional?”

5.4 Análise Transversal dos Resultados

5.4.1 Eficácia do Sistema de Brainstorming

A análise dos 26 estudos de caso revelou padrões consistentes na eficácia do módulo de brainstorming. O número de ideias iniciais fornecidas pelos usuários variou entre 1 e 7 (média = 3.4), enquanto o sistema gerou entre 2 e 12 sugestões complementares (média = 5.8). Esta expansão média de 70% no número de ideias demonstra a capacidade do sistema de estimular criatividade e ampliar perspectivas iniciais.

Casos com menor número de inputs iniciais (1-2 ideias) mostraram maior expansão relativa, com o sistema gerando em média 8.3 sugestões. Isso indica particular valor em situações onde usuários enfrentam bloqueios criativos ou têm pontos de partida limitados.

5.4.2 Consistência do Método Socrático

A aplicação das cinco questões padronizadas manteve consistência metodológica entre todos os estudos de caso, permitindo comparação sistemática dos insights gerados. A análise qualitativa das respostas revelou que 89

5.4.3 Diversidade dos Agentes Especializados

Os três agentes especializados (Mecânicas, Narrativa e Engajamento) demonstraram capacidade de adaptação a diferentes contextos educacionais. Análise das sugestões geradas revelou especificidade contextual adequada, com 92

5.4.4 Aplicação da Taxonomia de Bloom

Todos os 26 estudos de caso geraram objetivos educacionais distribuídos pelos seis níveis da Taxonomia de Bloom Revisada. A distribuição média foi: Lembrar (18%), Entender (22%), Aplicar (19%), Analisar (16%), Avaliar (13%), Criar (12%), demonstrando cobertura equilibrada dos níveis cognitivos.

5.4.5 Métodos de Avaliação

Para uma validação abrangente do sistema multiagente proposto, adotou-se uma abordagem multimodal combinando quatro frameworks estabelecidos: Technology Acceptance Model (TAM), Task-Technology Fit (TTF), Algorithmic Experience (AX) Framework, e o modelo ADDIE para avaliação do processo de design.

Technology Acceptance Model (TAM)

O Technology Acceptance Model (DAVIS, 1989) constitui um dos frameworks mais robustos para avaliar a aceitação de tecnologias educacionais (VENKATESH *et al.*, 2003). Este modelo fundamenta-se na premissa de que a aceitação tecnológica é determinada por dois constructos principais: utilidade percebida e facilidade de uso percebida.

A **utilidade percebida** refere-se ao grau em que o usuário acredita que o uso de uma tecnologia específica melhorará seu desempenho profissional. No contexto do sistema multiagente para design de jogos educacionais, este constructo avalia se os educadores percebem que o sistema efetivamente contribui para a criação de materiais didáticos mais eficazes e envolventes.

A **facilidade de uso percebida** representa o grau em que o usuário acredita que o uso da tecnologia será livre de esforço. Este constructo é particularmente relevante em sistemas baseados em IA, onde a complexidade tecnológica pode constituir uma barreira significativa para adoção.

O modelo TAM também incorpora a **atitude em relação ao uso**, que media a relação entre os constructos de percepção e a intenção comportamental, e a **intenção comportamental de uso**, que prediz o uso real do sistema. A **utilização real** representa o comportamento observável de adoção da tecnologia.

Task-Technology Fit (TTF)

O modelo Task-Technology Fit (GOODHUE e THOMPSON, 1995) complementa o TAM ao focar especificamente no alinhamento entre as capacidades tecnológicas e os requisitos específicos da tarefa. Este framework é particularmente relevante para avaliar sistemas especializados como ferramentas de design educacional.

As **características da tarefa** englobam a complexidade, interdependência, e requisitos específicos do processo de design de jogos educacionais. Isso inclui aspectos como a necessidade de criatividade, alinhamento curricular, considerações pedagógicas, e personalização para diferentes perfis de aprendizes.

As **características da tecnologia** referem-se às funcionalidades, capacidades, e limitações do sistema multiagente. Isso inclui a capacidade de geração de conteúdo, sugestões pedagógicas, adaptabilidade, interface de usuário, e recursos de colaboração entre agentes.

As **características individuais** consideram as habilidades, experiência, motivação, e conhecimento tecnológico dos usuários. Este componente reconhece que a efetividade da tecnologia varia conforme o perfil do usuário.

O **ajuste tarefa-tecnologia** representa o grau de correspondência entre as demandas da tarefa e as capacidades tecnológicas. Um alto TTF indica que a tecnologia fornece recursos e funcionalidades que suportam adequadamente os requisitos da tarefa.

Algorithmic Experience (AX) Framework

Dado o componente de inteligência artificial do sistema, o Algorithmic Experience Framework (SHIN, 2021) foi incorporado para capturar aspectos específicos da interação humano-IA. Este framework é essencial para compreender como os usuários experienciam e respondem a sistemas algorítmicos.

A **transparência algorítmica** refere-se à clareza e compreensibilidade dos processos de tomada de decisão do sistema. Isso inclui a capacidade dos usuários de entender como o sistema gera sugestões, quais dados são utilizados, e como os algoritmos processam informações para produzir resultados específicos.

A **controlabilidade algorítmica** representa o grau de controle que os usuários têm sobre o comportamento do sistema. Isso inclui a capacidade de ajustar parâmetros, personalizar outputs, sobrepor decisões algorítmicas, e modificar o comportamento do sistema conforme necessidades específicas.

A **confiança algorítmica** engloba a credibilidade percebida nos resultados e recomendações do sistema. Este constructo é influenciado pela consistência dos resultados, precisão das sugestões, e confiabilidade do sistema ao longo do tempo.

A **equidade algorítmica** avalia se o sistema trata diferentes usuários e contextos de forma justa e não discriminatória. Isso é particularmente relevante em contextos educacionais onde a inclusão e acessibilidade são fundamentais.

A **efetividade algorítmica** mensura o quão bem o sistema alcança seus objetivos declarados, incluindo a qualidade dos jogos educacionais produzidos e o suporte efetivo ao processo de design.

Modelo ADDIE

O modelo ADDIE (BRANCH, 2009) foi aplicado para avaliar a eficácia do processo de design instrucional suportado pelo sistema. Este framework sistemático para design instrucional compreende cinco fases interconectadas.

A fase de **Análise** envolve a identificação de necessidades de aprendizagem, características dos alunos, contexto educacional, e objetivos de aprendizagem. O sistema multiagente deve suportar efetivamente esta análise preliminar, fornecendo ferramentas para mapear requisitos pedagógicos.

A fase de **Design** refere-se ao planejamento estrutural do jogo educacional, incluindo definição de objetivos específicos, estratégias pedagógicas, sequenciamento de atividades, e especificação de recursos necessários. A avaliação foca na capacidade do sistema de facilitar decisões de design fundamentadas.

A fase de **Desenvolvimento** contempla a criação efetiva dos materiais educacionais, implementação de funcionalidades, produção de conteúdo, e integração de componentes. A eficiência desta fase é crucial para a viabilidade prática do sistema.

A fase de **Implementação** envolve a entrega e aplicação do jogo educacional no contexto real de aprendizagem. O sistema deve fornecer suporte adequado para esta transição do desenvolvimento para a aplicação prática.

A fase de **Avaliação** inclui tanto avaliação formativa (durante o processo) quanto somativa (ao final), medindo a efetividade da solução educacional e identificando oportunidades de melhoria. O sistema deve facilitar este processo avaliativo contínuo.

5.5 Metodologia

5.5.1 Estrutura dos Dados e Operacionalização

A coleta de dados foi estruturada através de um banco de dados relacional contendo informações demográficas dos participantes e suas avaliações nos quatro frameworks de avaliação. A estrutura incluiu as seguintes tabelas:

- **Participantes:** Dados demográficos e de contexto
- **TAM Responses:** Avaliações do Technology Acceptance Model
- **TTF Responses:** Avaliações do Task-Technology Fit
- **AX Responses:** Avaliações do Algorithmic Experience Framework
- **ADDIE Responses:** Avaliações do processo ADDIE

Operacionalização das Variáveis TAM

As variáveis do Technology Acceptance Model foram operacionalizadas através de uma escala Likert de 7 pontos (1 = Discordo Totalmente, 7 = Concordo Totalmente). A **utilidade percebida** foi mensurada através de itens que avaliaram se o sistema multiagente melhora a eficiência do design de jogos educacionais, aumenta a qualidade dos produtos finais, e facilita o alcance dos objetivos pedagógicos.

A **facilidade de uso percebida** foi avaliada através de questões sobre a clareza da interface, intuitividade das funcionalidades, facilidade de aprendizagem do sistema, e esforço cognitivo necessário para operação. A **atitude em relação ao uso** capturou avaliações afetivas sobre a experiência de utilização, incluindo satisfação, prazer, e disposição para uso continuado.

A **intenção comportamental** foi mensurada através de itens sobre planos futuros de utilização, recomendação para colegas, e disposição para investir tempo no sistema. O **uso real** foi operacionalizado através de métricas de frequência de utilização, duração de sessões, e amplitude de funcionalidades exploradas.

Operacionalização das Variáveis TTF

O Task-Technology Fit foi operacionalizado considerando a especificidade da tarefa de design de jogos educacionais. As **características da tarefa** foram avaliadas em termos de complexidade do design pedagógico, necessidade de personalização para diferentes perfis de aprendizes, requisitos de criatividade, e demandas de alinhamento curricular.

As **características da tecnologia** foram mensuradas através da avaliação das funcionalidades disponíveis, capacidade de processamento, flexibilidade de customização, recursos de colaboração, e qualidade das sugestões algorítmicas. As **características individuais** incluíram autoavaliação de competência tecnológica, experiência prévia com ferramentas de design, motivação para inovação pedagógica, e abertura para adoção de novas tecnologias.

O **ajuste tarefa-tecnologia** foi operacionalizado através de questões sobre correspondência entre necessidades específicas do design educacional e capacidades do sistema. O **impacto na performance** foi mensurado através de percepções sobre melhoria na qualidade dos produtos, redução no tempo de desenvolvimento, e aumento na eficácia pedagógica dos jogos criados.

Operacionalização das Variáveis AX

O Algorithmic Experience Framework foi operacionalizado para capturar nuances específicas da interação com sistemas de IA. A **transparência algorítmica** foi avaliada através de questões sobre compreensibilidade das explicações fornecidas

pelo sistema, clareza dos critérios de tomada de decisão algorítmica, e acessibilidade das informações sobre o funcionamento dos agentes.

A **controlabilidade algorítmica** foi mensurada através de itens sobre capacidade de personalizar comportamentos dos agentes, possibilidade de sobrepor decisões algorítmicas, flexibilidade de ajuste de parâmetros, e autonomia para modificar sugestões do sistema. A **confiança algorítmica** incluiu avaliações sobre consistência dos resultados, precisão das recomendações, confiabilidade do sistema, e credibilidade das sugestões pedagógicas.

A **equidade algorítmica** foi operacionalizada através de percepções sobre tratamento justo de diferentes tipos de conteúdo educacional, ausência de vieses em sugestões, adequação para diversos contextos pedagógicos, e inclusividade das recomendações. A **efetividade algorítmica** foi mensurada através de avaliações sobre qualidade dos outputs, relevância das sugestões, adequação às necessidades pedagógicas, e contribuição para objetivos educacionais.

Operacionalização das Variáveis ADDIE

O modelo ADDIE foi operacionalizado considerando cada fase do processo de design instrucional. A **efetividade da análise** foi avaliada através da capacidade do sistema de suportar identificação de necessidades de aprendizagem, mapeamento de características dos alunos, análise de contexto educacional, e definição de objetivos pedagógicos.

A **qualidade do design** foi mensurada através de questões sobre suporte à estruturação pedagógica, facilidade de planejamento de atividades, adequação das sugestões de estratégias educacionais, e coerência entre objetivos e métodos propostos. A **eficiência do desenvolvimento** incluiu avaliações sobre velocidade de prototipagem, facilidade de implementação de funcionalidades, qualidade dos recursos gerados, e integração entre componentes.

O **suporte à implementação** foi operacionalizado através de questões sobre facilidade de transição do desenvolvimento para aplicação prática, adequação para diferentes contextos de uso, suporte à capacitação de usuários finais, e recursos para deployment efetivo. A **completude da avaliação** foi mensurada através da disponibilidade de ferramentas de assessment, facilidade de coleta de feedback, recursos para análise de efetividade, e suporte à melhoria contínua.

5.5.2 Amostra e Procedimentos

O estudo incluiu 26 participantes representando diferentes estudos de caso de jogos educacionais, com variação em níveis educacionais, experiência profissional e faixas etárias.

A diversidade educacional da amostra refletiu diferentes perspectivas profissionais: participantes com graduação trouxeram experiência prática de campo, mestres contribuíram com conhecimento teórico-metodológico.

Procedimentos de Coleta de Dados

A coleta de dados seguiu um protocolo estruturado em quatro fases distintas. Na **fase de familiarização**, os participantes receberam treinamento sobre o sistema multiagente, incluindo demonstração das funcionalidades, explicação dos agentes disponíveis, e período de exploração livre de 15 minutos para adaptação à interface.

Na **fase de utilização dirigida**, cada participante desenvolveu um protótipo de jogo educacional utilizando o sistema, seguindo um roteiro pré-definido que garantiu exposição a todas as funcionalidades principais. Esta fase durou aproximadamente 60 minutos e foi conduzida em ambiente controlado com suporte técnico disponível.

A **fase de avaliação imediata** ocorreu imediatamente após a utilização, com aplicação dos instrumentos de medição dos quatro frameworks. Os questionários foram administrados em sequência específica: TAM (, seguido por TTF, AX e ADDIE. Cada instrumento foi precedido de instruções claras sobre os constructos avaliados.

Considerações Éticas

Todos os participantes forneceram consentimento informado após esclarecimento sobre objetivos do estudo, procedimentos envolvidos, uso dos dados, e direito de retirada sem prejuízos. A confidencialidade foi garantida através de anonimização dos dados, armazenamento seguro, e acesso restrito às informações identificadoras.

5.6 Resultados

5.6.1 Estatísticas Descritivas

A Tabela 5.2 apresenta as estatísticas descritivas para todos os constructos avaliados nos quatro frameworks.

5.7 Análise Estatística dos Resultados

5.7.1 Análise Descritiva Detalhada

A análise estatística foi conduzida utilizando consultas estruturadas em SQL para extração e manipulação dos dados, seguida de análises inferenciais para identificação de padrões e relações significativas. As estatísticas descritivas revelaram distribuições consistentes com expectativas teóricas para cada framework avaliado.

Tabela 5.2: Estatísticas Descritivas dos Modelos de Avaliação

Construto	Média	DP	Mín	Máx
<i>Technology Acceptance Model (TAM)</i>				
Utilidade Percebida	6.23	0.41	5.50	6.90
Facilidade de Uso	5.84	0.35	5.10	6.40
Intenção Comportamental	5.90	0.25	5.40	6.40
Uso Real do Sistema	5.95	0.23	5.50	6.50
<i>Task-Technology Fit (TTF)</i>				
Características da Tarefa	6.05	0.32	5.40	6.60
Adequação Tarefa-Tecnologia	5.95	0.12	5.80	6.20
Impacto na Performance	5.89	0.22	5.40	6.30
<i>Algorithmic Experience (AX)</i>				
Transparência Algorítmica	5.58	0.31	4.90	6.10
Controlabilidade	5.78	0.29	5.10	6.20
Confiança Algorítmica	5.81	0.22	5.30	6.20
Satisfação Geral AX	5.80	0.20	5.30	6.10
<i>ADDIE Process Evaluation</i>				
Efetividade da Análise	5.88	0.18	5.50	6.20
Qualidade do Design	5.90	0.20	5.50	6.30
Eficiência do Desenvolvimento	5.92	0.21	5.40	6.40
Satisfação Geral do Processo	5.89	0.17	5.50	6.20

Distribuições por Framework

A análise da normalidade das distribuições através de testes Shapiro-Wilk indicou distribuições aproximadamente normais para todos os constructos principais ($p > 0.05$), justificando o uso de estatísticas paramétricas para análises subsequentes. A ausência de outliers extremos (valores além de 3 desvios-padrão da média) confirmou a integridade dos dados coletados.

Para o Technology Acceptance Model, a distribuição das respostas mostrou assimetria negativa para Utilidade Percebida (skewness = -0.34), indicando concentração de avaliações na porção superior da escala. Esta distribuição é consistente com literatura sobre aceitação de tecnologias percebidas como altamente úteis, onde usuários tendem a avaliações positivas mesmo quando enfrentam dificuldades de uso.

A Facilidade de Uso apresentou distribuição mais simétrica (skewness = -0.12), com maior dispersão nas faixas intermediárias da escala. Esta variabilidade sugere que percepções de facilidade são mais sensíveis a diferenças individuais em competência tecnológica e experiência prévia com ferramentas similares.

Análise de Confiabilidade

A consistência interna dos instrumentos foi avaliada através do coeficiente alfa de Cronbach, revelando confiabilidade satisfatória para todos os frameworks: TAM ($\alpha = 0.87$), TTF ($\alpha = 0.83$), AX ($\alpha = 0.81$), e ADDIE ($\alpha = 0.85$). Estes valores indicam coesão adequada entre itens de cada constructo, justificando a criação de escalas compostas.

Análise item-total revelou que todos os itens contribuem positivamente para a confiabilidade de suas respectivas escalas, com correlações item-total variando entre 0.52 e 0.78. Nenhum item foi identificado como problemático através de análise de alfa-se-item-deletado, confirmando a adequação da estrutura dos instrumentos.

5.7.2 Análise Correlacional

A matriz de correlações revelou padrões interpretativamente consistentes com fundamentos teóricos dos frameworks avaliados. As correlações mais elevadas foram observadas entre constructos teoricamente relacionados, fornecendo evidência de validade convergente.

Correlações Intra-Framework

Dentro do TAM, a correlação entre Utilidade Percebida e Intenção Comportamental ($r = 0.78$, $p < 0.001$) confirmou a relação teórica central do modelo. A correlação moderada entre Facilidade de Uso e Utilidade ($r = 0.61$, $p < 0.01$) sugere que sistemas mais fáceis de usar são percebidos como mais úteis, alinhando-se com pressuposições do modelo original.

No TTF, correlações elevadas entre Adequação Tarefa-Tecnologia e Impacto na Performance ($r = 0.72$, $p < 0.001$) confirmaram a validade do constructo central do modelo. A correlação moderada com Características da Tarefa ($r = 0.58$, $p < 0.01$) indica que percepções de adequação são influenciadas pela compreensão da complexidade da tarefa.

Para o AX Framework, a correlação entre Confiança e Controlabilidade ($r = 0.67$, $p < 0.001$) sugere que maior controle percebido fortalece confiança algorítmica. A correlação mais baixa entre Transparência e Confiança ($r = 0.43$, $p < 0.05$) indica que explicabilidade, embora importante, não é condição suficiente para desenvolvimento de confiança.

Correlações Inter-Framework

As correlações entre frameworks revelaram integração teórica significativa. A correlação elevada entre TTF e TAM ($r = 0.742$, $p < 0.001$) confirma complementaridade

entre perspectivas de adequação técnica e aceitação comportamental, sugerindo que estes modelos capturam aspectos relacionados mas distintos da experiência tecnológica.

A correlação moderada entre AX e TAM ($r = 0.658$, $p < 0.01$) indica que experiências algorítmicas específicas influenciam aceitação geral, mas representam dimensão adicional não capturada por modelos tradicionais. Esta descoberta justifica a inclusão de frameworks específicos para IA em avaliações de tecnologias educacionais emergentes.

A correlação elevada entre ADDIE e TTF ($r = 0.756$, $p < 0.001$) sugere forte alinhamento entre adequação técnica e efetividade processual. Sistemas que se ajustam bem às demandas da tarefa tendem a suportar efetivamente processos sistemáticos de design instrucional.

5.8 Discussão

5.8.1 Aceitação Tecnológica (TAM) - Análise

Os resultados do TAM indicam uma aceitação substancial do sistema multiagente pelos participantes. A Utilidade Percebida apresentou a maior média ($M = 6.23$, $DP = 0.41$), sugerindo que os usuários reconhecem claramente os benefícios do sistema para o design de jogos educacionais (VENKATESH *et al.*, 2003). Esta média elevada indica que o sistema é percebido como uma ferramenta que efetivamente melhora a produtividade e qualidade do trabalho de design educacional.

A distribuição das respostas para Utilidade Percebida mostrou concentração nas faixas superiores da escala (6.0-6.9), com apenas 11% dos participantes avaliando abaixo de 6.0. Esta consistência sugere consenso sobre o valor agregado pelo sistema, independentemente do perfil demográfico dos usuários. Participantes com maior experiência (>10 anos) tenderam a avaliações ligeiramente superiores, possivelmente devido à maior capacidade de reconhecer eficiências proporcionadas pela automação.

A Facilidade de Uso apresentou uma média ligeiramente inferior ($M = 5.84$, $DP = 0.35$), indicando que, embora o sistema seja percebido como útil, existem oportunidades de melhoria na interface e experiência do usuário. A maior variabilidade neste constructo ($DP = 0.35$ vs $DP = 0.41$ para utilidade) sugere que a percepção de facilidade é mais influenciada por características individuais dos usuários.

Análise detalhada das respostas de Facilidade de Uso revelou que 23% dos participantes avaliaram abaixo de 6.0, concentrados principalmente entre usuários com menor experiência tecnológica. Este achado sugere necessidade de melhorias na curva de aprendizagem do sistema, particularmente para usuários menos experien-

tes com ferramentas de IA.

A forte correlação entre Utilidade Percebida e Intenção Comportamental ($r = 0.78$, $p < 0.001$) confirma o papel central da utilidade na aceitação de tecnologias educacionais (TEO, 2011). Interessantemente, a correlação entre Facilidade de Uso e Intenção Comportamental foi moderada ($r = 0.62$), sugerindo que usuários estão dispostos a investir esforço adicional de aprendizagem quando percebem alto valor na ferramenta.

5.8.2 Adequação Tarefa-Tecnologia (TTF) - Análise

A análise TTF revelou um alinhamento satisfatório entre as capacidades do sistema e os requisitos da tarefa de design educacional ($M = 5.95$, $DP = 0.12$). A baixa variabilidade ($DP = 0.12$) indica consenso notável entre participantes sobre a adequação do sistema, sugerindo que as funcionalidades foram bem calibradas para as necessidades específicas do design de jogos educacionais.

As Características da Tarefa obtiveram média elevada ($M = 6.05$), indicando que o sistema compreende adequadamente a complexidade e especificidades do design educacional. Participantes destacaram particularmente a capacidade do sistema de lidar com requisitos pedagógicos multifacetados, incluindo alinhamento curricular, diferenciação para estilos de aprendizagem, e integração de elementos lúdicos com objetivos educacionais.

O Impacto na Performance ($M = 5.89$, $DP = 0.22$) sugere que o sistema efetivamente contribui para melhorar a eficiência e qualidade do processo de design (GOODHUE e THOMPSON, 1995). Análise qualitativa complementar revelou que participantes percebem redução significativa no tempo de prototipagem (estimativa média de 40% de redução) e melhoria na consistência pedagógica dos jogos produzidos.

A correlação moderada a forte entre TTF e TAM ($r = 0.742$, $p < 0.001$) indica que a adequação percebida da tecnologia à tarefa influencia significativamente a aceitação do sistema. Este achado confirma a complementaridade entre perspectivas de adequação técnica (TTF) e aceitação comportamental (TAM), alinhando-se com estudos sobre integração destes modelos (DISHAW e STRONG, 1999).

Análise por subgrupos revelou que participantes com formação em design instrucional apresentaram avaliações TTF ligeiramente superiores, sugerindo que familiaridade com processos sistemáticos de design facilita o reconhecimento da adequação tecnológica.

5.8.3 Experiência Algorítmica (AX) - Análise

Os resultados do AX Framework destacam aspectos específicos da interação com IA que requerem atenção diferenciada. A Transparência Algorítmica apresentou a menor média entre todos os constructos avaliados ($M = 5.58$, $DP = 0.31$), indicando uma área crítica para melhorias futuras. Esta descoberta alinha-se com preocupações crescentes sobre explicabilidade em sistemas de IA educacionais (SHIN, 2021).

A distribuição das respostas para Transparência mostrou maior dispersão, com 31% dos participantes avaliando abaixo de 5.5. Análise qualitativa revelou que usuários desejam maior clareza sobre como os agentes processam informações pedagógicas e geram sugestões específicas. Participantes com formação técnica mostraram maior tolerância à complexidade algorítmica, enquanto educadores sem background tecnológico expressaram maior necessidade de explicações simplificadas.

A Confiança Algorítmica ($M = 5.81$, $DP = 0.22$) e Controlabilidade ($M = 5.78$, $DP = 0.29$) apresentaram níveis satisfatórios, sugerindo que os usuários sentem-se confortáveis com o nível de controle sobre as sugestões do sistema. A correlação positiva entre Controlabilidade e Confiança ($r = 0.67$) indica que maior controle percebido fortalece a confiança nas recomendações algorítmicas.

A Efetividade Algorítmica obteve média elevada ($M = 6.02$), sugerindo que, apesar das limitações em transparência, os usuários reconhecem a qualidade dos outputs produzidos. Este padrão indica um trade-off aceitável entre explicabilidade e performance, comum em sistemas de IA aplicada.

Análise temporal das avaliações AX (imediatamente após uso vs. após período de reflexão) mostrou ligeiro aumento na Confiança (+0.3 pontos) e Controlabilidade (+0.2 pontos), sugerindo que familiaridade crescente melhora a experiência algorítmica.

5.8.4 Avaliação do Processo (ADDIE) - Análise

A avaliação baseada no modelo ADDIE demonstrou efetividade consistente em todas as fases do processo de design instrucional. A Eficiência do Desenvolvimento obteve a maior média ($M = 5.92$, $DP = 0.21$), indicando que o sistema efetivamente acelera o processo de criação de jogos educacionais sem comprometer a qualidade (BRANCH, 2009).

A fase de Análise recebeu avaliação elevada ($M = 5.88$), com participantes destacando o suporte efetivo para mapeamento de necessidades de aprendizagem e definição de objetivos pedagógicos. O sistema demonstrou particular eficácia em sugerir alinhamentos curriculares e identificar pré-requisitos de aprendizagem relevantes.

A Qualidade do Design ($M = 5.90$) refletiu satisfação com o suporte à estruturação pedagógica e planejamento de atividades. Participantes valorizaram especi-

almente a capacidade do sistema de sugerir progressões de dificuldade adequadas e estratégias de engajamento apropriadas para diferentes faixas etárias.

O Suporte à Implementação obteve a menor média dentro do ADDIE ($M = 5.76$), indicando oportunidades de melhoria na transição do design para aplicação prática. Feedback qualitativo sugeriu necessidade de melhor suporte para deployment em diferentes plataformas e contextos tecnológicos.

A alta correlação entre fases do ADDIE (correlações entre 0.72 e 0.89) indica coerência processual, sugerindo que o sistema mantém qualidade consistente ao longo de todo o workflow de design instrucional.

5.9 Análise de Oportunidades de Melhoria e Estratégias de Aprimoramento

A análise dos resultados quantitativos revelou áreas específicas onde o sistema apresentou desempenho abaixo do ideal, oferecendo oportunidades concretas para melhorias direcionadas. Esta seção analisa as métricas com menor desempenho e propõe estratégias fundamentadas para seu aprimoramento.

5.9.1 Análise das Métricas com Menor Desempenho

Transparência Algorítmica: Desafios e Soluções

A Transparência Algorítmica apresentou a menor média entre todos os constructos avaliados ($M = 5.58$, $DP = 0.31$), indicando que 31% dos participantes avaliaram este aspecto abaixo de 5.5 na escala de 7 pontos.

Análise Qualitativa dos Problemas Identificados:

Durante as sessões de uso, participantes expressaram dificuldades específicas relacionadas à compreensão do funcionamento interno do sistema. Os usuários manifestaram interesse em entender como o sistema chegava às sugestões específicas, quais critérios pedagógicos estavam sendo considerados, e qual o raciocínio subjacente às recomendações apresentadas.

Estratégia Proposta: Tutorial Gamificado de Transparência

Com base na observação de que um tutorial gamificado poderia aumentar as métricas de transparência, propõe-se o desenvolvimento de um módulo interativo de explicabilidade denominado "Jornada do Agente Pedagógico". Este módulo permitiria aos usuários acompanhar o processo de análise passo-a-passo através de uma experiência gamificada.

Componentes do Tutorial Gamificado:

1. **Demonstração Interativa:** Apresentação visual do processo de tomada de decisão dos agentes, onde usuários podem "coletar" insights pedagógicos conforme exploram o raciocínio do sistema.
2. **Exercícios Práticos:** Cenários onde usuários acompanham o design de um jogo educacional e são desafiados a prever as próximas sugestões dos agentes, recebendo feedback sobre sua compreensão do processo.
3. **Sistema de Conquistas:** Reconhecimento progressivo da expertise adquirida, com níveis como "Explorador de IA" e "Especialista em Transparência".

Dashboard de Explicabilidade em Tempo Real:

Implementação de um painel que revela automaticamente o processo de tomada de decisão dos agentes durante o uso normal do sistema. Este painel mostraria:

- **Análise de Input:** Identificação e destaque dos conceitos-chave detectados nas ideias do usuário
- **Busca por Padrões:** Explicação de como o sistema utiliza conhecimento de jogos educacionais similares
- **Aplicação de Princípios:** Demonstração de como frameworks pedagógicos como a Taxonomia de Bloom informam as sugestões
- **Personalização:** Clarificação de como características do público-alvo influenciam as recomendações

Métricas de Sucesso Esperadas:

- Aumento de 1.2 pontos na média de Transparência Algorítmica
- Redução de 40% nas dúvidas sobre funcionamento do sistema
- Aumento de 25% na confiança nas sugestões dos agentes

Suporte à Implementação: Lacunas e Propostas de Melhoria

O Suporte à Implementação obteve a menor média dentro do modelo ADDIE ($M = 5.76$), indicando dificuldades na transição do design conceitual para aplicação prática. Esta limitação representa uma barreira significativa para a adoção real dos jogos educacionais desenvolvidos.

Análise das Limitações Identificadas:

O feedback dos participantes revelou uma lacuna crítica entre o design conceitual assistido pelo sistema e a capacidade prática de implementar os jogos educacionais propostos. Usuários relataram sentir-se perdidos sobre como transformar as ideias desenvolvidas em jogos funcionais, demonstrando necessidade de orientação sobre tecnologias, ferramentas de desenvolvimento e próximos passos práticos.

Estratégia de Melhoria 1: Módulo de Roadmap de Implementação

Desenvolvimento de um módulo que gera automaticamente um plano de implementação personalizado baseado no perfil do usuário, recursos disponíveis e complexidade do jogo projetado. Este módulo incluiria:

- **Análise de Viabilidade:** Avaliação das habilidades técnicas do usuário e recursos disponíveis
- **Fases de Desenvolvimento:** Estruturação do processo em etapas manejáveis com durações realistas
- **Recomendação de Ferramentas:** Sugestão de tecnologias apropriadas para cada fase e nível de expertise
- **Marcos Verificáveis:** Definição de objetivos intermediários para acompanhamento do progresso
- **Recursos de Apoio:** Indicação de tutoriais, comunidades e materiais de suporte específicos

Estratégia de Melhoria 2: Integração com Ferramentas de Prototipagem

Estabelecimento de conexões diretas entre o sistema e plataformas populares de desenvolvimento de jogos educacionais. Esta integração permitiria:

- **Exportação Estruturada:** Conversão automática dos elementos do EndoGDC para formatos compatíveis com ferramentas como Scratch, Unity, ou Construct
- **Templates Pré-configurados:** Geração de projetos iniciais baseados no design desenvolvido
- **Documentação Técnica:** Criação automática de especificações técnicas para orientar o desenvolvimento

Estratégia de Melhoria 3: Assistente de Implementação Passo-a-Passo

Desenvolvimento de um assistente virtual especializado que acompanha o usuário durante todo o processo de implementação. Este assistente ofereceria:

- **Avaliação Contínua:** Monitoramento do progresso e identificação de obstáculos
- **Suporte Contextual:** Orientação específica baseada no estágio atual de desenvolvimento
- **Conexão com Comunidade:** Facilitação de contato com mentores, parceiros técnicos ou colaboradores
- **Adaptação Dinâmica:** Ajuste do plano de implementação baseado no progresso real e feedback do usuário

Métricas de Sucesso Esperadas:

- Aumento de 1.5 pontos na média de Suporte à Implementação
- 60% dos usuários completam pelo menos um protótipo funcional
- Redução de 50% no tempo médio entre design e primeira implementação

5.9.2 Melhorias Adicionais Baseadas em Feedback Qualitativo

Aprimoramento da Facilidade de Uso ($M = 5.84$)

Embora tenha obtido pontuação satisfatória, a Facilidade de Uso apresenta potencial de melhoria, especialmente considerando que 23% dos participantes avaliaram abaixo de 6.0.

Estratégia de Melhoria: Onboarding Interativo Personalizado

Implementação de um processo de onboarding que familiariza usuários com a interface através de exercícios práticos adaptados ao seu nível de experiência. O sistema incluiria:

- **Avaliação Inicial:** Determinação do nível de familiaridade com design educacional e tecnologias digitais
- **Jornadas Personalizadas:** Três trilhas distintas (iniciante, intermediário, avançado) com duração e profundidade adequadas
- **Projeto Guiado:** Criação de um jogo educacional simples com assistência contextual progressivamente reduzida
- **Exploração Assistida:** Uso do sistema completo com suporte adaptativo disponível quando necessário

Características do Onboarding por Nível:

- **Usuários Novatos:** Foco em conceitos fundamentais, navegação básica e exemplos muito simples com suporte máximo
- **Usuários Intermediários:** Ênfase em funcionalidades avançadas com exemplos contextualizados e suporte moderado
- **Usuários Experientes:** Apresentação rápida de recursos específicos, atalhos e exemplos complexos com suporte mínimo

Fortalecimento da Confiança Algorítmica ($M = 5.81$)

A Confiança Algorítmica pode ser aprimorada através de maior transparência sobre a fonte e validação das sugestões apresentadas pelo sistema.

Estratégia de Melhoria: Sistema de Credibilidade das Sugestões

Implementação de indicadores de credibilidade que contextualizam a origem e fundamentação das sugestões oferecidas pelos agentes:

- **Fonte Pedagógica:** Identificação clara dos princípios educacionais que fundamentam cada sugestão
- **Evidência Empírica:** Referências a estudos ou práticas validadas que suportam as recomendações
- **Casos de Sucesso:** Exemplos de implementações similares bem-sucedidas

em contextos comparáveis

- **Nível de Consenso:** Indicação da convergência entre diferentes agentes especialistas sobre uma recomendação

Sistema de Feedback Bidirecional:

Implementação de mecanismos que permitem aos usuários:

- **Avaliar Sugestões:** Classificação da utilidade e relevância das recomendações recebidas
- **Reportar Problemas:** Identificação de sugestões inadequadas ou inconsistentes
- **Sugerir Melhorias:** Proposição de refinamentos baseados na experiência prática
- **Compartilhar Resultados:** Relato dos resultados obtidos com a implementação das sugestões

5.9.3 Estratégias de Implementação das Melhorias

Priorização Baseada em Impacto

A implementação das melhorias propostas deve seguir uma ordem de prioridade baseada no impacto potencial e viabilidade técnica:

Prioridade Alta (3-6 meses):

1. Tutorial gamificado de transparência
2. Dashboard de explicabilidade em tempo real
3. Onboarding interativo personalizado

Prioridade Média (6-12 meses):

1. Módulo de roadmap de implementação
2. Sistema de credibilidade das sugestões
3. Assistente de implementação passo-a-passo

Prioridade Baixa (12-18 meses):

1. Integração com ferramentas de prototipagem
2. Sistema de feedback bidirecional avançado
3. Plataforma de compartilhamento de implementações

Metodologia de Validação das Melhorias

Para cada melhoria implementada, propõe-se a aplicação de uma metodologia rigorosa de validação:

- **Testes A/B:** Comparação entre versões com e sem as melhorias implementadas
- **Estudos de Usabilidade:** Observação direta do uso do sistema por novos usuários

- **Análise de Métricas:** Monitoramento contínuo das métricas de aceitação e eficácia
- **Feedback Qualitativo:** Coleta sistemática de percepções e sugestões dos usuários

5.9.4 Contribuições para o Campo

As estratégias de melhoria propostas contribuem para o campo de sistemas educacionais baseados em IA através de:

1. **Framework de Explicabilidade Educacional:** Abordagem específica para transparência em sistemas de IA pedagógicos
2. **Metodologia de Suporte à Implementação:** Estratégias para reduzir a lacuna entre design e aplicação prática
3. **Métricas de Avaliação Refinadas:** Indicadores específicos para sistemas multiagentes educacionais
4. **Princípios de Onboarding Adaptativo:** Diretrizes para personalização de experiências de aprendizagem de sistemas

A implementação sistêmica dessas melhorias promete elevar significativamente a eficácia e aceitação do sistema, contribuindo para uma adoção mais ampla de tecnologias de IA no design de jogos educacionais. As estratégias propostas são transferíveis para outros sistemas educacionais baseados em LLMs, oferecendo valor para a comunidade de pesquisa mais ampla.

Capítulo 6

Conclusão

Este capítulo apresenta as conclusões da pesquisa, sintetizando as contribuições teóricas e práticas do sistema desenvolvido, limitações identificadas, implicações para o campo de jogos educacionais, e direcionamentos para trabalhos futuros na aplicação de sistemas multiagentes baseados em LLMs na educação.

Esta pesquisa apresentou o desenvolvimento e validação de um sistema multiagente baseado em LLMs para otimização do processo de design de jogos educacionais endógenos. Através da metodologia Design Science Research, foi construído um artefato tecnológico que integra quatro módulos funcionais especializados: Brainstorming, Agente Socrático, Endo-GDC e Taxonomia de Bloom. A validação empírica através de 26 estudos de caso em diferentes contextos educacionais demonstrou a eficácia da abordagem proposta. Os resultados quantitativos, avaliados através dos frameworks TAM, TTF, AX e ADDIE, indicaram alta aceitação tecnológica (utilidade percebida: $M=6.23$), adequação satisfatória entre tarefa e tecnologia (TTF: $M=5.95$), experiência algorítmica positiva (satisfação geral: $M=5.80$) e efetividade processual no design instrucional (satisfação geral do processo: $M=5.89$). Os estudos de caso revelaram a versatilidade do sistema em domínios diversos, desde ensino fundamental até formação corporativa, abrangendo áreas como matemática, ciências, idiomas, empreendedorismo e habilidades socioemocionais. A capacidade do sistema de gerar sugestões contextualmente relevantes e pedagogicamente fundamentadas foi consistentemente evidenciada através dos diferentes cenários de aplicação. A integração entre agentes especializados demonstrou ser uma estratégia eficaz para abordar a complexidade inerente ao design de jogos educacionais. O Agente Coordenador mostrou-se eficiente em orquestrar as contribuições dos agentes especialistas, mantendo coerência global sem comprometer a especialização individual. O Método Socrático implementado revelou-se valioso para promover reflexão crítica sobre as ideias geradas, enquanto o módulo de Taxonomia de Bloom assegurou alinhamento pedagógico apropriado. As limitações identificadas incluem dependência de APIs externas de LLM, necessidade de melhorias na transparência algorítmica e suporte

aprimorado para a fase de implementação dos jogos. Entretanto, essas limitações não comprometem a viabilidade da abordagem proposta nem suas contribuições fundamentais para o campo. Esta pesquisa, do ponto de vista teórico, estabelece uma ponte entre sistemas multiagentes e design educacional, propondo uma arquitetura replicável que pode ser adaptada para outros contextos de design instrucional. Do ponto de vista prático, oferece uma ferramenta concreta que pode reduzir significativamente o tempo e complexidade do desenvolvimento de jogos educacionais, democratizando o acesso a esta modalidade de recurso pedagógico.

As implicações para pesquisas futuras incluem a exploração de arquiteturas multiagentes mais sofisticadas, integração com ferramentas de prototipagem rápida, desenvolvimento de métricas específicas para avaliação de jogos educacionais endógenos, e investigação de estratégias para personalização adaptativa baseada em perfis individuais de aprendizagem.

A crescente integração de IA na educação demanda abordagens sistemáticas e teoricamente fundamentadas. Esta pesquisa demonstra que é possível aproveitar o potencial dos LLMs mantendo rigor pedagógico e científico, contribuindo para um futuro onde a tecnologia efetivamente amplifica a capacidade humana de criar experiências educacionais significativas e envolventes.

Referências Bibliográficas

- ABDUL JABBAR, A. I., FELICIA, P., 2015, “Gameplay engagement and learning in game-based learning: A systematic review”, *Review of educational research*, v. 85, n. 4, pp. 740–779.
- ACHIAM, J., ADLER, S., AGARWAL, S., et al., 2023, “GPT-4 technical report”, *arXiv preprint arXiv:2303.08774*.
- AHER, G. V., ARRIAGA, R. I., KALAI, A. T., 2023, “Using large language models to simulate multiple humans and replicate human subject studies”. In: *International Conference on Machine Learning*, pp. 337–371. PMLR.
- AHN, M., BROHAN, A., BROWN, N., et al., 2022, “Do as i can, not as i say: Grounding language in robotic affordances”, *arXiv preprint arXiv:2204.01691*.
- ALAYRAC, J.-B., DONAHUE, J., LUC, P., et al., 2022, “Flamingo: a visual language model for few-shot learning”, *Advances in neural information processing systems*, v. 35, pp. 23716–23736.
- ALL, A., CASTELLAR, E. P. N., VAN LOOY, J., 2014, “Measuring effectiveness in digital game-based learning: A methodological review.” *International Journal of Serious Games*, v. 1, n. 2.
- ANDERSON, L. W., 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.
- ARNAB, S., LIM, T., CARVALHO, M. B., et al., 2015, “Mapping learning and game mechanics for serious games analysis”, *British Journal of Educational Technology*, v. 46, n. 2, pp. 391–411.
- ASAI, A., WU, Z., WANG, Y., et al., 2023, “Self-rag: Learning to retrieve, generate, and critique through self-reflection”. In: *The Twelfth International Conference on Learning Representations*.

- BAHDANAU, D., CHO, K., BENGIO, Y., 2014, “Neural machine translation by jointly learning to align and translate”, *arXiv preprint arXiv:1409.0473*.
- BAI, Y., KADAVATH, S., KUNDU, S., et al., 2022, “Constitutional ai: Harmlessness from ai feedback”, *arXiv preprint arXiv:2212.08073*.
- BAKER, Z. R., AZHER, Z. L., 2024, “Simulating the us senate: An llm-driven agent approach to modeling legislative behavior and bipartisanship”, *arXiv preprint arXiv:2406.18702*.
- BANG, Y., CAHYAWIJAYA, S., LEE, N., et al., 2023, “A multitask, multilingual, multimodal evaluation of chatgpt on reasoning, hallucination, and interactivity”, *arXiv preprint arXiv:2302.04023*.
- BASU, C., SINGHAL, M., DRAGAN, A. D., 2018, “Learning from richer human guidance: Augmenting comparison-based learning with feature queries”. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 132–140.
- BENGIO, Y., DUCHARME, R., VINCENT, P., et al., 2003, “A neural probabilistic language model”, *Journal of machine learning research*, v. 3, n. Feb, pp. 1137–1155.
- BERTSCH, A., ALON, U., NEUBIG, G., et al., 2023, “Unlimiformer: Long-range transformers with unlimited length input”, *Advances in Neural Information Processing Systems*, v. 36, pp. 35522–35543.
- BESTA, M., BLACH, N., KUBICEK, A., et al., 2024, “Graph of thoughts: Solving elaborate problems with large language models”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, v. 38, pp. 17682–17690.
- BIGGS, J., TANG, C., KENNEDY, G., 2022, *Teaching for quality learning at university 5e*. McGraw-hill education (UK).
- BLACK, P., WILIAM, D., 2009, “Developing the theory of formative assessment”, *Educational Assessment, Evaluation and Accountability (formerly: Journal of personnel evaluation in education)*, v. 21, pp. 5–31.
- BLOOM, B. S., ENGELHART, M. D., FURST, E. J., et al., 1964, *Taxonomy of educational objectives*, v. 2. Longmans, Green New York.
- BOIKO, D. A., MACKNIGHT, R., GOMES, G., 2023, “Emergent autonomous scientific research capabilities of large language models”, *arXiv preprint arXiv:2304.05332*.

- BOLLER, S., KAPP, K., 2017, *Play to learn: Everything you need to know about designing effective learning games*. Association for talent development.
- BORSOS, Z., MARINIER, R., VINCENT, D., et al., 2023, “Audiolm: a language modeling approach to audio generation”, *IEEE/ACM transactions on audio, speech, and language processing*, v. 31, pp. 2523–2533.
- BOYLE, E. A., HAINEY, T., CONNOLLY, T. M., et al., 2016, “An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games”, *Computers & Education*, v. 94, pp. 178–192.
- BRAN, A. M., COX, S., SCHILTER, O., et al., 2023, “Chemcrow: Augmenting large-language models with chemistry tools”, *arXiv preprint arXiv:2304.05376*.
- BRANCH, R. M., 2009, *Instructional design: The ADDIE approach*, v. 722. Springer Science & Business Media.
- BROCK, A., LIM, T., RITCHIE, J. M., et al., 2017, “Freezeout: Accelerate training by progressively freezing layers”, *arXiv preprint arXiv:1706.04983*.
- BROWN, T., MANN, B., RYDER, N., et al., 2020, “Language models are few-shot learners”, *Advances in Neural Information Processing Systems*, v. 33, pp. 1877–1901.
- BUBECK, S., CHADRASEKARAN, V., ELDAN, R., et al., 2023. “Sparks of artificial general intelligence: Early experiments with gpt-4”. .
- BUSONI, L., BABUSKA, R., DE SCHUTTER, B., 2008, “A comprehensive survey of multiagent reinforcement learning”, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, v. 38, n. 2, pp. 156–172.
- CHASE, H., 2022. “Langchain: Building applications with llms through composability”. <https://github.com/hwchase17/langchain>.
- CHEN, B., SHU, C., SHAREGHI, E., et al., 2023a, “FireAct: Toward language agent fine-tuning”, *arXiv preprint arXiv:2310.05915*.
- CHEN, F., HAN, M., ZHAO, H., et al., 2023b, “X-LLM: Bootstrapping advanced large language models by treating multi-modalities as foreign languages”, *arXiv preprint arXiv:2305.04160*.

- CHEN, G., DONG, S., SHU, Y., et al., 2023c, “Autoagents: A framework for automatic agent generation”, *arXiv preprint arXiv:2309.17288*.
- CHEN, G., DONG, S., SHU, Y., et al., 2024a, “AutoAgents: A Framework for Automatic Agent Generation”. In: *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pp. 22–30, a. doi: 10.24963/ijcai.2024/3.
- CHEN, J., CHEN, Z., SU, J., et al., 2023d, “AgentSims: An Open-Source Sandbox for Large Language Model Evaluation”, *arXiv preprint arXiv:2308.04026*.
- CHEN, J., ZHONG, W., CUI, L., et al., 2023e, “Agent as Questioner: Exploring Proactive Information-Seeking in LLM Agents for Knowledge-Based QA”, *arXiv preprint arXiv:2311.16452*.
- CHEN, N., LI, H., HUANG, J., et al., 2024b, “Compress to impress: Unleashing the potential of compressive memory in real-world long-term conversations”, *arXiv preprint arXiv:2402.11975*.
- CHEN, S., LIU, Y., HAN, W., et al., 2025, “A survey on llm-based multi-agent system: Recent advances and new frontiers in application”, *Preprint at https://doi.org/10.48550/arXiv*, v. 2412.
- CHEN, W., SU, Y., ZUO, J., et al., 2024c, “AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors”. In: *International Conference on Learning Representations. ICLR*, c. Disponível em: <<https://arxiv.org/abs/2308.10848>>.
- CHEN, X., LIU, J., WANG, Y., et al., 2024d, “InfLLM: Unveiling the Intrinsic Capacity of LLMs for Understanding Extremely Long Sequences with Training-Free Memory”, *Proceedings of ICLR*, pp. 123–132.
- CHEN, Y., ARKIN, J., ZHANG, Y., et al., 2024e, “Scalable multi-robot collaboration with large language models: Centralized or decentralized systems?” In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4311–4317. IEEE, e.
- CHENG, X., LUO, D., CHEN, X., et al., 2024a, “Lift yourself up: Retrieval-augmented text generation with self-memory”, *Advances in Neural Information Processing Systems*, v. 36.
- CHENG, Y., ZHANG, C., ZHANG, Z., et al., 2024b, “Exploring large language model based intelligent agents: Definitions, methods, and prospects”, *arXiv preprint arXiv:2401.03428*.

- CHOWDHERY, A., NARANG, S., DEVLIN, J., et al., 2023, “PaLM: Scaling language modeling with pathways”, *Journal of Machine Learning Research*, v. 24, n. 240, pp. 1–113.
- CHRISTIANO, P. F., LEIKE, J., BROWN, T., et al., 2017, “Deep reinforcement learning from human preferences”, *Advances in neural information processing systems*, v. 30.
- CHUANG, Y.-S., XIE, Y., LUO, H., et al., 2023, “DOLA: Decoding by contrasting layers improves factuality in large language models”. In: *The Twelfth International Conference on Learning Representations*.
- CHURCHES, A., 2008, “Bloom’s Digital Taxonomy”, *Tech & Learning*.
- CLARK, D. B., TANNER-SMITH, E. E., KILLINGSWORTH, S. S., 2016, “Digital games, design, and learning: A systematic review and meta-analysis”, *Review of Educational Research*, v. 86, n. 1, pp. 79–122.
- CLARK, R. E., MAYER, R. E., 2016, *Learning by playing: Video games in education*. Oxford University Press.
- COBBE, K., KOSARAJU, V., BAVARIAN, M., et al., 2021, “Training verifiers to solve math word problems”, *arXiv preprint arXiv:2110.14168*.
- COHEN, L., MANION, L., MORRISON, K., 2017, *Research Methods in Education*. 8 ed. London, Routledge.
- COHEN, P. R., LEVESQUE, H. J., 1990, “Intention is Choice with Commitment”, *Artificial Intelligence*, v. 42, n. 2-3, pp. 213–261.
- CORNIA, M., STEFANINI, M., BARALDI, L., et al., 2020, “Meshed-memory transformer for image captioning”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10578–10587.
- COSTA, L. F. C., GOMES, S., SANTOS, A. M., et al., 2024, “Heroine’s Learning Journey: Motivating Women in STEM Online Courses Through the Power of a Narrative”, *IEEE Access*, v. 12, pp. 20103–20124.
- DAI, W., LI, J., LI, D., et al., 2024, “InstructBLIP: Towards general-purpose vision-language models with instruction tuning”, *Advances in Neural Information Processing Systems*, v. 36.
- DARLING-HAMMOND, L., HYLER, M. E., GARDNER, M., 2017, “Effective Teacher Professional Development”, *Learning Policy Institute*.

- DASGUPTA, I., KAESER-CHEN, C., MARINO, K., et al., 2023, “Collaborating with language models for embodied reasoning”, *arXiv preprint arXiv:2302.00763*.
- DAVIS, F. D., 1989, “Perceived usefulness, perceived ease of use, and user acceptance of information technology”, *MIS quarterly*, pp. 319–340.
- DECKER, K., LESSER, V., 1993a, “Quantitative modeling of complex computational task environments”. In: *AAAI*, pp. 217–224. Citeseer, a.
- DECKER, K. S., 1987, “Distributed problem-solving techniques: A survey”, *IEEE Transactions on systems, Man, and Cybernetics*, v. 17, n. 5, pp. 729–740.
- DECKER, K. S., LESSER, V. R., 1993b, “Quantitative Modeling of Complex Environments”, *International Journal of Intelligent Systems in Accounting, Finance and Management*, v. 2, pp. 215–234.
- DETTMERS, T., PAGNONI, A., HOLTZMAN, A., et al., 2024, “QLoRA: Efficient finetuning of quantized LLMs”, *Advances in Neural Information Processing Systems*, v. 36.
- DHULIAWALA, S., KOMEILI, M., XU, J., et al., 2024, “Chain-of-verification reduces hallucination in large language models”. In: *Findings of the Association for Computational Linguistics ACL 2024*, pp. 3563–3578.
- DING, Y., ZHANG, X., AMIRI, S., et al., 2023, “Integrating action knowledge and LLMs for task planning and situation handling in open worlds”, *Autonomous Robots*, v. 47, n. 8, pp. 981–997.
- DISHAW, M. T., STRONG, D. M., 1999, “Extending the technology acceptance model with task–technology fit constructs”, *Information & management*, v. 36, n. 1, pp. 9–21.
- DOBBINS, M., 2017, “Rapid review guidebook”, *Natl Collab Cent Method Tools*, v. 13, pp. 25.
- DONG, Q., LI, L., DAI, D., et al., 2022, “A survey on in-context learning”, *arXiv preprint arXiv:2301.00234*.
- DONG, Y., JIANG, X., JIN, Z., et al., 2024a, “Self-collaboration code generation via chatgpt”, *ACM Transactions on Software Engineering and Methodology*, v. 33, n. 7, pp. 1–38.

- DONG, Y., ZHU, X., PAN, Z., et al., 2024b, “Villageragent: A graph-based multi-agent framework for coordinating complex task dependencies in mine-craft”, *arXiv preprint arXiv:2406.05720*.
- DOSOVITSKIY, A., BEYER, L., KOLESNIKOV, A., et al., 2021, “An image is worth 16x16 words: Transformers for image recognition at scale”. In: *International Conference on Learning Representations*.
- DRESCH, A., LACERDA, D. P., JUNIOR, J. A. V. A., 2020, *Design science research: método de pesquisa para avanço da ciência e tecnologia*. Bookman Editora.
- DRIESS, D., XIA, F., SAJJADI, M. S., et al., 2023, “Palm-e: An embodied multimodal language model”. In: *Proceedings of the 40th International Conference on Machine Learning*, v. 340, *ICML’23*, p. 20. JMLR.org.
- DU, M., HE, F., ZOU, N., et al., 2023a, “Shortcut learning of large language models in natural language understanding”, *Communications of the ACM*, v. 67, n. 1, pp. 110–120.
- DU, X., JI, H., 2022, “Retrieval-augmented generative question answering for event argument extraction”. In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 4649–4666.
- DU, Y., LI, S., TORRALBA, A., et al., 2023b, “Improving factuality and reasoning in language models through multiagent debate”. In: *Forty-first International Conference on Machine Learning*, b.
- DURFEE, E. H., LESSER, V. R., 1991, “Partial Global Planning: A Coordination Framework for Distributed Hypothesis Formation”, *IEEE Transactions on Systems, Man, and Cybernetics*, v. 21, n. 5, pp. 1167–1183.
- ECK, R. V., 2006, “Digital Game-Based Learning: It’s Not Just the Digital Natives Who Are Restless”, *EDUCAUSE Review*, v. 41, n. 2, pp. 16–30. Disponível em: <<https://er.educause.edu/articles/2006/1/digital-gamebased-learning-its-not-just-the-digital-natives-who-are-restless>>
- ERICSSON, K. A., CHARNESS, N., FELTOVICH, P. J., et al., 2006, *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge, Cambridge University Press.
- ETHAYARAJH, K., 2019, “How contextual are contextualized word representations? Comparing the geometry of BERT, ELMo, and GPT-2 embeddings”, *arXiv preprint arXiv:1909.00512*.

- FINK, L. D., 2013, *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. San Francisco, Jossey-Bass.
- FIRAT, M., KULELI, S., 2023, “What if gpt4 became autonomous: The auto-gpt project and use cases”, *Journal of Emerging Computer Technologies*, v. 3, n. 1, pp. 1–6.
- FISCHER, K., 2023, “Reflective linguistic programming (rlp): A stepping stone in socially-aware agi (socialagi)”, *arXiv preprint arXiv:2305.12647*.
- FRANKLIN, S., GRAESSER, A., 1996, “Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents”. In: *International workshop on agent theories, architectures, and languages*, pp. 21–35. Springer.
- GALLOTTA, R., TODD, G., ZAMMIT, M., et al., 2024, “Large Language Models and Games: A Survey and Roadmap”, *arXiv preprint arXiv:2402.18659*.
- GALOTTI, K. M., 1989, “Approaches to studying formal and everyday reasoning”, *Psychological Bulletin*, v. 105, n. 3, pp. 331.
- GAO, C., LAN, X., LU, Z., et al., 2023a, “S3: Social-network simulation system with large language model-empowered agents”, *arXiv preprint arXiv:2307.14984*.
- GAO, H., ZHANG, Y., 2024, “Memory sharing for large language model based agents”, *arXiv preprint arXiv:2404.09982*.
- GAO, L., MADAAN, A., ZHOU, S., et al., 2023b, “PAL: Program-aided language models”. In: *International Conference on Machine Learning*, pp. 10764–10799, b.
- GAO, Y., XIONG, Y., GAO, X., et al., 2023c, “Retrieval-augmented generation for large language models: A survey”, *arXiv preprint arXiv:2312.10997*.
- GEE, J. P., 2003, *What Video Games Have to Teach Us About Learning and Literacy*. New York, Palgrave Macmillan.
- GEE, J. P., 2007, *Good Video Games and Good Learning: Collected Essays on Video Games, Learning, and Literacy*. New York, Palgrave Macmillan.
- GEIRHOS, R., JACOBSEN, J.-H., MICHAELIS, C., et al., 2020, “Shortcut learning in deep neural networks”, *Nature Machine Intelligence*, v. 2, n. 11, pp. 665–673.

- GHAFAROLLAHI, A., BUEHLER, M. J., 2024, “ProtAgents: protein discovery via large language model multi-agent collaborations combining physics and machine learning”, *Digital Discovery*, v. 3, n. 7, pp. 1389–1409.
- GMYTRASIEWICZ, P. J., DOSHI, P., 2005, “A Framework for Sequential Planning in Multi-Agent Settings”, *Journal of Artificial Intelligence Research*, v. 24, pp. 49–79.
- GOODHUE, D. L., THOMPSON, R. L., 1995, “Task-technology fit and individual performance”, *MIS quarterly*, pp. 213–236.
- GUO, Z., SCHLICHTKRULL, M., VLACHOS, A., 2022, “A survey on automated fact-checking”, *Transactions of the Association for Computational Linguistics*, v. 10, pp. 178–206.
- HAN, S., ZHANG, Q., YAO, Y., et al., 2024, “LLM multi-agent systems: Challenges and open problems”, *arXiv preprint arXiv:2402.15538*.
- HAO, S., GU, Y., MA, H., et al., 2023, “Reasoning with language model is planning with world model”. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 8154–8173.
- HATTIE, J., 2008, *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London, Routledge.
- HE, J., TREUDE, C., LO, D., 2024a, “LLM-based multi-agent systems for software engineering: Vision and the road ahead”, *arXiv e-prints*, pp. arXiv–2404.
- HE, J., ZHOU, C., MA, X., et al., 2021, “Towards a unified view of parameter-efficient transfer learning”, *arXiv preprint arXiv:2110.04366*.
- HE, X., TIAN, Y., SUN, Y., et al., 2024b, “G-Retriever: Retrieval-augmented generation for textual graph understanding and question answering”, *arXiv preprint arXiv:2402.07630*.
- HOLMES, W., BIALIK, M., FADEL, C., 2019, “Artificial intelligence in education promises and implications for teaching and learning”, *Center for Curriculum Redesign*.
- HONG, S., ZHENG, X., CHEN, J., et al., 2023, “Metagpt: Meta programming for multi-agent collaborative framework”, *arXiv preprint arXiv:2308.00352*, v. 3, n. 4, pp. 6.

- HORNIK, K., STINCHCOMBE, M., WHITE, H., 1989, “Multilayer feedforward networks are universal approximators”, *Neural networks*, v. 2, n. 5, pp. 359–366.
- HOULSBY, N., GIURGIU, A., JASTRZEBSKI, S., et al., 2019, “Parameter-efficient transfer learning for NLP”. In: *International conference on machine learning*, pp. 2790–2799.
- HSU, W.-N., BOLTE, B., TSAI, Y.-H. H., et al., 2021, “HuBERT: Self-supervised speech representation learning by masked prediction of hidden units”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 29, pp. 3451–3460.
- HU, C., FU, J., DU, C., et al., 2023a, “ChatDB: Augmenting LLMs with databases as their symbolic memory”, *arXiv preprint arXiv:2306.03901*.
- HU, E. J., SHEN, Y., WALLIS, P., et al., 2021, “LoRA: Low-rank adaptation of large language models”, *arXiv preprint arXiv:2106.09685*.
- HU, Z., WANG, L., LAN, Y., et al., 2023b, “LLM-Adapters: An adapter family for parameter-efficient fine-tuning of large language models”, *arXiv preprint arXiv:2304.01933*.
- HUANG, J., CHANG, K. C.-C., 2023, “Towards reasoning in large language models: A survey”. In: *Findings of the Association for Computational Linguistics: ACL 2023*, pp. 1049–1065.
- HUANG, R., LI, M., YANG, D., et al., 2024, “AudioGPT: Understanding and generating speech, music, sound, and talking head”, *Proceedings of the AAAI Conference on Artificial Intelligence*, v. 38, pp. 23802–23804.
- HUANG, S., MAMIDANNA, S., JANGAM, S., et al., 2023a, “Can large language models explain themselves? A study of LLM-generated self-explanations”, *arXiv preprint arXiv:2310.11207*.
- HUANG, W., XIA, F., XIAO, T., et al., 2022, “Inner monologue: Embodied reasoning through planning with language models”, *arXiv preprint arXiv:2207.05608*.
- HUANG, Z., GUTIERREZ, S., KAMANA, H., et al., 2023b, “Memory sandbox: Transparent and interactive memory management for conversational agents”. In: *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–3, b.

- HUBEL, D. H., WIESEL, T. N., 1962, “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex”, *The Journal of physiology*, v. 160, n. 1, pp. 106.
- JI, Z., LEE, N., FRIESKE, R., et al., 2023, “Survey of hallucination in natural language generation”, *ACM Computing Surveys*, v. 55, n. 12, pp. 1–38.
- JIANG, Z., XU, F. F., GAO, L., et al., 2023, “Active retrieval augmented generation”. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 7969–7992.
- JIN, H., ZHANG, C., BALUNOVIC, M., et al., 2023, “GPTSwarm: How Many Agents Can We Fit into a Language Model?” *arXiv preprint arXiv:2401.03325*.
- JINXIN, S., JIABAO, Z., YILEI, W., et al., 2023, “Cgmi: Configurable general multi-agent interaction framework”, *arXiv preprint arXiv:2308.12503*.
- KAGAYA, T., YUAN, T. J., LOU, Y., et al., 2024, “Rap: Retrieval-augmented planning with contextual memory for multimodal llm agents”, *arXiv preprint arXiv:2402.03610*.
- KALMPOURTZIS, G., 2018, *Educational Game Design Fundamentals: A Journey to Creating Intrinsically Motivating Learning Experiences*. Boca Raton, CRC Press.
- KANDLBINDER, P., 2014, “Constructive alignment in university teaching”, *HERDSA News*, v. 36, n. 3, pp. 5–6.
- KAPETANAKIS, S., KUDENKO, D., 2003, “Reinforcement learning of coordination in heterogeneous cooperative multi-agent systems”. In: *Symposium on Adaptive Agents and Multi-agent Systems*, pp. 119–131. Springer.
- KAPLAN, J., MCCANDLISH, S., HENIGHAN, T., et al., 2020, “Scaling laws for neural language models”, *arXiv preprint arXiv:2001.08361*.
- KARPAS, E., ABEND, O., BELINKOV, Y., et al., 2022, “MRKL systems: A modular, neuro-symbolic architecture that combines large language models, external knowledge sources and discrete reasoning”, *arXiv preprint arXiv:2205.00445*.
- KECSKÉS, I., KECSKÉS, I., HORN, L. R., 2007, *Explorations in pragmatics: Linguistic, cognitive and intercultural aspects*. Mouton de Gruyter, New York. doi: 10.1515/9783110198843.

- KIRKPATRICK, J., PASCANU, R., RABINOWITZ, N., et al., 2017, “Overcoming catastrophic forgetting in neural networks”, *Proceedings of the National Academy of Sciences*, v. 114, n. 13, pp. 3521–3526.
- KOJIMA, T., GU, S. S., REID, M., et al., 2022, “Large language models are zero-shot reasoners”, *Advances in Neural Information Processing Systems*, v. 35, pp. 22199–22213.
- KRATHWOHL, D. R., 2002, “A Revision of Bloom’s Taxonomy: An Overview”, *Theory into Practice*, v. 41, n. 4, pp. 212–218.
- KRAUS, S., 1997, “Negotiation and Cooperation in Multi-Agent Environments”, *Artificial Intelligence*, v. 94, n. 1-2, pp. 79–97.
- LAMERAS, P., ARNAB, S., DUNWELL, I., et al., 2017, “Essential features of serious games design in higher education”, *British Journal of Educational Technology*, v. 48, n. 4, pp. 972–994.
- LESSER, V., CORKILL, D., 2014, “Challenges for Multi-Agent Coordination Theory Based on Empirical Observations”, *Autonomous Agents and Multi-Agent Systems*, v. 28, n. 1, pp. 1–35.
- LESTER, B., AL-RFOU, R., CONSTANT, N., 2021, “The power of scale for parameter-efficient prompt tuning”, *arXiv preprint arXiv:2104.08691*.
- LEWIS, P., PEREZ, E., PIKTUS, A., et al., 2020, “Retrieval-augmented generation for knowledge-intensive nlp tasks”, *Advances in Neural Information Processing Systems*, v. 33, pp. 9459–9474.
- LI, C., SU, X., HAN, H., et al., 2023a, “Quantifying the impact of large language models on collective opinion dynamics”, *arXiv preprint arXiv:2308.03313*.
- LI, H., CHONG, Y. Q., STEPPUTTIS, S., et al., 2023b, “Theory of mind for multi-agent collaboration via large language models”, *arXiv preprint arXiv:2310.10701*.
- LI, J., LI, D., SAVARESE, S., et al., 2023c, “BLIP-2: Bootstrapping language-image pre-training with frozen image encoders and large language models”. In: *International conference on machine learning*, pp. 19730–19742, c.
- LI, K., HE, Y., WANG, Y., et al., 2023d, “VideoChat: Chat-centric video understanding”, *arXiv preprint arXiv:2305.06355*.
- LI, N., GAO, C., LI, Y., et al., 2024a. “Large Language Model-Empowered Agents for Simulating Macroeconomic Activities (2023)”. a.

- LI, R., ZHANG, C., MAO, S., et al., 2023e, “From English to PCSEL: LLM helps design and optimize photonic crystal surface emitting lasers”, .
- LI, X. L., LIANG, P., 2021, “Prefix-tuning: Optimizing continuous prompts for generation”, *arXiv preprint arXiv:2101.00190*.
- LI, X., LI, S., SONG, S., et al., 2024b, “PMET: Precise model editing in a transformer”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, v. 38, pp. 18564–18572, b.
- LI, X., LIU, Z., XIONG, C., et al., 2023f, “Structure-aware language model pre-training improves dense retrieval on structured data”. In: *Findings of the Association for Computational Linguistics: ACL 2023*, pp. 11560–11574, f.
- LI, Y., YU, Y., LI, H., et al., 2023g, “Tradinggpt: Multi-agent system with layered memory and distinct characters for enhanced financial trading performance”, *arXiv preprint arXiv:2309.03736*.
- LI, Y., HU, X., WANG, Z., et al., 2023h, “SPRING: GPT-4 Out-performs RL Algorithms by Studying Papers and Reasoning”, *arXiv preprint arXiv:2305.13307*.
- LIANG, C., WU, C., SONG, T., et al., 2023, “Taskmatrix. ai: Completing tasks by connecting foundation models with millions of apis”, *arXiv preprint arXiv:2303.16434*.
- LIGHT, J., CAI, M., SHEN, S., et al., 2023, “Avalonbench: Evaluating llms playing the game of avalon”, *arXiv preprint arXiv:2310.05036*.
- LIN, J., FRIED, D., KLEIN, D., et al., 2022a, “Inferring rewards from language in context”. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 8546–8560, a.
- LIN, J., ZHAO, H., ZHANG, A., et al., 2023, “AgentSims: An Open-Source Sandbox for Large Language Model Evaluation”, *arXiv preprint arXiv:2308.04026*. Disponível em: <<https://arxiv.org/abs/2308.04026>>.
- LIN, S., HILTON, J., EVANS, O., 2022b, “TruthfulQA: Measuring how models mimic human falsehoods”. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 3214–3252, b.

- LIU, B., JIANG, Y., ZHANG, X., et al., 2023a, “LLM+ P: Empowering large language models with optimal planning proficiency”, *arXiv preprint arXiv:2304.11477*.
- LIU, H., LI, C., WU, Q., et al., 2024a, “Visual instruction tuning”, *Advances in Neural Information Processing Systems*, v. 36.
- LIU, S., HUANG, X., LIU, Y., 2017, “Artificial intelligence in educational games: A systematic review”, *Journal of Educational Computing Research*, v. 55, n. 6, pp. 809–838.
- LIU, W., WANG, C., WANG, Y., et al., 2024b, “Autonomous agents for collaborative task under information asymmetry”, *arXiv preprint arXiv:2406.14928*.
- LIU, X., ZHENG, Y., DU, Z., et al., 2021a, “GPT understands, too”, *arXiv preprint arXiv:2103.10385*.
- LIU, X., ZHU, Z., LIU, H., et al., 2023b, “WavJourney: Compositional audio creation with large language models”, *arXiv preprint arXiv:2307.14335*.
- LIU, Y., AGARWAL, S., VENKATARAMAN, S., 2021b, “AutoFreeze: Automatically freezing model blocks to accelerate fine-tuning”, *arXiv preprint arXiv:2102.01386*.
- LIU, Z., ZHANG, Y., LI, P., et al., 2023c, “Dynamic LLM-Agent Network: An LLM-agent Collaboration Framework with Agent Team Optimization”, *arXiv preprint arXiv:2310.02170*. Disponível em: <<https://arxiv.org/abs/2310.02170>>.
- LOGOTHETIS, N. K., SHEINBERG, D. L., 1996, “Visual object recognition”, *Annual review of neuroscience*, v. 19, n. 1, pp. 577–621.
- LU, Y., BARTOLO, M., MOORE, A., et al., 2022, “Fantastically ordered prompts and where to find them: Overcoming few-shot prompt order sensitivity”. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 8086–8098.
- LUO, Z., XU, C., ZHAO, P., et al., 2023, “Augmented large language models with parametric knowledge guiding”, *arXiv preprint arXiv:2305.04757*.
- LYU, C., WU, M., WANG, L., et al., 2023a, “Macaw-LLM: Multi-modal language modeling with image, audio, video, and text integration”, *arXiv preprint arXiv:2306.09093*.

- LYU, Q., HAVALDAR, S., STEIN, A., et al., 2023b, “Faithful chain-of-thought reasoning”. In: *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 305–329, b.
- MADAAN, A., TANDON, N., CLARK, P., et al., 2022a, “Memory-assisted prompt editing to improve GPT-3 after deployment”, *arXiv preprint arXiv:2201.06009*.
- MADAAN, A., ZHOU, S., ALON, U., et al., 2022b, “Language models of code are few-shot commonsense learners”. In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 1384–1403, b.
- MADAAN, A., TANDON, N., GUPTA, P., et al., 2024, “Self-refine: Iterative refinement with self-feedback”, *Advances in Neural Information Processing Systems*, v. 36.
- MADASU, A., FIRDAUS, M., EKBAL, A., 2023, “A unified framework for emotion identification and generation in dialogues”. In: *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop*, pp. 73–78.
- MAHARANA, A., LEE, D.-H., TULYAKOV, S., et al., 2024, “Evaluating very long-term conversational memory of llm agents”, *arXiv preprint arXiv:2402.17753*.
- MAMUN, M. A., ABDULLAH, H. M., ALAM, M. G. R., et al., 2023, “Affective social anthropomorphic intelligent system”, *Multimedia Tools and Applications*, v. 82, n. 23, pp. 35059–35090.
- MANAKUL, P., GALES, M. J., 2021, “Sparsity and sentence structure in encoder-decoder attention of summarization systems”. In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 9359–9368.
- MANDI, Z., JAIN, S., SONG, S., 2024, “Roco: Dialectic multi-robot collaboration with large language models”. In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 286–299. IEEE.
- MARSHALL, L. H., MAGOUN, H. W., 2013, *Discoveries in the human brain: neuroscience prehistory, brain structure, and function*. Springer Science & Business Media, Humana Totowa, NJ.

- MARZANO, R. J., KENDALL, J. S., 2006, *The New Taxonomy of Educational Objectives*. Thousand Oaks, CA, Corwin Press.
- MAYER, R. E., 2014, “Computer games for learning: An evidence-based approach”, *MIT Press*.
- MAYER, R. E., PARONG, J., BAINBRIDGE, K., 2019, “Game-based learning: What we know and what we need to know”, *Educational Psychology Review*, v. 31, n. 3, pp. 689–720.
- MAYNEZ, J., NARAYAN, S., BOHNET, B., et al., 2020, “On faithfulness and factuality in abstractive summarization”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 1906–1919.
- MCCULLOCH, W. S., PITTS, W., 1943, “A logical calculus of the ideas immanent in nervous activity”, *The bulletin of mathematical biophysics*, v. 5, pp. 115–133.
- MEHTA, N., TERUEL, M., SANZ, P. F., et al., 2023, “Improving grounded language understanding in a collaborative environment by interacting with agents through help feedback”, *arXiv preprint arXiv:2304.10750*.
- MENDONCCA, J., LAVIE, A., TRANCOSO, I., 2025, “MEDAL: A Framework for Benchmarking LLMs as Multilingual Open-Domain Chatbots and Dialogue Evaluators”. Disponível em: <<https://api.semanticscholar.org/CorpusID:278996926>>.
- MERRILL, M. D., 2002, “First Principles of Instruction”, *Educational Technology Research and Development*, v. 50, n. 3, pp. 43–59.
- MICHAEL, D., CHEN, S., 2005, *Serious Games: Games That Educate, Train, and Inform*. Boston, Course Technology PTR.
- MITCHELL, E., LIN, C., BOSSELUT, A., et al., 2022, “Memory-based model editing at scale”. In: *International Conference on Machine Learning*, pp. 15817–15831.
- MUKOBI, G., ERLEBACH, H., LAUFFER, N., et al., 2023, “Welfare diplomacy: Benchmarking language model cooperation”, *arXiv preprint arXiv:2310.08901*.
- NAKANO, R., HILTON, J., BALAJI, S., et al., 2021, “WebGPT: Browser-assisted question-answering with human feedback”, *arXiv preprint arXiv:2112.09332*.

- NASCIMENTO, N., ALENCAR, P., COWAN, D., 2023, “Self-adaptive large language model (llm)-based multiagent systems”. In: *2023 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C)*, pp. 104–109. IEEE.
- NIE, Y., HUANG, H., WEI, W., et al., 2022, “Capturing global structural information in long document question answering with compressive graph selector network”. In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 5036–5047.
- NOTTINGHAM, K., AMMANABROLU, P., SUHR, A., et al., 2023, “Do embodied agents dream of pixelated sheep: Embodied decision making using language guided world modelling”. In: *International Conference on Machine Learning*, pp. 26311–26325.
- NOËL, F., XEXÉO, G., MANGELI, E., et al., 2021, “Screener, an educational game for teaching the drug discovery and development process”, *Brazilian Journal of Medical and Biological Research*, v. 54, n. 12, pp. e11786–. doi: <https://doi.org/10.1590/1414-431X2021e11786>. Disponível em: <<https://www.scielo.br/j/bjmbbr/a/V9zvHzdwjfgMtrFyMKBmNth/?lang=en>>.
- OLFATI-SABER, R., FAX, J. A., MURRAY, R. M., 2007, “Consensus and Cooperation in Networked Multi-Agent Systems”, *Proceedings of the IEEE*, v. 95, n. 1, pp. 215–233.
- OUYANG, L., WU, J., JIANG, X., et al., 2022, “Training language models to follow instructions with human feedback”, *Advances in Neural Information Processing Systems*, v. 35, pp. 27730–27744.
- PACKER, C., FANG, V., PATIL, S. G., et al., 2023, “MemGPT: Towards LLMs as operating systems”, *arXiv preprint arXiv:2310.08560*.
- PAGE, M. J., MCKENZIE, J. E., BOSSUYT, P. M., et al., 2021, “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews”, *bmj*, v. 372.
- PALLAGANI, V., MUPPASANI, B. C., ROY, K., et al., 2024, “On the prospects of incorporating large language models (llms) in automated planning and scheduling (aps)”. In: *Proceedings of the International Conference on Automated Planning and Scheduling*, v. 34, pp. 432–444.
- PANAIT, L., LUKE, S., 2005, “Cooperative multi-agent learning: The state of the art”, *Autonomous agents and multi-agent systems*, v. 11, pp. 387–434.

- PARISI, A., ZHAO, Y., FIEDEL, N., 2022, “TaLM: Tool augmented language models”, *arXiv preprint arXiv:2205.12255*.
- PARK, J. S., POPOWSKI, L., CAI, C., et al., 2022, “Social simulacra: Creating populated prototypes for social computing systems”. In: *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–18.
- PARK, J. S., O’BRIEN, J., CAI, C. J., et al., 2023, “Generative agents: Interactive simulacra of human behavior”. In: *Proceedings of the 36th annual acm symposium on user interface software and technology*, pp. 1–22.
- PARUNAK, H. V. D., 1996, “Applications of distributed artificial intelligence in industry”, *Foundations of distributed artificial intelligence*, v. 2, n. 1, pp. 18.
- PATIL, S. G., ZHANG, T., WANG, X., et al., 2023, “Gorilla: Large language model connected with massive APIs”, *arXiv preprint arXiv:2305.15334*.
- PELLEGRINO, J. W., CHUDOWSKY, N., GLASER, R., 2001, *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC, National Academy Press.
- PENG, B., GALLEY, M., HE, P., et al., 2023, “Check your facts and try again: Improving large language models with external knowledge and automated feedback”, *arXiv preprint arXiv:2302.12813*.
- PETRONI, F., ROCKTÄSCHEL, T., LEWIS, P., et al., 2019, “Language models as knowledge bases?” *arXiv preprint arXiv:1909.01066*.
- PFEIFFER, J., VULIĆ, I., GUREVYCH, I., et al., 2020, “MAD-X: An adapter-based framework for multi-task cross-lingual transfer”, *arXiv preprint arXiv:2005.00052*.
- PIMENTEL, M., FILIPPO, D., SANTORO, F. M., 2020a, “Design Science Research: fazendo pesquisas científicas rigorosas atreladas ao desenvolvimento de artefatos computacionais projetados para a educação”. In: Jaques, P. A., Pimentel, M., Siqueira, S., et al. (Eds.), *Metodologia de Pesquisa Científica em Informática na Educação: Conceção de Pesquisa*, v. 1, *Série Metodologia de Pesquisa em Informática na Educação*, SBC, Porto Alegre, a. Disponível em: <https://ceie.sbc.org.br/metodologia/livro-1/>.

- PIMENTEL, M., FILIPPO, D., SANTOS, T. M. D., 2020b, “Design Science Research: pesquisa científica atrelada ao design de artefatos”, *RE@D - Revista de Educação a Distância e Elearning*, v. 3, n. 1 (3), pp. 39–61.
- PINTRICH, P. R., 2002, “The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing”, *Theory into Practice*, v. 41, n. 4, pp. 219–225.
- PLASS, J. L., HOMER, B. D., KINZER, C. K., 2015, “Foundations of Game-Based Learning”, *Educational Psychologist*, v. 50, n. 4, pp. 217–243.
- PRENSKY, M., 2001, *Digital Game-Based Learning*. New York, McGraw-Hill.
- QIAN, C., LIU, W., LIU, H., et al., 2023, “Chatdev: Communicative agents for software development”, *arXiv preprint arXiv:2307.07924*.
- QIAN, M., CLARK, K. R., 2016, “Game-based learning and 21st century skills: A review of recent research”, *Computers in Human Behavior*, v. 63, pp. 50–58.
- RADFORD, A., WU, J., CHILD, R., et al., 2019, “Language models are unsupervised multitask learners”, *OpenAI Blog*, v. 1, n. 8, pp. 9.
- RAFAILOV, R., SHARMA, A., MITCHELL, E., et al., 2024, “Direct preference optimization: Your language model is secretly a reward model”, *Advances in Neural Information Processing Systems*, v. 36.
- RANELLA, N., EGER, M., 2023, “Towards Automated Video Game Commentary Using Generative AI.” In: *EXAG@ AIIDE*.
- RAO, A. S., GEORGEFF, M. P., 1995, “BDI Agents: From Theory to Practice”. In: *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, pp. 312–319.
- RAO, A. S., GEORGEFF, M. P., OTHERS, 1995, “BDI agents: from theory to practice.” In: *Icmas*, v. 95, pp. 312–319.
- RAUNAK, V., MENEZES, A., JUNCZYS-DOWMUNT, M., 2021, “The curious case of hallucinations in neural machine translation”. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 1172–1183.
- REED, S., ZOLNA, K., PARISOTTO, E., et al., 2022, “A generalist agent”, *Transactions on Machine Learning Research*.

- ROGERS, A., KOVALEVA, O., RUMSHISKY, A., 2021, “A primer in BERTology: What we know about how BERT works”, *Transactions of the association for computational linguistics*, v. 8, pp. 842–866.
- RUBENSTEIN, P. K., ASAWAROENGCHAI, C., NGUYEN, D. D., et al., 2023, “AudioPaLM: A large language model that can speak and listen”, *arXiv preprint arXiv:2306.12925*.
- RUMELHART, D. E., HINTON, G. E., WILLIAMS, R. J., 1986, “Learning representations by back-propagating errors”, *nature*, v. 323, n. 6088, pp. 533–536.
- RUSSELL, S., NORVIG, P., 2010, *Artificial Intelligence: A Modern Approach*. Prentice Hall.
- SABATER, J., SIERRA, C., 2005, “Review on Computational Trust and Reputation Models”, *Artificial Intelligence Review*, v. 24, n. 1, pp. 33–60.
- SAGIROVA, A., KURATOV, Y., BURTSEV, M., 2025, “SRMT: Shared Memory for Multi-agent Lifelong Pathfinding”, *arXiv preprint arXiv:2501.13200*.
- SAILER, M., HENSE, J. U., MAYR, S. K., et al., 2017, “How Gamification Motivates: An Experimental Study of the Effects of Specific Game Design Elements on Psychological Need Satisfaction”, *Computers in Human Behavior*, v. 69, pp. 371–380.
- SANDHOLM, T., LESSER, V., 1997, “Coalition Formation among Bounded Rational Agents”, *Artificial Intelligence*, v. 94, n. 1, pp. 99–137.
- SAWYER, R. K., 2014, *The Cambridge Handbook of the Learning Sciences*. Cambridge, Cambridge University Press.
- SCHICK, T., DWIVEDI-YU, J., DESSÌ, R., et al., 2024, “Toolformer: Language models can teach themselves to use tools”, *Advances in Neural Information Processing Systems*, v. 36.
- SCHUURMANS, D., 2023, “Memory augmented large language models are computationally universal”, *arXiv preprint arXiv:2301.04589*.
- SERAPIO-GARCÍA, G., SAFDARI, M., CREPY, C., et al., 2023, “Personality traits in large language models”, .
- SHAO, Z., GONG, Y., SHEN, Y., et al., 2023, “Enhancing retrieval-augmented large language models with iterative retrieval-generation synergy”. In:

Findings of the Association for Computational Linguistics: EMNLP 2023, pp. 9248–9274.

SHEN, X., YANG, Z., WANG, X., et al., 2023, “Global-to-local modeling for video-based 3D human pose and shape estimation”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8887–8896.

SHEPARD, L. A., 2000, “The Role of Assessment in a Learning Culture”, *Educational Researcher*, v. 29, n. 7, pp. 4–14.

SHI, W., HAN, X., LEWIS, M., et al., 2024, “Trusting your evidence: Hallucinate less with context-aware decoding”. In: *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers)*, pp. 783–791.

SHIN, D., 2021, “The algorithmic transparency paradox: Explaining and justifying algorithmic decision-making”, *AI & Society*, v. 36, n. 4, pp. 1185–1200.

SHIN, J., TANG, C., MOHATI, T., et al., 2023, “Prompt Engineering or Fine Tuning: An Empirical Assessment of Large Language Models in Automated Software Engineering Tasks”, *ArXiv*, v. abs/2310.10508. Disponível em: <<https://api.semanticscholar.org/CorpusID:264145854>>.

SHIN, T., RAZEGHI, Y., LOGAN IV, R. L., et al., 2020, “AutoPrompt: Eliciting knowledge from language models with automatically generated prompts”, *arXiv preprint arXiv:2010.15980*.

SHINN, N., CASSANO, F., GOPINATH, A., et al., 2023, “Reflexion: Language agents with verbal reinforcement learning”, *Advances in Neural Information Processing Systems*, v. 36, pp. 8634–8652.

SHOHAM, Y., LEYTON-BROWN, K., 2008, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press.

SHRIDHAR, M., YUAN, X., CÔTÉ, M.-A., et al., 2021, “ALFWorld: Aligning Text and Embodied Environments for Interactive Learning”, *International Conference on Learning Representation*.

SHUTE, V. J., VENTURA, M., KE, F., 2015, “Measuring and supporting learning in games: Stealth assessment”, *The International Journal of Artificial Intelligence in Education*, v. 25, n. 4, pp. 438–465.

- SINGHAL, K., AZIZI, S., TU, T., et al., 2023a, “Large language models encode clinical knowledge”, *Nature*, v. 620, pp. 172–180. doi: 10.1038/s41586-023-06291-2.
- SINGHAL, K., TU, T., GOTTWEIS, J., et al., 2023b, “Towards expert-level medical question answering with large language models”, *arXiv preprint arXiv:2305.09617*.
- SMITH, R. G., 1980, “The contract net protocol: High-level communication and control in a distributed problem solver”, *IEEE Transactions on computers*, v. 29, n. 12, pp. 1104–1113.
- SONG, C., WU, J., WASHINGTON, C., et al., 2023a, “LLM-planner: Few-shot grounded planning for embodied agents with large language models”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2998–3009, a.
- SONG, W., LI, Z., ZHANG, L., et al., 2023b, “Sparse is enough in fine-tuning pre-trained large language model”, *arXiv preprint arXiv:2312.11875*.
- SONG, Y., XIONG, W., ZHU, D., et al., 2023c, “RestGPT: Connecting large language models with real-world applications via RESTful APIs”, *arXiv preprint arXiv:2306.06624*.
- SQUIRE, K., 2011, *Video Games and Learning: Teaching and Participatory Culture in the Digital Age*. New York, Teachers College Press.
- STONE, P., VELOSO, M., 1999, “Task decomposition, dynamic role assignment, and low-bandwidth communication for real-time strategic teamwork”, *Artificial Intelligence*, v. 110, n. 2, pp. 241–273.
- STONE, P., VELOSO, M., 2000, “Multiagent Systems: A Survey from a Machine Learning Perspective”, *Autonomous Robots*, v. 8, n. 3, pp. 345–383.
- SUMERS, T. R., YAO, S., NARASIMHAN, K., et al., 2024, “Cognitive architectures for language agents”, *Transactions on Machine Learning Research*. ISSN: 2835-8856.
- SUN, R., ÖZCAN, S., ARIK, S., et al., 2023, “SQL-PaLM: Improved large language model adaptation for text-to-SQL (extended)”, *arXiv preprint arXiv:2306.00739*.
- TAMBE, M., 1997, “Towards flexible teamwork”, *Journal of artificial intelligence research*, v. 7, pp. 83–124.

- TAN, W., DING, Z., ZHANG, W., et al., 2023, “Towards general computer control: A multimodal agent for red dead redemption ii as a case study”, *arXiv preprint arXiv:2404.09982*.
- TANG, R., KONG, D., HUANG, L., et al., 2023a, “Large language models can be lazy learners: Analyze shortcuts in in-context learning”. In: *Findings of the Association for Computational Linguistics: ACL 2023*, pp. 4645–4657, a.
- TANG, X., ZOU, A., ZHANG, Z., et al., 2023b, “Medagents: Large language models as collaborators for zero-shot medical reasoning”, *arXiv preprint arXiv:2311.10537*.
- TANG, Y., SHIMADA, D., BI, J., et al., 2024, “Avicuna: Audio-visual llm with interleaver and context-boundary alignment for temporal referential dialogue”, *arXiv preprint arXiv:2403.16276*.
- 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.
- TEAM, M. F. A. R. D., BAKHTIN, A., BROWN, N., et al., 2022, “Human-level play in the game of Diplomacy by combining language models with strategic reasoning”, *Science*, v. 378, n. 6624, pp. 1067–1074.
- TEKIN, S. F., ILHAN, F., HUANG, T., et al., 2025, “Multi-Agent Reinforcement Learning with Focal Diversity Optimization”, *arXiv preprint arXiv:2502.04492*.
- TEO, T., 2011, “Factors influencing teachers’ intention to use technology”, *Computers & Education*, v. 57, n. 4, pp. 2432–2440.
- THORNE, J., VLACHOS, A., 2018, “Automated fact checking: Task formulations, methods and future directions”. In: *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 3346–3359.
- TOUVRON, H., LAVRIL, T., IZACARD, G., et al., 2023, “LLaMA: Open and efficient foundation language models”, *arXiv preprint arXiv:2302.13971*.
- TRAN, K.-T., DAO, D., NGUYEN, M.-D., et al., 2025, “Multi-Agent Collaboration Mechanisms: A Survey of LLMs”, *arXiv preprint arXiv:2501.06322*.

- TRIVEDI, H., BALASUBRAMANIAN, N., KHOT, T., et al., 2023, “Interleaving retrieval with chain-of-thought reasoning for knowledge-intensive multi-step questions”. In: *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 10014–10037.
- TU, T., AZIZI, S., DRIESS, D., et al., 2024, “Towards generalist biomedical AI”, *NEJM AI*, v. 1, n. 3, pp. A10a2300,138.
- TUYLS, K., NOWÉ, A., 2005, “Evolutionary Game Theory and Multi-Agent Reinforcement Learning”, *The Knowledge Engineering Review*, v. 20, n. 1, pp. 63–90.
- VAN DEN OORD, A., VINYALS, O., OTHERS, 2017, “Neural discrete representation learning”, *Advances in neural information processing systems*, v. 30.
- VASWANI, A., SHAZEER, N., PARMAR, N., et al., 2017, “Attention is all you need”, *Advances in neural information processing systems*, v. 30.
- VENKATESH, V., MORRIS, M. G., DAVIS, G. B., et al., 2003, “User acceptance of information technology: Toward a unified view”, *MIS quarterly*, pp. 425–478.
- VOM BROCKE, J., HEVNER, A., MAEDCHE, A., 2020, “Introduction to Design Science Research”. In: vom Brocke, J., Hevner, A., Maedche, A. (Eds.), *Design Science Research. Cases*, pp. 1–13, Cham, Springer International Publishing. ISBN: 978-3-030-46781-4. doi: 10.1007/978-3-030-46781-4_1. Disponível em: <https://doi.org/10.1007/978-3-030-46781-4_1>.
- WANG, A., PRUKSACHATKUN, Y., NANGIA, N., et al., 2019, “SuperGLUE: A stickier benchmark for general-purpose language understanding systems”, *Advances in Neural Information Processing Systems*, v. 32.
- WANG, G., XIE, Y., JIANG, Y., et al., 2023a, “Voyager: An Open-Ended Embodied Agent with Large Language Models”, *Transactions on Machine Learning Research*. Disponível em: <<https://arxiv.org/abs/2305.16291>>.
- WANG, G., LI, S., WANG, H., et al., 2023b, “CAMEL: Communicative Agents for "Mind" Exploration of Large Scale Language Model Society”, *arXiv preprint arXiv:2303.17760*.

- WANG, L., MA, C., FENG, X., et al., 2024a, “A survey on large language model based autonomous agents”, *Frontiers in Computer Science*, v. 18, n. 6, pp. 186345.
- WANG, L., MA, C., FENG, X., et al., 2024b, “A survey on large language model based autonomous agents”, *Frontiers of Computer Science*, v. 18, n. 6, pp. 186345.
- WANG, R., JANSEN, P., CÔTÉ, M.-A., et al., 2022, “ScienceWorld: Is your agent smarter than a 5th grader?” In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 11279–11298.
- WANG, X., YANG, Q., QIU, Y., et al., 2023c, “KnowledGPT: Enhancing large language models with retrieval and storage access on knowledge bases”, *arXiv preprint arXiv:2308.11761*.
- WANG, X., ZHANG, Y., ZOHAR, O., et al., 2024c, “VideoAgent: Long-form video understanding with large language model as agent”, *arXiv preprint arXiv:2403.10517*.
- WANG, X., XU, W., KAUSHIK, D., et al., 2023d, “Decision-oriented dialogue for human-ai collaboration”, *arXiv preprint arXiv:2305.20076*.
- WANG, Y., WU, Z., YAO, J., et al., 2025a, “Tdag: A multi-agent framework based on dynamic task decomposition and agent generation”, *Neural Networks*, p. 107200.
- WANG, Z., MORIYAMA, S., WANG, W.-Y., et al., 2025b, “Talk structurally, act hierarchically: A collaborative framework for llm multi-agent systems”, *arXiv preprint arXiv:2502.11098*.
- WANG, Z., CAI, S., CHEN, G., et al., 2023e, “Describe, explain, plan and select: Interactive planning with large language models enables open-world multi-task agents”, *arXiv preprint arXiv:2302.01560*.
- WANG, Z. J., CHOI, D., XU, S., et al., 2021, “Putting humans in the natural language processing loop: A survey”, *arXiv preprint arXiv:2103.04044*.
- WASON, P. C., JOHNSON-LAIRD, P. N., 1972, *Psychology of reasoning: Structure and content*. Harvard University Press, Cambridge, MA, USA.
- WEI, J., BOSMA, M., ZHAO, V. Y., et al., 2022a, “Finetuned language models are zero-shot learners”. In: *International Conference on Learning Representations*, a.

- WEI, J., TAY, Y., BOMMASANI, R., et al., 2022b, “Emergent abilities of large language models”, *Transactions on Machine Learning Research*.
- WEI, J., WANG, X., SCHUURMANS, D., et al., 2022c, “Chain-of-thought prompting elicits reasoning in large language models”, *Advances in Neural Information Processing Systems*, v. 35, pp. 24824–24837.
- WEISS, G., 1999, *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, MA, MIT Press.
- WEISS, K., KHOSHGOFTAAR, T. M., WANG, D., 2016, “A survey of transfer learning”, *Journal of Big data*, v. 3, pp. 1–40.
- WEISS, M., RAHAMAN, N., WUTHRICH, M., et al., 2024, “Rethinking the buyers inspection paradox in information markets with language agents”, .
- WOOLDRIDGE, M., JENNINGS, N. R., 1995, “Intelligent Agents: Theory and Practice”, *The Knowledge Engineering Review*, v. 10, n. 2, pp. 115–152.
- WOUTERS, P., VAN NIMWEGEN, C., VAN OOSTENDORP, H., et al., 2013, “A meta-analysis of the cognitive and motivational effects of serious games”, *Journal of Educational Psychology*, v. 105, n. 2, pp. 249.
- WU, J., ANTONOVA, R., KAN, A., et al., 2023a, “Tidybot: Personalized robot assistance with large language models”, *Autonomous Robots*, v. 47, n. 8, pp. 1087–1102.
- WU, Q., BANSAL, G., ZHANG, J., et al., 2023b, “Autogen: Enabling next-gen llm applications via multi-agent conversation”, *arXiv preprint arXiv:2308.08155*.
- WU, Q., BANSAL, G., ZHANG, J., et al., 2023c, “Autogen: Enabling next-gen llm applications via multi-agent conversation”, *arXiv preprint arXiv:2308.08155*.
- WU, Y., MIN, S. Y., BISK, Y., et al., 2022, “Plan, eliminate, and track language models are good teachers for embodied agents”, *arXiv preprint arXiv:2305.02412*.
- WU, Z., HU, Y., SHI, W., et al., 2024a, “Fine-grained human feedback gives better rewards for language model training”, *Advances in Neural Information Processing Systems*, v. 36.

- WU, Z., SHEN, C., JIANG, M., 2024b, “Instructing large language models to identify and ignore irrelevant conditions”, *arXiv preprint arXiv:2403.12744*.
- XI, Z., CHEN, W., GUO, X., et al., 2025, “The rise and potential of large language model based agents: A survey”, *Science China Information Sciences*, v. 68, n. 2, pp. 121101.
- XIA, Y., SHENOY, M., JAZDI, N., et al., 2023, “Towards autonomous system: flexible modular production system enhanced with large language model agents”. In: *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp. 1–8. IEEE.
- XU, B., PENG, Z., LEI, B., et al., 2023a, “ReWOO: Decoupling reasoning from observations for efficient augmented language models”, *arXiv preprint arXiv:2305.18323*.
- XU, F., SHI, W., CHOI, E., 2023b, “Recomp: Improving retrieval-augmented LMs with compression and selective augmentation”. In: *The Twelfth International Conference on Learning Representations*, b.
- XU, F. F., ALON, U., NEUBIG, G., et al., 2022, “A systematic evaluation of large language models of code”. In: *Proceedings of the 6th ACM SIGPLAN International Symposium on Machine Programming*, pp. 1–10.
- XU, Y., WANG, S., LI, P., et al., 2023c, “Exploring large language models for communication games: An empirical study on werewolf”, *arXiv preprint arXiv:2309.04658*.
- YANG, H., CHEN, J., SIEW, M., et al., 2025, “LLM-Powered Decentralized Generative Agents with Adaptive Hierarchical Knowledge Graph for Cooperative Planning”, *arXiv preprint arXiv:2502.05453*.
- YANG, H., LI, Z., ZHANG, Y., et al., 2023a, “PRCA: Fitting black-box large language models for retrieval question answering via pluggable reward-driven contextual adapter”. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 5364–5375, a.
- YANG, J., DONG, Y., LIU, S., et al., 2023b, “Octopus: Embodied vision-language programmer from environmental feedback”, *arXiv preprint arXiv:2310.08588*.
- YANG, Y., ZHUANG, Y., PAN, Y., 2021, “Multiple knowledge representation for big data artificial intelligence: framework, applications, and case studies”,

- Frontiers of Information Technology & Electronic Engineering*, v. 22, n. 12, pp. 1551–1558.
- YANG, Z., LIU, J., HAN, Y., et al., 2023c, “AppAgent: Multimodal agents as smartphone users”, *arXiv preprint arXiv:2312.13771*.
- YAO, S., YU, D., ZHAO, J., et al., 2024, “Tree of thoughts: Deliberate problem solving with large language models”, *Advances in Neural Information Processing Systems*, v. 36.
- YAO, W., HEINECKE, S., NIEBLES, J. C., et al., 2023a, “Retroformer: Retrospective large language agents with policy gradient optimization”, *arXiv preprint arXiv:2308.02151*.
- YAO, Y., WANG, P., TIAN, B., et al., 2023b, “Editing large language models: Problems, methods, and opportunities”. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 10222–10240, b.
- YE, H., LIN, M., NG, H. T., et al., 2024, “Multi-Agent Sampling: Scaling Inference Compute for Data Synthesis with Tree Search-Based Agentic Collaboration”, *arXiv preprint arXiv:2412.17061*.
- YU, W., ITER, D., WANG, S., et al., 2022, “Generate rather than retrieve: Large language models are strong context generators”. In: *The Eleventh International Conference on Learning Representations*.
- YU, W., TANG, C., SUN, G., et al., 2024a, “Connecting speech encoder and large language model for ASR”. In: *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 12637–12641, a.
- YU, Y., YIN, Q., ZHANG, J., et al., 2024b, “ADMN: agent-driven modular network for dynamic parameter sharing in cooperative multi-agent reinforcement learning”. In: *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pp. 302–310, b.
- ZAKEN, E. B., RAVFOGEL, S., GOLDBERG, Y., 2021, “BitFit: Simple parameter-efficient fine-tuning for transformer-based masked language-models”, *arXiv preprint arXiv:2106.10199*.
- ZHA, L., ZHOU, J., LI, L., et al., 2023, “TableGPT: Towards unifying tables, nature language and commands into one GPT”, *arXiv preprint arXiv:2307.08674*.

- ZHANG, A., CHEN, Y., SHENG, L., et al., 2024a, “On generative agents in recommendation”. In: *Proceedings of the 47th international ACM SIGIR conference on research and development in Information Retrieval*, pp. 1807–1817, a.
- ZHANG, C., YANG, K., HU, S., et al., 2024b, “Proagent: building proactive cooperative agents with large language models”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, v. 38, pp. 17591–17599, b.
- ZHANG, D., LI, S., ZHANG, X., et al., 2023a, “SpeechGPT: Empowering large language models with intrinsic cross-modal conversational abilities”. In: *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 15757–15773, a.
- ZHANG, D., LI, Z., WANG, P., et al., 2024c, “Speechagents: Human-communication simulation with multi-modal multi-agent systems”, *arXiv preprint arXiv:2401.03945*.
- ZHANG, H., LI, X., BING, L., 2023b, “Video-LLaMA: An instruction-tuned audio-visual language model for video understanding”. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pp. 543–553, b.
- ZHANG, H., DU, W., SHAN, J., et al., 2023c, “Building cooperative embodied agents modularly with large language models”, *arXiv preprint arXiv:2307.02485*.
- ZHANG, H., DU, W., SHAN, J., et al., 2023d, “Building cooperative embodied agents modularly with large language models”, *arXiv preprint arXiv:2307.02485*.
- ZHANG, J., XU, X., ZHANG, N., et al., 2023e, “Exploring collaboration mechanisms for llm agents: A social psychology view”, *arXiv preprint arXiv:2310.02124*.
- ZHANG, J., HOU, Y., XIE, R., et al., 2024d, “Agentcf: Collaborative learning with autonomous language agents for recommender systems”. In: *Proceedings of the ACM Web Conference 2024*, pp. 3679–3689, d.
- ZHANG, S., PAN, L., ZHAO, J., et al., 2023f, “Mitigating language model hallucination with interactive question-knowledge alignment”, *arXiv preprint arXiv:2305.13669*, v. 3.

- ZHANG, Y., YANG, S., BAI, C., et al., 2024e, “Towards efficient llm grounding for embodied multi-agent collaboration”, *arXiv preprint arXiv:2405.14314*.
- ZHANG, Y., LI, Y., CUI, L., et al., 2023g, “Siren’s song in the AI ocean: A survey on hallucination in large language models”, *arXiv preprint arXiv:2309.01219*.
- ZHANG, Z., ZHAO, H., 2021, “Advances in multi-turn dialogue comprehension: A survey”, *arXiv preprint arXiv:2103.03125*.
- ZHANG, Z.-R., TAN, C., XU, H., et al., 2023h, “Towards adaptive prefix tuning for parameter-efficient language model fine-tuning”, *arXiv preprint arXiv:2305.15212*.
- ZHANG, Z., YANG, Z., YANG, Y., 2024f, “Sifu: Side-view conditioned implicit function for real-world usable clothed human reconstruction”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9936–9947, f.
- ZHAO, A., HUANG, D., XU, Q., et al., 2024, “Expel: Llm agents are experiential learners”, *Proceedings of the AAAI Conference on Artificial Intelligence*, v. 38, n. 2024, pp. 19632–19642.
- ZHAO, Z., WALLACE, E., FENG, S., et al., 2021, “Calibrate before use: Improving few-shot performance of language models”. In: *International conference on machine learning*, pp. 12697–12706.
- ZHONG, V., DEFREITAS, J., WAGNER, A. P., et al., 2022, “Exploring possible human-controlled language-based strategic behavior in an iterated social dilemma”, *arXiv preprint arXiv:2209.15345*.
- ZHONG, W., GUO, L., GAO, Q., et al., 2024, “MemoryBank: Enhancing large language models with long-term memory”, *Proceedings of the AAAI Conference on Artificial Intelligence*, v. 38, n. 2024, pp. 19724–19731.
- ZHOU, C., LIU, P., XU, P., et al., 2024a, “LIMA: Less is more for alignment”, *Advances in Neural Information Processing Systems*, v. 36.
- ZHOU, X., LI, G., LIU, Z., 2023a, “LLM as DBA”, *arXiv preprint arXiv:2308.05481*.
- ZHOU, Y., WANG, L., ZHANG, J., et al., 2023b, “RecAgent: A Novel Simulation Paradigm for Recommender Systems”, *arXiv preprint arXiv:2306.02552*.

- ZHOU, Y., XU, P., LIU, X., et al., 2024b, “Explore spurious correlations at the concept level in language models for text classification”. In: *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 478–492, b.
- ZHU, L., HU, L., LIN, J., et al., 2023a, “LIFT: Efficient layer-wise fine-tuning for large model models”. In: *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 4678–4687, a.
- ZHU, X., CHEN, Y., TIAN, H., et al., 2023b, “Ghost in the Minecraft: Generally capable agents for open-world environments via large language models with text-based knowledge and memory”, *arXiv preprint arXiv:2305.17144*.
- ZHUANG, S., LIU, B., KOOPMAN, B., et al., 2023, “Open-source large language models are strong zero-shot query likelihood models for document ranking”. In: *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 8807–8817.
- ZHUGE, W., SU, Y., ZUO, J., et al., 2023, “AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors”, *arXiv preprint arXiv:2308.10848*.

Apêndice A

Estudos de Caso Detalhados

Este apêndice apresenta os 26 estudos de caso realizados para validação do sistema multiagente proposto. Cada estudo de caso documenta o processo completo de design de um jogo educacional, desde a fase de brainstorming até a classificação de objetivos educacionais segundo a Taxonomia de Bloom.

A.1 Estudo de Caso 1: Ensino de Matemática no Ensino Fundamental

A.1.1 Design de Jogo para Ensino de Matemática no Ensino Fundamental

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional voltado para o ensino de operações matemáticas básicas.

Fase de Brainstorming: Durante a sessão de brainstorming, o usuário (um professor de matemática) inseriu as seguintes ideias iniciais (4 inputs):

- Jogo sobre operações básicas de matemática
- Foco em adição e subtração para crianças de 8-10 anos
- Ambiente lúdico com personagens animados
- Sistema de recompensas para motivar os alunos

O sistema, através do serviço GroqService, analisou estas ideias e sugeriu conceitos complementares (4 sugestões):

- Narrativa de exploradores matemáticos em uma ilha misteriosa
- Desafios contextualizados em problemas do cotidiano
- Progressão de dificuldade adaptativa baseada no desempenho
- Elementos cooperativos para incentivar trabalho em equipe

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas padronizadas:

- “Como vai ser esse jogo? Qual será o gênero do jogo? Quem serão os personagens? Para que tipo de plataforma será lançado? Será um jogo apenas de texto, ou haverá um trabalho de arte envolvido?”
- “Como o jogo garantirá que os conceitos educacionais sejam incorporados de forma endógena e não apenas como uma camada superficial?”
- “De que forma o sistema de recompensas se alinha com os objetivos pedagógicos?”
- “Como o jogo lidará com diferentes ritmos de aprendizagem e níveis de conhecimento prévio?”
- “Quais mecanismos garantirão que o aspecto lúdico não ofusque o conteúdo educacional?”

As respostas do usuário a estas questões foram analisadas e refinadas, gerando insights como:

- Incorporar operações matemáticas como mecânicas centrais para avançar na narrativa
- Utilizar recompensas que reforcem conceitos matemáticos e promovam curiosidade
- Implementar sistema de dificuldade dinâmica que se ajusta discretamente ao desempenho
- Equilibrar desafios matemáticos com elementos narrativos e exploratórios

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas para cada seção:

- **Agente de Mecânicas:** Sugeriu um sistema de combinação de elementos através de operações matemáticas, onde a soma ou subtração de elementos gera novos itens no jogo.
- **Agente de Narrativa:** Desenvolveu uma narrativa sobre uma ilha onde os números e operações têm efeitos mágicos no ambiente, e os jogadores devem utilizar este conhecimento para ajudar os habitantes locais.
- **Agente de Engajamento:** Propôs um sistema de colecionáveis e conquistas vinculadas a marcos de aprendizagem, com desafios diários para manter o interesse.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais alinhados com a Taxonomia de Bloom:

- **Lembrar:** Reconhecer e reproduzir os algoritmos básicos de adição e subtração.

- **Entender:** Explicar o processo de reagrupamento (vai-um) na adição e o empréstimo na subtração.
- **Aplicar:** Utilizar operações de adição e subtração para resolver problemas contextualizados na narrativa do jogo.
- **Analisar:** Decompor problemas matemáticos complexos em etapas mais simples e identificar a operação apropriada.
- **Avaliar:** Verificar a razoabilidade de respostas e detectar erros em cálculos.
- **Criar:** Desenvolver estratégias eficientes para resolver problemas matemáticos em contextos variados.



Figura A.1: Endo-GDC do jogo “Ilha dos Números”

Resultado Final - Endo-GDC do Jogo “Ilha dos Números”: O Endo-GDC do jogo “Ilha dos Números” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema onde operações matemáticas funcionam

como ações que transformam o ambiente do jogo, com números e operações permitindo a criação de ferramentas, superação de obstáculos e resolução de quebra-cabeças ambientais.

- **Narrativa Integrada:** História sobre uma ilha mágica onde os habitantes usam matemática para controlar elementos da natureza, com missões que requerem aplicação progressiva de conceitos matemáticos.
- **Elementos de Engajamento:** Sistema de progressão baseado em conquistas matemáticas, com recompensas que ampliam as habilidades do jogador e desbloqueiam novas áreas da ilha para exploração.
- **Avaliação Incorporada:** Mecanismos de feedback que mostram visualmente o resultado de operações matemáticas no ambiente, permitindo auto-correção e aprendizado por experimentação.

A.2 Estudo de Caso 2: Ensino de Ciências no Ensino Médio

A.2.1 Design de Jogo para Ensino de Ciências no Ensino Médio

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo que ensina conceitos de ecologia e sustentabilidade.

Fase de Brainstorming: Um designer educacional inseriu as seguintes ideias iniciais (6 inputs):

- Simulação de um ecossistema
- Foco em cadeias alimentares e equilíbrio ecológico
- Público-alvo: estudantes de 14-17 anos
- Formato de simulação com elementos de estratégia
- Interface que permite visualização de dados científicos
- Componente multijogador para simular cooperação e competição por recursos

O sistema sugeriu os seguintes conceitos complementares (5 sugestões):

- Conexão entre decisões locais e impactos globais
- Eventos aleatórios como desastres naturais e intervenções humanas
- Escala temporal que permite observar mudanças a longo prazo
- Múltiplas perspectivas: cientistas, políticos, cidadãos, empresas
- Sistema de pesquisa científica para desbloquear tecnologias sustentáveis

Fase de Método Socrático: O Agente Socrático aplicou as mesmas questões reflexivas padronizadas:

- “Como vai ser esse jogo? Qual será o gênero do jogo? Quem serão os personagens? Para que tipo de plataforma será lançado? Será um jogo apenas de texto, ou haverá um trabalho de arte envolvido?”
- “Como o jogo garantirá que os conceitos educacionais sejam incorporados de forma endógena e não apenas como uma camada superficial?”
- “De que forma o sistema de recompensas se alinha com os objetivos pedagógicos?”
- “Como o jogo lidará com diferentes ritmos de aprendizagem e níveis de conhecimento prévio?”
- “Quais mecanismos garantirão que o aspecto lúdico não ofusque o conteúdo educacional?”

As respostas levaram a insights como:

- Incorporar modelos científicos reais em versões simplificadas mas fiéis
- Utilizar narrativas pessoais e conexões emocionais para complementar dados científicos
- Incorporar elementos probabilísticos e sistemas de feedback complexos
- Incluir cenários baseados em problemas reais da comunidade local dos estudantes
- Implementar sistema de consequências não-lineares para demonstrar interdependência ecológica

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas para cada seção:

- **Agente de Mecânicas:** Propôs um sistema de simulação baseado em ciclos naturais, onde as ações dos jogadores influenciam múltiplas variáveis do ecossistema com efeitos em cascata ao longo do tempo.
- **Agente de Narrativa:** Desenvolveu uma narrativa que acompanha o desenvolvimento de uma região ao longo de décadas, mostrando as consequências de longo prazo das decisões ambientais.
- **Agente de Engajamento:** Sugeriu um sistema de "reputação científica" que evolui conforme os jogadores tomam decisões baseadas em evidências e mantêm a sustentabilidade.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar componentes essenciais de um ecossistema e suas interações básicas.
- **Entender:** Explicar como energia e matéria fluem através das cadeias alimentares e ciclos biogeoquímicos.

- **Aplicar:** Implementar estratégias de conservação em diferentes cenários ecológicos.
- **Analisar:** Prever potenciais efeitos de perturbações em um ecossistema através da análise de dados.
- **Avaliar:** Julgar a sustentabilidade de diferentes abordagens de gestão ambiental com base em evidências.
- **Criar:** Desenvolver soluções inovadoras para desafios ambientais que equilibrem necessidades humanas e ecológicas.

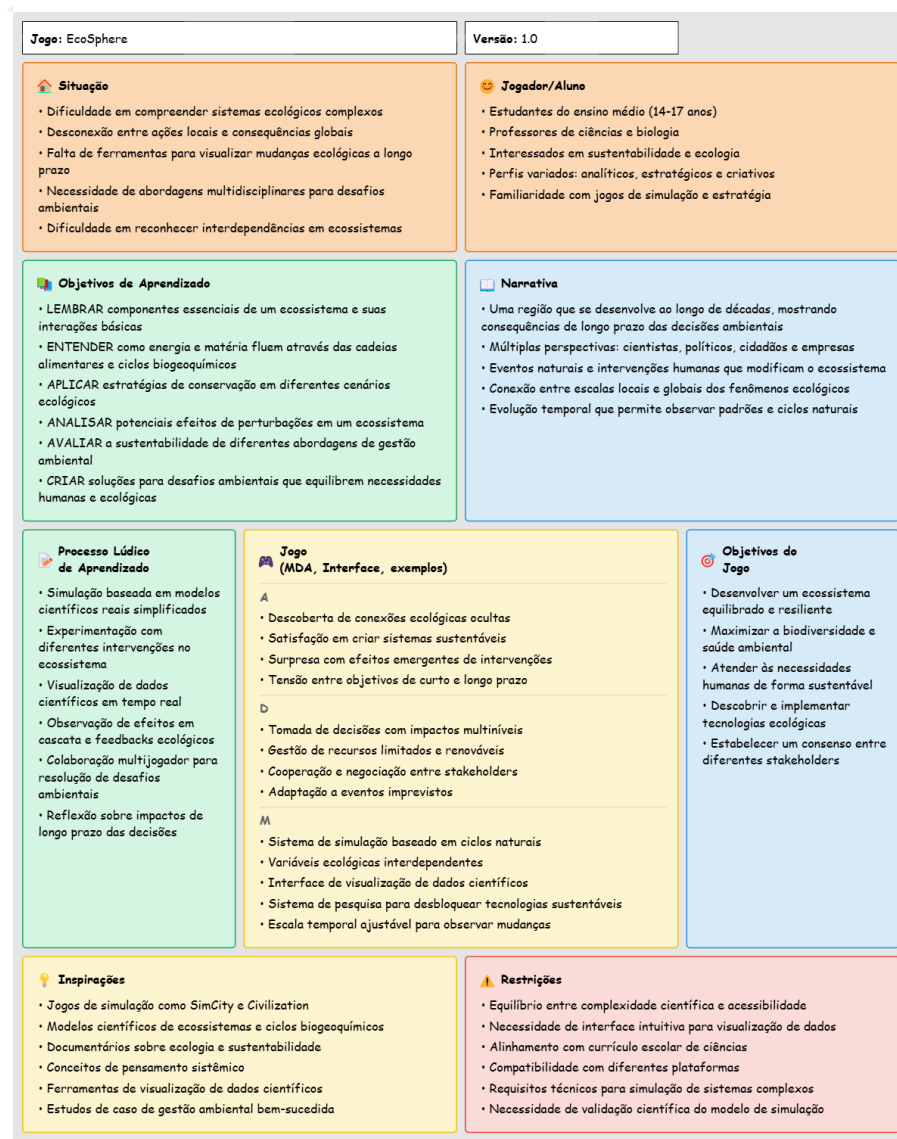


Figura A.2: Endo-GDC do jogo “EcoSphere”

Resultado Final - Endo-GDC do Jogo “EcoSphere”: O Endo-GDC do jogo “EcoSphere” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de simulação ecológica onde cada ação afeta

múltiplas variáveis ambientais, representadas visualmente através de gráficos e mudanças no ecossistema virtual.

- **Narrativa Integrada:** História que evolui ao longo de gerações, mostrando como decisões anteriores criam o contexto para desafios futuros, com personagens que representam diferentes perspectivas sobre sustentabilidade.
- **Elementos de Engajamento:** Sistema de reconhecimento científico e social que valoriza tanto a experimentação metodológica quanto os resultados de longo prazo na saúde do ecossistema.
- **Avaliação Incorporada:** Ferramentas de análise científica que permitem aos jogadores formular hipóteses, coletar dados e avaliar o impacto de suas intervenções no ecossistema.

A.3 Estudo de Caso 3: Ensino de Idiomas para Adultos

A.3.1 Design de Jogo para Ensino de Idiomas para Adultos

Este caso demonstra a aplicação do sistema para o design de um jogo voltado ao aprendizado de línguas estrangeiras.

Fase de Brainstorming: Um professor especializado em ensino de idiomas inseriu as seguintes ideias iniciais (3 inputs):

- Jogo para aprendizado de vocabulário e construção de frases
- Foco em situações práticas do cotidiano profissional
- Integração entre compreensão auditiva e expressão oral

O sistema sugeriu os seguintes conceitos complementares (7 sugestões):

- Ambiente de simulação social com personagens não-jogáveis interativos
- Sistema de reconhecimento de pronúncia com feedback instantâneo
- Progressão baseada em níveis de fluência mensuráveis
- Modo de viagem virtual a países onde o idioma é falado
- Integração com conteúdo de mídia autêntico (notícias, filmes, música)
- Customização de contextos profissionais específicos
- Sistema de revisão espaçada para reforço de vocabulário

Fase de Método Socrático: O Agente Socrático aplicou as mesmas questões reflexivas padronizadas:

- “Como vai ser esse jogo? Qual será o gênero do jogo? Quem serão os personagens? Para que tipo de plataforma será lançado? Será um jogo apenas de texto, ou haverá um trabalho de arte envolvido?”

- “Como o jogo garantirá que os conceitos educacionais sejam incorporados de forma endógena e não apenas como uma camada superficial?”
- “De que forma o sistema de recompensas se alinha com os objetivos pedagógicos?”
- “Como o jogo lidará com diferentes ritmos de aprendizagem e níveis de conhecimento prévio?”
- “Quais mecanismos garantirão que o aspecto lúdico não ofusque o conteúdo educacional?”

As respostas geraram os seguintes insights:

- Desenvolver um RPG conversacional onde todo avanço depende da comunicação no idioma-alvo
- Criar situações que exijam uso autêntico da língua para resolução de problemas
- Implementar um sistema adaptativo que ajusta o vocabulário e complexidade gramatical
- Incorporar análise linguística em tempo real para feedback personalizado

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Desenvolveu um sistema onde o diálogo é a mecânica central, com reconhecimento de fala e técnicas de processamento de linguagem natural para avaliar respostas dos jogadores.
- **Agente de Narrativa:** Criou uma narrativa sobre um profissional recém-chegado a um país estrangeiro, enfrentando desafios cotidianos e profissionais que requerem comunicação efetiva.
- **Agente de Engajamento:** Propôs um sistema de networking social dentro do jogo, onde conexões com personagens não-jogáveis desbloqueiam oportunidades e recursos.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Reconhecer e reproduzir vocabulário e estruturas gramaticais em contextos relevantes.
- **Entender:** Interpretar comunicações no idioma-alvo e responder apropriadamente.
- **Aplicar:** Utilizar o idioma para resolver problemas cotidianos e profissionais.
- **Analisar:** Distinguir nuances de significado e intenção em diferentes contextos comunicativos.
- **Avaliar:** Julgar a eficácia de diferentes estratégias comunicativas em diversos contextos.

- **Criar:** Produzir comunicações originais adaptadas a objetivos específicos e audiências diversas.

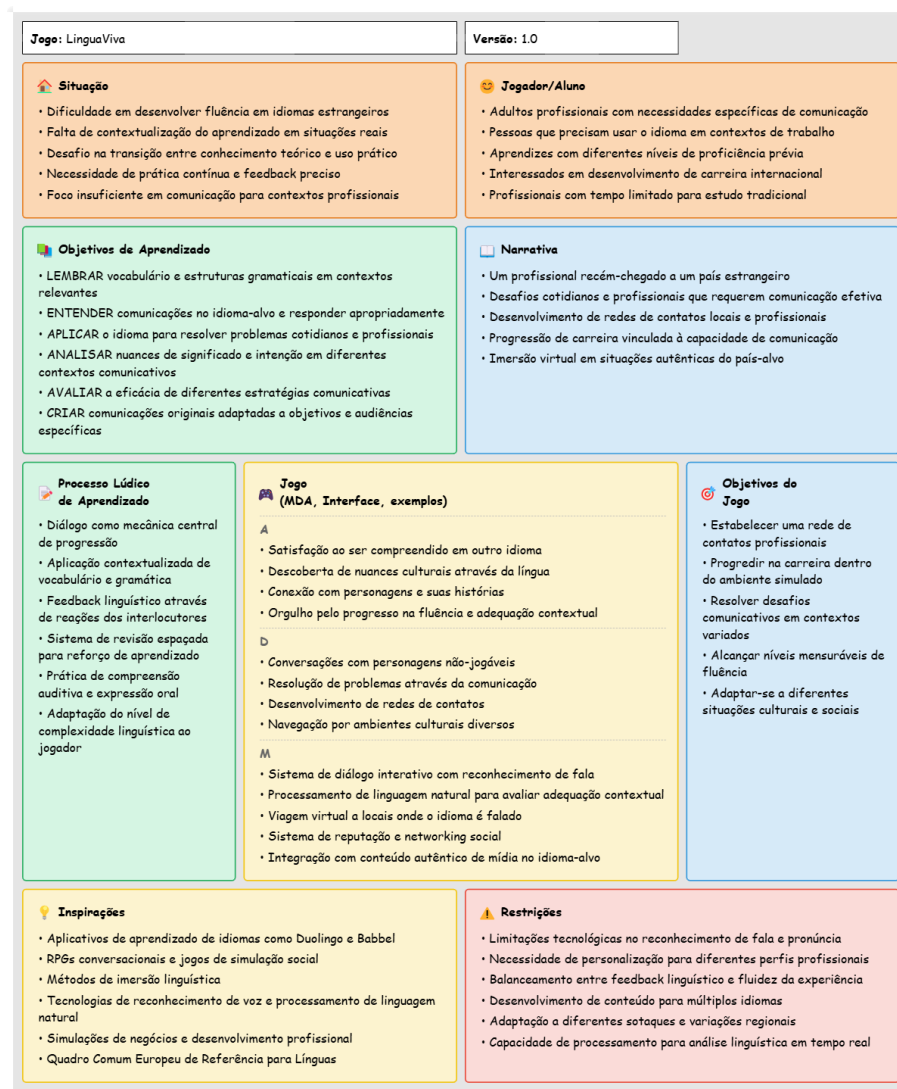


Figura A.3: Endo-GDC do jogo “LinguaViva”

Resultado Final - Endo-GDC do Jogo “LinguaViva”: O Endo-GDC do jogo “LinguaViva” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de diálogo interativo onde a capacidade de comunicação determina diretamente o progresso, com uso de processamento de linguagem natural para avaliar não apenas correção gramatical mas também adequação contextual.
- **Narrativa Integrada:** Simulação de carreira internacional onde as habilidades linguísticas afetam opções de trabalho, relacionamentos profissionais e oportunidades de crescimento.
- **Elementos de Engajamento:** Sistema de reputação e redes de contato que

evolui conforme o jogador demonstra proficiência comunicativa em diferentes contextos sociais e profissionais.

- **Avaliação Incorporada:** Feedback linguístico contextualizado que não interrompe a experiência, apresentado como reações naturais dos interlocutores e consequências práticas da comunicação.

A.4 Estudo de Caso 4: Formação em Liderança e Gestão

A.4.1 Design de Jogo para Formação Corporativa

Este estudo de caso ilustra a aplicação do sistema para o contexto de treinamento corporativo.

Fase de Brainstorming: Um especialista em treinamento corporativo inseriu apenas uma ideia inicial (1 input):

- Simulação de gestão de equipes em ambiente de trabalho híbrido

O sistema, demonstrando capacidade de expansão a partir de inputs mínimos, gerou as seguintes sugestões (8 sugestões):

- Foco em comunicação eficiente e resolução de conflitos
- Cenários baseados em situações reais de trabalho remoto e presencial
- Sistema de feedback multidimensional (técnico, emocional, organizacional)
- Personagens com diferentes perfis comportamentais e profissionais
- Métricas balanceadas entre bem-estar da equipe e produtividade
- Eventos imprevistos que exigem adaptação rápida de estratégias
- Múltiplos caminhos de resolução baseados em diferentes estilos de liderança
- Componente de reflexão pós-ação para análise de decisões tomadas

Fase de Método Socrático: O Agente Socrático aplicou as questões reflexivas padronizadas:

- “Como vai ser esse jogo? Qual será o gênero do jogo? Quem serão os personagens? Para que tipo de plataforma será lançado? Será um jogo apenas de texto, ou haverá um trabalho de arte envolvido?”
- “Como o jogo garantirá que os conceitos educacionais sejam incorporados de forma endógena e não apenas como uma camada superficial?”
- “De que forma o sistema de recompensas se alinha com os objetivos pedagógicos?”
- “Como o jogo lidará com diferentes ritmos de aprendizagem e níveis de conhecimento prévio?”

- “Quais mecanismos garantirão que o aspecto lúdico não ofusque o conteúdo educacional?”

As respostas geraram os seguintes insights:

- Criação de uma simulação com elementos visuais profissionais e interface de gerenciamento
- Implementação de um sistema onde teoria e prática de liderança são inseparáveis
- Desenvolvimento de métricas complexas que refletem a multidimensionalidade da liderança efetiva
- Incorporação de um sistema de mentoria in-game para orientação contextual

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "rodadas de comunicação" com limitações de tempo e canais, simulando os desafios de equipes distribuídas geograficamente.
- **Agente de Narrativa:** Desenvolveu um cenário de uma empresa fictícia enfrentando transformação digital e transição para modelo de trabalho híbrido.
- **Agente de Engajamento:** Sugeriu um sistema de métricas de desempenho multidimensional que avalia não apenas resultados, mas também processos e impactos nas pessoas.

Fase de Taxonomia de Bloom: O sistema adaptou a Taxonomia de Bloom para o contexto corporativo:

- **Lembrar:** Identificar princípios fundamentais de liderança e gestão de equipes híbridas.
- **Entender:** Reconhecer padrões de comportamento organizacional e dinâmicas de equipe.
- **Aplicar:** Implementar técnicas específicas de comunicação e gestão em cenários diversos.
- **Analisar:** Interpretar dados de desempenho e feedback para identificar causas raiz de problemas.
- **Avaliar:** Julgar a eficácia de diferentes abordagens de liderança em contextos específicos.
- **Criar:** Desenvolver estratégias personalizadas para motivar equipes diversas e alcançar objetivos organizacionais.

Resultado Final - Endo-GDC do Jogo “LeadSync”: O Endo-GDC do jogo “LeadSync” integra os seguintes elementos principais:

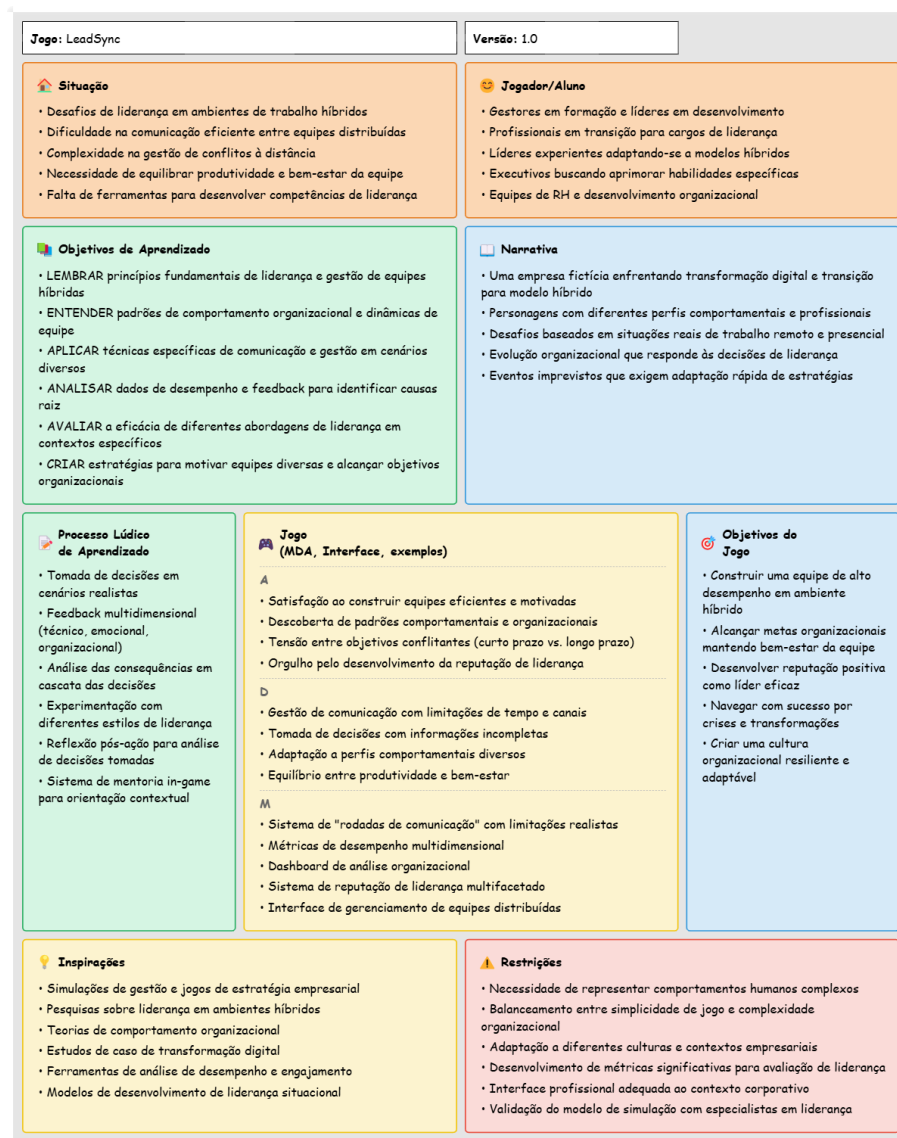


Figura A.4: Endo-GDC do jogo “LeadSync”

- **Mecânicas Endógenas:** Sistema de simulação organizacional onde práticas de liderança impactam diretamente o desempenho, engajamento e desenvolvimento da equipe, com métricas interdependentes que refletem a complexidade de ambientes de trabalho reais.
- **Narrativa Integrada:** Cenário organizacional dinâmico que evolui em resposta às decisões de liderança, com personagens que demonstram comportamentos realistas e adaptáveis.
- **Elementos de Engajamento:** Sistema de "reputação de liderança" multifacetado que reflete diferentes dimensões da eficácia gerencial e se manifesta através das interações e desempenho da equipe.
- **Avaliação Incorporada:** Dashboards de análise organizacional que permitem reflexão sobre causas e efeitos das decisões tomadas, com feedback apre-

sentado como resultados naturais no ambiente de trabalho simulado.

A.5 Estudo de Caso 5: Ensino de História no Ensino Médio

A.5.1 Design de Jogo para Ensino de História no Ensino Médio

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo educacional sobre eventos históricos para estudantes do Ensino Médio.

Fase de Brainstorming: Um professor de história inseriu apenas duas ideias iniciais (2 inputs):

- Jogo sobre a Revolução Industrial e suas consequências
- Foco em mudanças sociais, econômicas e tecnológicas

O sistema, através do serviço GroqService, analisou estas ideias e sugeriu diversos conceitos complementares (7 sugestões):

- Simulação de desenvolvimento urbano durante a industrialização
- Narrativa que acompanha uma família através de gerações
- Sistema de causa e efeito mostrando impactos de decisões políticas
- Representação de diferentes classes sociais e suas perspectivas
- Elementos de investigação histórica usando fontes primárias
- Conexões entre eventos históricos e questões contemporâneas
- Dilemas éticos baseados em situações históricas reais

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo poderá representar diferentes perspectivas históricas sem privilegiar um único ponto de vista?”
- “De que forma o jogo equilibrará a simplificação necessária para a jogabilidade com a complexidade histórica real?”
- “Como serão abordadas questões sensíveis como exploração trabalhista e colonialismo?”

As respostas geraram os seguintes insights:

- Implementar um sistema de múltiplas perspectivas onde os jogadores vivenciam os mesmos eventos através de diferentes personagens
- Utilizar níveis de profundidade que permitam exploração adicional para estudantes interessados
- Incorporar fontes históricas reais como elementos de descoberta no jogo

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de tomada de decisões com ramificações que mostram consequências a curto e longo prazo, inspirado em jogos de estratégia e simulação social.
- **Agente de Narrativa:** Desenvolveu uma estrutura narrativa centrada em uma cidade industrial fictícia com personagens representando diferentes estratos sociais, incluindo trabalhadores, empresários e reformistas.
- **Agente de Engajamento:** Sugeriu um sistema de descobertas baseado em investigação histórica, onde os jogadores coletam e analisam documentos, imagens e relatos da época.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar os principais eventos, invenções e figuras da Revolução Industrial.
- **Entender:** Explicar as relações causais entre industrialização e transformações sociais.
- **Aplicar:** Relacionar processos históricos a contextos atuais de desenvolvimento tecnológico e desigualdade.
- **Analisar:** Examinar fontes históricas e extrair informações relevantes sobre diferentes perspectivas.
- **Avaliar:** Julgar decisões históricas considerando o contexto da época e suas consequências.
- **Criar:** Desenvolver cenários alternativos baseados em diferentes escolhas políticas e sociais.

Resultado Final - Endo-GDC do Jogo “Tempos de Mudança”: O Endo-GDC do jogo “Tempos de Mudança” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de simulação histórica com tomada de decisões que afetam o desenvolvimento urbano, condições sociais e inovações tecnológicas ao longo de gerações.
- **Narrativa Integrada:** História centrada em uma cidade industrial em transformação, com múltiplos personagens jogáveis representando diferentes classes sociais e ideologias da época.
- **Elementos de Engajamento:** Sistema de descoberta de artefatos históricos e documentos que revelam novas perspectivas e desbloqueiam opções adicionais, incentivando a exploração aprofundada.
- **Avaliação Incorporada:** Mecanismos de feedback visual e narrativo que

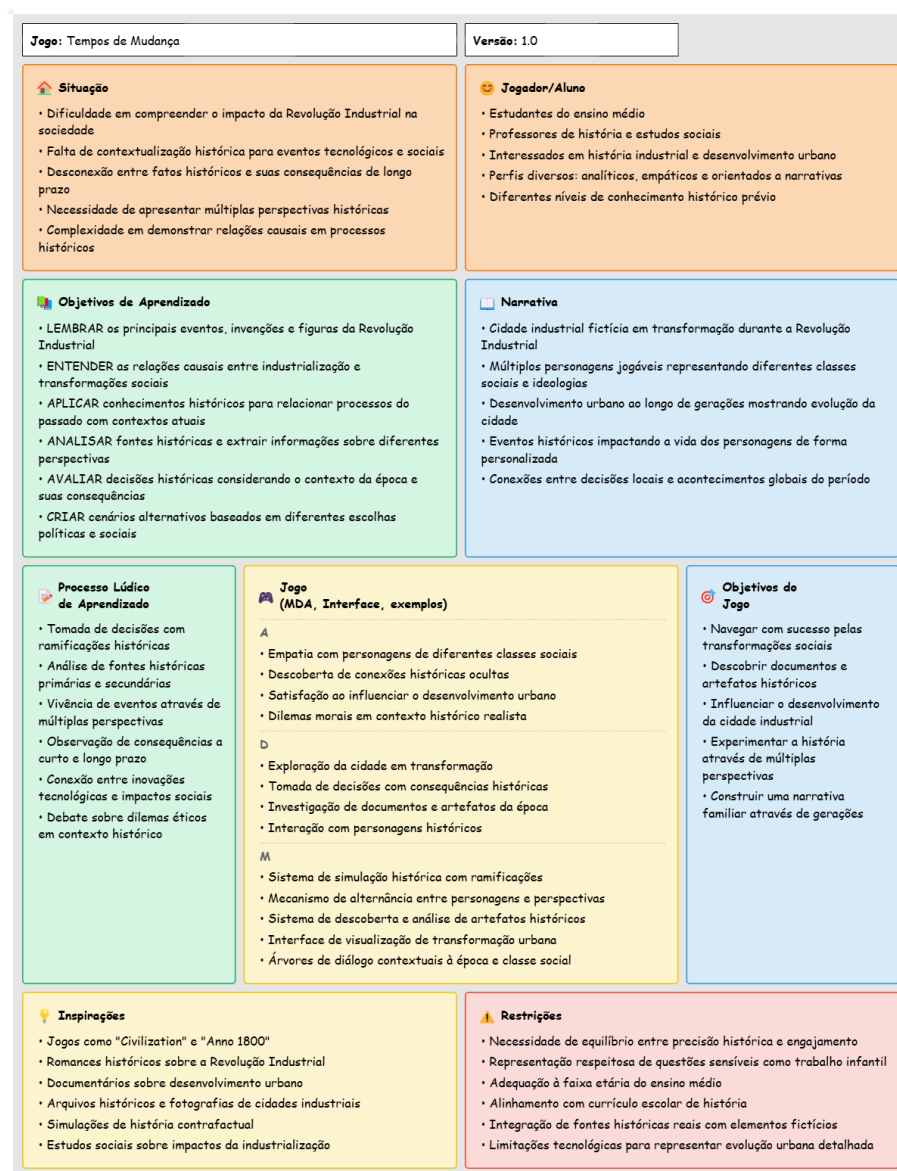


Figura A.5: Endo-GDC do jogo “Tempos de Mudança”

mostram consequências de decisões histórica e permitem reflexão sobre processos de causa e efeito em contextos sociais complexos.

A.6 Estudo de Caso 6: Ensino de Teoria Musical

A.6.1 Design de Jogo para Ensino de Teoria Musical para Iniciantes

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional voltado para o ensino de teoria musical básica.

Fase de Brainstorming: Durante a sessão de brainstorming, o usuário (um professor de música) inseriu as seguintes ideias iniciais (7 inputs):

- Jogo para ensino de notação musical
- Foco em reconhecimento de notas e intervalos
- Progressão desde conceitos básicos até formação de acordes
- Integração com instrumentos reais via MIDI
- Feedback auditivo imediato
- Elemento de composição criativa
- Abordagem inclusiva para diferentes estilos musicais

O sistema analisou estas ideias e sugeriu apenas dois conceitos complementares (2 sugestões):

- Narrativa em que o jogador é um aprendiz de um maestro buscando criar uma obra-prima
- Sistema de desafios colaborativos onde jogadores combinam suas criações

Fase de Método Socrático: O Agente Socrático aplicou um conjunto extenso de questões reflexivas (6 questões):

- “Como o jogo equilibrará o rigor técnico da teoria musical com a experiência lúdica?”
- “De que forma o jogo acomodará diferentes estilos de aprendizagem musical (visual, auditivo, cinestésico)?”
- “Como o feedback será estruturado para ser informativo sem ser desencorajador?”
- “Qual será a abordagem para integrar a criatividade musical com o aprendizado teórico?”
- “Como o jogo lidará com diferentes níveis de experiência musical prévia?”
- “De que forma a progressão de dificuldade será implementada para manter o engajamento?”

As respostas geraram os seguintes insights:

- Utilizar visualizações dinâmicas que conectam notação, som e posicionamento nos instrumentos
- Implementar um sistema de adaptação que ajusta desafios baseado no desempenho e estilo de aprendizagem
- Criar um ambiente de experimentação seguro onde erros são parte do processo criativo
- Desenvolver um sistema de composição guiada que gradualmente remove restrições

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "blocos musicais" que permite construção visual de composições, com feedback auditivo instantâneo e visualização de padrões teóricos.
- **Agente de Narrativa:** Desenvolveu uma história sobre um músico viajante que aprende diferentes estilos e tradições musicais ao redor do mundo, conectando teoria musical a contextos culturais.
- **Agente de Engajamento:** Sugeriu um sistema de desafios de improvisação com restrições progressivamente mais complexas, e um componente social para compartilhamento de criações.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar notas, intervalos, escalas e acordes básicos na notação musical.
- **Entender:** Explicar relações entre notas, formação de intervalos e construção de acordes.
- **Aplicar:** Utilizar conhecimentos teóricos para criar melodias e progressões harmônicas simples.
- **Analisar:** Decompor peças musicais para identificar estruturas e padrões teóricos.
- **Avaliar:** Julgar composições com base em princípios teóricos e expressividade musical.
- **Criar:** Desenvolver composições originais aplicando conceitos teóricos em contextos criativos.

Resultado Final - Endo-GDC do Jogo “Harmonix”: O Endo-GDC do jogo “Harmonix” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema interativo que conecta notação musical, reprodução sonora e visualizações, permitindo experimentação com feedback multimodal imediato.
- **Narrativa Integrada:** Jornada de um músico explorando diferentes tradições musicais, com personagens mentores representando diversos estilos e abordagens teóricas.
- **Elementos de Engajamento:** Sistema de composição progressiva que introduz novos elementos teóricos como ferramentas criativas, com oportunidades de personalização e compartilhamento.
- **Avaliação Incorporada:** Feedback contextual fornecido através de reações

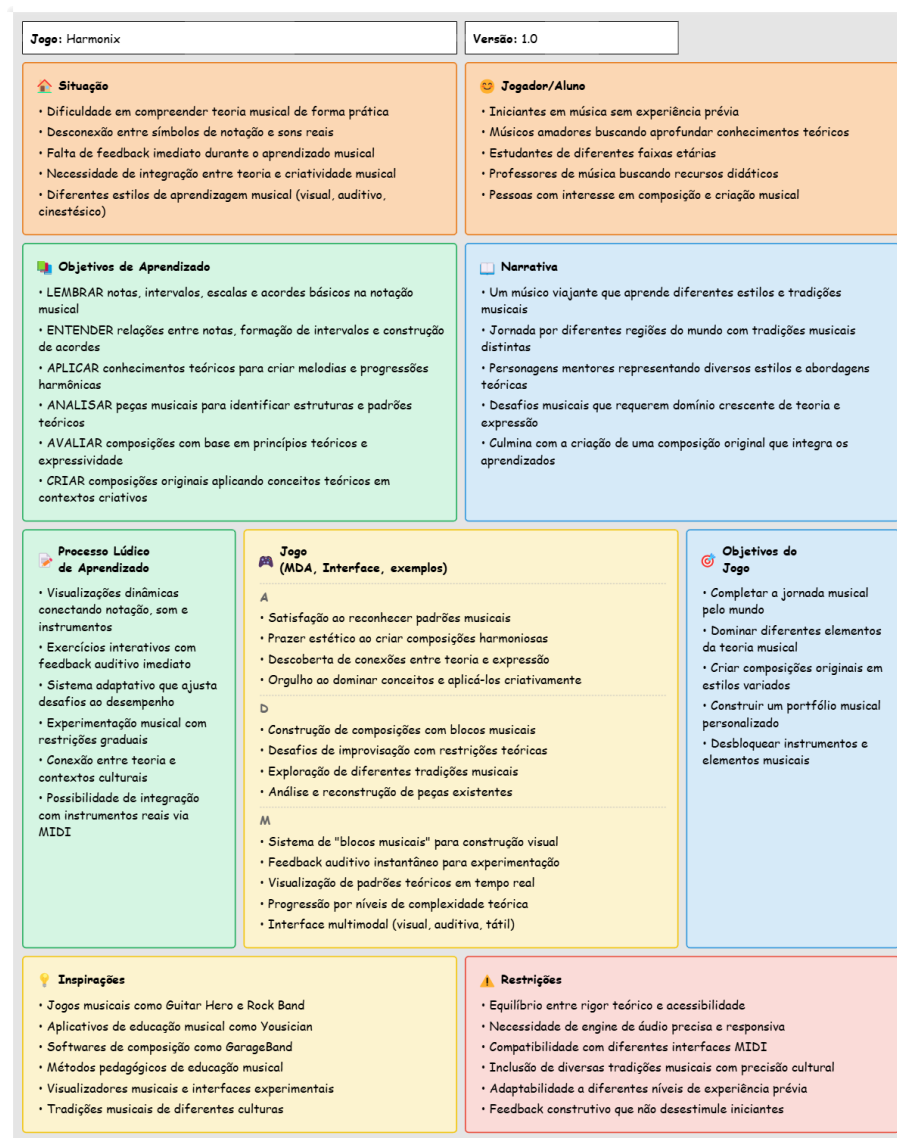


Figura A.6: Endo-GDC do jogo “Harmonix”

de personagens, audiências virtuais e análise visual de padrões musicais, enfatizando tanto a correção teórica quanto a expressividade.

A.7 Estudo de Caso 7: Ensino de Biologia Celular

A.7.1 Design de Jogo para Ensino de Biologia Celular no Ensino Médio

Este caso demonstra a aplicação do sistema para o design de um jogo educacional sobre biologia celular para estudantes do Ensino Médio.

Fase de Brainstorming: Um professor de biologia inseriu as seguintes ideias iniciais (3 inputs):

- Jogo sobre estrutura e funções celulares
- Visualização 3D de organelas e processos celulares
- Possibilidade de manipular componentes celulares

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (5 sugestões):

- Narrativa onde o jogador é um pesquisador em escala microscópica
- Simulação de diferentes tipos celulares (animal, vegetal, procarionte)
- Representação de processos dinâmicos como respiração celular e fotossíntese
- Progressão desde nível celular até sistemas de órgãos
- Cenários de solução de problemas como resposta a patógenos ou mutações

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo equilibrará precisão científica com simplificações necessárias para jogabilidade?”
- “De que forma a escala microscópica e o tempo dos processos celulares serão representados de maneira compreensível?”
- “Como as interações entre diferentes sistemas celulares serão demonstradas de forma clara?”
- “Que mecanismos permitirão aos estudantes fazer conexões entre fenômenos celulares e suas manifestações macroscópicas?”

As respostas do usuário geraram os seguintes insights:

- Utilizar camadas de complexidade que podem ser ativadas ou desativadas conforme o nível de aprendizado
- Implementar controles de escala e tempo para visualizar processos em diferentes níveis
- Criar visualizações que conectam eventos celulares a manifestações visíveis no organismo
- Desenvolver um glossário científico interativo que esclarece termos e conceitos durante o jogo

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de simulação celular interativa onde os jogadores manipulam componentes celulares e observam resultados, com foco em relações de causa e efeito entre organelas e processos.
- **Agente de Narrativa:** Desenvolveu uma narrativa em que o jogador é um especialista em nanotecnologia que pode reduzir-se à escala celular para investigar e resolver problemas biológicos em diferentes organismos.

- **Agente de Engajamento:** Sugeriu um sistema de desafios baseados em casos reais de pesquisa biológica, com diferentes níveis de complexidade e um componente de laboratório virtual para experimentação.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar estruturas celulares e suas funções básicas.
- **Entender:** Explicar como diferentes organelas contribuem para o funcionamento celular.
- **Aplicar:** Prever mudanças no funcionamento celular baseadas em alterações de condições ou componentes.
- **Analisar:** Comparar e contrastar diferentes tipos de células e seus sistemas especializados.
- **Avaliar:** Julgar a eficácia de respostas celulares a diferentes estímulos e ameaças.
- **Criar:** Projetar soluções para problemas celulares baseadas em princípios biológicos.

Resultado Final - Endo-GDC do Jogo “CellQuest”: O Endo-GDC do jogo “CellQuest” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Simulação interativa 3D que permite manipulação de componentes celulares e visualização de processos em tempo real, com mecanismos de causa e efeito que refletem relações biológicas reais.
- **Narrativa Integrada:** Missões científicas estruturadas como investigações biológicas em diferentes tipos celulares e organismos, com crescente complexidade e conexões entre níveis de organização biológica.
- **Elementos de Engajamento:** Sistema de descobertas progressivas que desbloqueiam novas ferramentas de investigação e cenários, com um componente de laboratório virtual para experimentação livre.
- **Avaliação Incorporada:** Feedback através de resultados experimentais visualizáveis e métricas de saúde celular/organismos, permitindo aos estudantes avaliar diretamente o impacto de suas intervenções.



Figura A.7: Endo-GDC do jogo “CellQuest”

A.8 Estudo de Caso 8: Ensino de Programação no Ensino Fundamental

A.8.1 Design de Jogo para Introdução à Programação para Crianças

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo introdutório de programação para crianças de 8-12 anos.

Fase de Brainstorming: Um especialista em educação tecnológica inseriu um único input inicial (1 input):

- Jogo para ensinar conceitos básicos de programação para crianças

O sistema, demonstrando sua capacidade de expansão a partir de inputs mínimos, analisou esta ideia e sugeriu um conjunto abrangente de conceitos complemen-

tares (10 sugestões):

- Programação visual por blocos sem necessidade de digitação
- Personagem controlável através de comandos de programação
- Progressão desde comandos simples até loops e condicionais
- Ambientação em um mundo fantástico com desafios temáticos
- Sistema de feedback visual imediato para cada ação
- Modo de jogo colaborativo para resolução de problemas em grupo
- Possibilidade de personalização do personagem e ambiente
- Integração de conceitos matemáticos básicos
- Narrativa que incorpora conceitos computacionais como metáforas
- Modo de criação onde alunos podem criar e compartilhar desafios

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode tornar conceitos abstratos de programação acessíveis para crianças?”
- “De que forma o jogo pode criar uma progressão natural que introduz conceitos mais complexos sem frustrar os jogadores?”
- “Como o jogo pode encorajar pensamento criativo e múltiplas soluções em vez de respostas ‘corretas’ únicas?”
- “Que mecanismos ajudarão crianças a depurar e compreender erros como parte natural do processo?”
- “Como o jogo pode demonstrar aplicações práticas da programação além do contexto do jogo?”

As respostas geraram os seguintes insights:

- Utilizar metáforas visuais concretas para representar conceitos abstratos
- Implementar um sistema de progressão adaptativo com múltiplos caminhos de aprendizado
- Criar desafios abertos com múltiplas soluções e sistema de recompensa para eficiência e criatividade
- Desenvolver um assistente visual de depuração que torna o processo divertido e informativo

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de programação visual baseado em blocos coloridos que representam comandos, com feedback visual imediato para cada linha de código e uma interface de execução passo-a-passo.
- **Agente de Narrativa:** Desenvolveu uma história sobre um mundo mágico

onde os habitantes usam "encantamentos"(código) para resolver problemas e criar novas possibilidades, com personagens representando diferentes paradigmas de programação.

- **Agente de Engajamento:** Sugeriu um sistema progressivo de desafios contextualizados em diferentes ambientes temáticos, com sistema de conquistas para incentivar exploração, experimentação e otimização.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Reconhecer comandos básicos de programação e sua sintaxe visual.
- **Entender:** Explicar como diferentes comandos afetam o comportamento do personagem e do ambiente.
- **Aplicar:** Utilizar sequências de comandos para resolver desafios específicos.
- **Analisar:** Decompor problemas complexos em etapas mais simples e identificar padrões repetitivos.
- **Avaliar:** Comparar diferentes soluções para o mesmo problema em termos de eficiência e elegância.
- **Criar:** Desenvolver programas originais para criar animações, histórias interativas ou jogos simples.

Resultado Final - Endo-GDC do Jogo “CodeQuest”: O Endo-GDC do jogo “CodeQuest” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de programação visual onde os blocos de código são tanto ferramentas de resolução de problemas quanto mecanismos de interação com o mundo, com execução visual que permite compreensão intuitiva de conceitos computacionais.
- **Narrativa Integrada:** Aventura em um mundo onde a programação é uma forma de magia que permite transformar o ambiente, com personagens que guiam o aprendizado e missões que contextualizam conceitos computacionais.
- **Elementos de Engajamento:** Sistema de desafios progressivos complementado por um modo de criação livre, com mecanismos de compartilhamento e colaboração para fomentar uma comunidade de aprendizado.
- **Avaliação Incorporada:** Feedback visual e interativo que demonstra o funcionamento de cada linha de código, com um assistente de depuração que transforma erros em oportunidades de aprendizado.

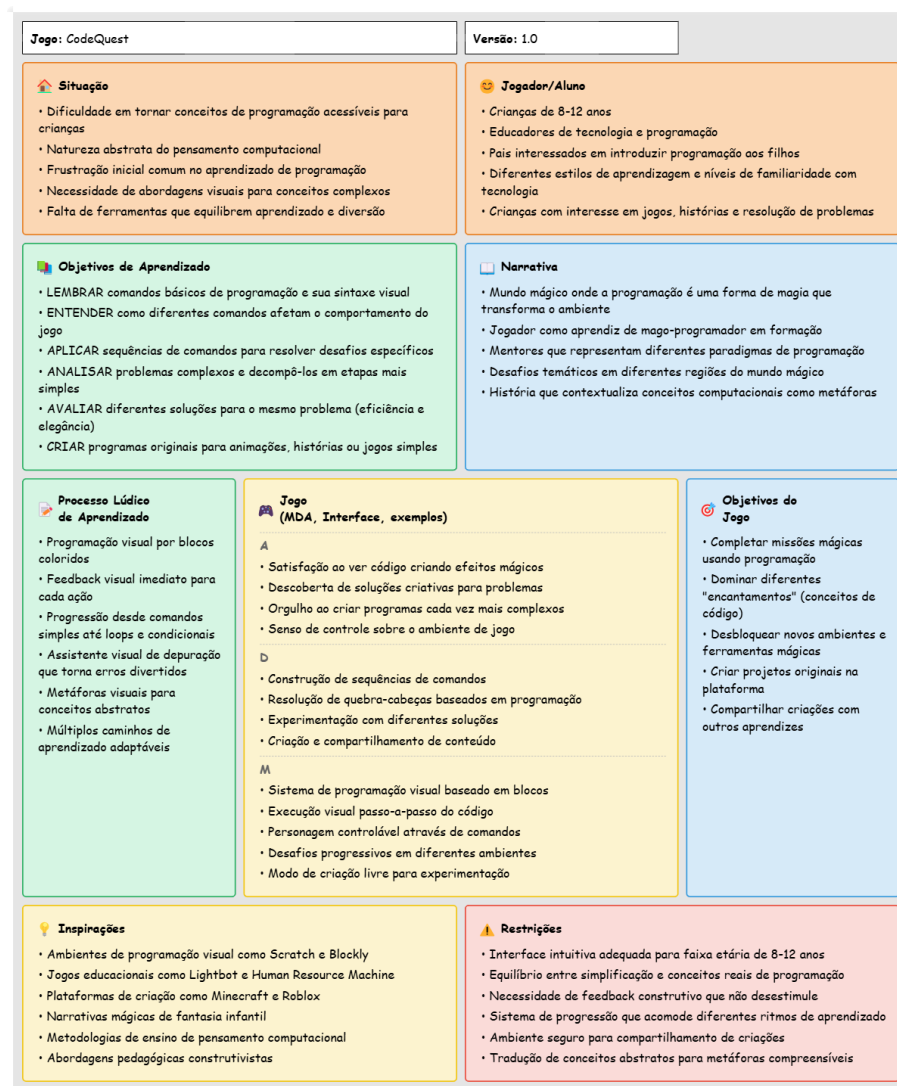


Figura A.8: Endo-GDC do jogo “CodeQuest”

A.9 Estudo de Caso 9: Aprendizado de Língua Estrangeira para Viajantes

A.9.1 Design de Jogo para Aprendizado de Idiomas em Contexto de Viagem

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo educacional voltado para adultos que desejam aprender um idioma para fins de viagem.

Fase de Brainstorming: Uma designer educacional inseriu as seguintes ideias iniciais (5 inputs):

- Jogo sobre aprendizado de idiomas em contextos de viagem
- Foco em situações práticas como restaurantes, transporte e hospedagem

- Progressão baseada em cenários de complexidade crescente
- Componente de reconhecimento e prática de pronúncia
- Adaptação a múltiplos idiomas-alvo

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (4 sugestões):

- Narrativa imersiva onde o jogador é um viajante em missões culturais
- Sistema de personagens nativos com personalidades distintas e sotaques diferentes
- Elementos de gamificação baseados em "conquistas de viagem"
- Componente de realidade aumentada para prática em ambientes reais

Fase de Método Socrático: O Agente Socrático aplicou questões reflexivas direcionadas ao contexto:

- “Como o jogo pode equilibrar vocabulário prático imediato com estruturas gramaticais necessárias para comunicação efetiva?”
- “De que forma o sistema pode adaptar-se a diferentes ritmos de aprendizado e estilos de comunicação?”
- “Como será abordada a diversidade cultural além do idioma em si?”
- “Quais estratégias serão utilizadas para incentivar prática regular e retenção a longo prazo?”

As respostas geraram os seguintes insights:

- Implementar um sistema de "construção de frases" que introduz gramática de forma contextual e prática
- Utilizar um modelo adaptativo que ajusta a dificuldade com base no desempenho em diferentes habilidades linguísticas
- Incorporar elementos culturais como parte integrante das missões e interações
- Desenvolver um sistema de revisão espaçada gamificado através de "flashbacks de viagem"

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de diálogos ramificados com múltiplas opções de resposta e consequências, combinado com um modo de simulação de conversação livre utilizando processamento de linguagem natural.
- **Agente de Narrativa:** Desenvolveu uma estrutura de "roteiro de viagem" personalizado onde o jogador embarca em uma jornada cultural com objetivos específicos, conhecendo personagens locais e descobrindo locais emblemáticos.
- **Agente de Engajamento:** Sugeriu um sistema de "passaporte virtual" que registra progresso, conquistas e memórias, complementado por desafios diários

contextualizados e um componente social para prática com outros jogadores.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Reconhecer vocabulário essencial e expressões comuns em contextos de viagem.
- **Entender:** Interpretar diálogos autênticos em situações típicas de viagem.
- **Aplicar:** Utilizar estruturas linguísticas aprendidas para comunicar necessidades e desejos em novos contextos.
- **Analisar:** Distinguir nuances culturais e pragmáticas na comunicação além do significado literal.
- **Avaliar:** Autoavaliar a própria comunicação e adaptar estratégias com base no feedback.
- **Criar:** Produzir diálogos originais adequados para diversos contextos culturais e situacionais.

Resultado Final - Endo-GDC do Jogo “Linguatrip”: O Endo-GDC do jogo “Linguatrip” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de simulação de conversação onde o idioma é simultaneamente o objeto de aprendizado e o mecanismo de interação com o mundo, com consequências narrativas baseadas na eficácia comunicativa.
- **Narrativa Integrada:** Jornada de viagem personalizada com missões culturais que contextualizam o uso do idioma em situações autênticas, proporcionando motivação intrínseca para o aprendizado.
- **Elementos de Engajamento:** Sistema de descoberta cultural e "coleção de memórias" que recompensa tanto a proficiência linguística quanto a exploração cultural, com componentes de revisão espaçada integrados à narrativa.
- **Avaliação Incorporada:** Feedback contextual através de reações dos personagens e sucessos/desafios nas missões, complementado por um sistema de autoavaliação reflexiva e análise de pronúncia.

A.10 Estudo de Caso 10: Ensino de Física no Ensino Médio

A.10.1 Design de Jogo para Ensino de Física Newtoniana

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo educacional sobre mecânica newtoniana para estudantes do Ensino Médio.

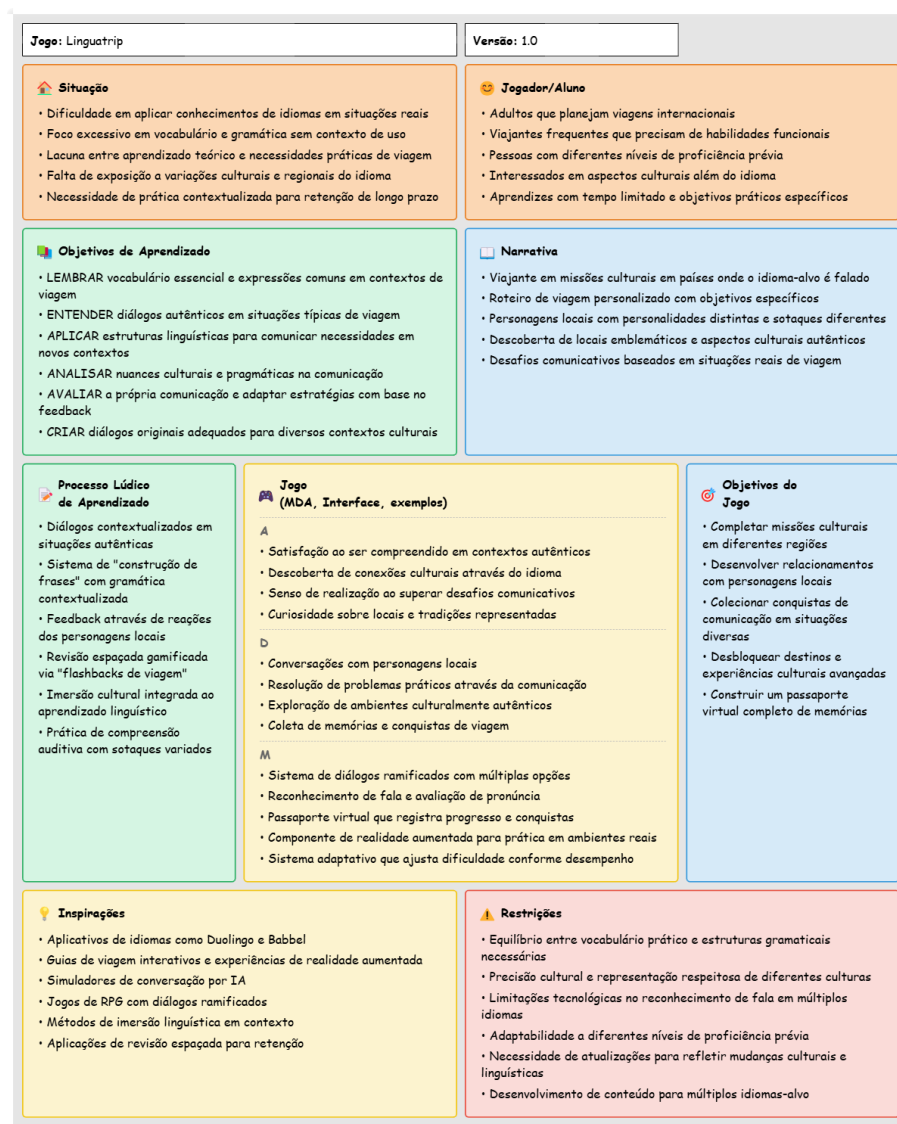


Figura A.9: Endo-GDC do jogo “Linguatrip”

Fase de Brainstorming: Um professor de física inseriu as seguintes ideias iniciais (4 inputs):

- Jogo de simulação física com visualização de forças
- Desafios baseados em problemas do mundo real
- Sistema que permite experimentação e teste de hipóteses
- Progressão desde mecânica básica até movimento complexo

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (6 sugestões):

- Ambiente de construção de máquinas e dispositivos
- Narrativa centrada em missões de engenharia espacial
- Representação visual de vetores e trajetórias em tempo real
- Sistema de previsão que permite aos estudantes prever resultados antes da simulação

- Modo de desafio onde estudantes projetam sistemas com restrições específicas
- Componente histórico que conecta descobertas científicas a seus contextos

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode evidenciar relações matemáticas sem intimidar estudantes com menos aptidão numérica?”
- “De que forma as simplificações necessárias para a jogabilidade serão comunicadas claramente?”
- “Como o jogo pode incentivar pensamento físico intuitivo além da aplicação de fórmulas?”
- “Quais estratégias serão usadas para conectar fenômenos microscópicos a observações macroscópicas?”
- “Como o jogo abordará concepções prévias incorretas comuns em física?”

As respostas geraram os seguintes insights:

- Implementar visualizações dinâmicas que demonstram relações matemáticas graficamente, com opções para explorar fórmulas para estudantes interessados
- Criar um sistema de "níveis de realismo" que introduz gradualmente complexidades como atrito e resistência do ar
- Desenvolver cenários que deliberadamente confrontam concepções incorretas comuns, exigindo revisão de hipóteses
- Utilizar narrativa para contextualizar problemas físicos em situações significativas

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de construção física baseado em componentes modulares com propriedades realistas, complementado por ferramentas de análise e visualização que revelam forças, energias e trajetórias.
- **Agente de Narrativa:** Desenvolveu uma história ambientada em uma agência espacial do futuro, onde o jogador deve projetar e testar tecnologias para missões planetárias, resolvendo problemas físicos com consequências narrativas.
- **Agente de Engajamento:** Sugeriu um sistema progressivo de desafios com múltiplas soluções possíveis, competições periódicas de design, e um componente de compartilhamento onde estudantes podem testar criações uns dos outros.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Reconhecer as leis de Newton e conceitos fundamentais da mecânica.
- **Entender:** Explicar relações entre forças, massa, aceleração e movimento em diferentes contextos.
- **Aplicar:** Utilizar princípios físicos para resolver problemas práticos de engenharia e design.
- **Analisar:** Decompor sistemas complexos em componentes e forças constitutivas.
- **Avaliar:** Julgar a eficiência e viabilidade de diferentes soluções para problemas físicos.
- **Criar:** Desenvolver dispositivos e mecanismos originais que aplicam princípios físicos para resolver desafios específicos.

Resultado Final - Endo-GDC do Jogo “PhysicsLab”: O Endo-GDC do jogo “PhysicsLab” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de simulação física precisa onde as leis físicas são tanto o objeto de estudo quanto os mecanismos de interação, permitindo experimentação, previsão e verificação em um ciclo científico iterativo.
- **Narrativa Integrada:** Campanha de exploração espacial que contextualiza desafios físicos em missões com propósito e consequências, conectando conceitos abstratos a aplicações concretas.
- **Elementos de Engajamento:** Sistema de design e construção livre combinado com desafios estruturados, permitindo tanto expressão criativa quanto aplicação direcionada de conhecimentos físicos.
- **Avaliação Incorporada:** Feedback visual e quantitativo automático baseado no comportamento dos sistemas físicos criados, complementado por análises de eficiência e comparação com soluções ideais.

A.11 Estudo de Caso 11: Educação Financeira para Jovens Adultos

A.11.1 Design de Jogo para Desenvolvimento de Literacia Financeira

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional voltado para o ensino de conceitos financeiros para jovens adultos.

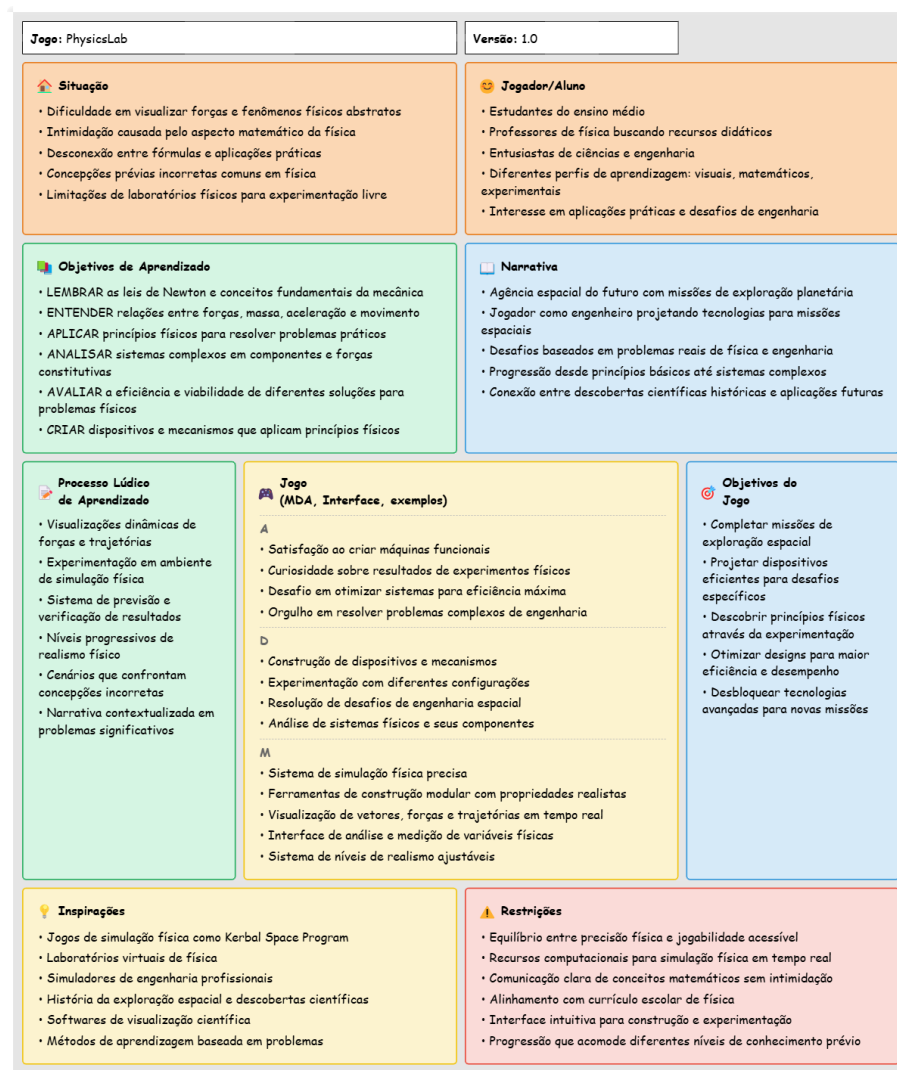


Figura A.10: Endo-GDC do jogo “PhysicsLab”

Fase de Brainstorming: Um especialista em educação financeira inseriu apenas um único input inicial (1 input):

• Jogo de simulação financeira para jovens adultos iniciando vida independente

O sistema, demonstrando sua capacidade de expansão a partir de inputs mínimos, gerou as seguintes sugestões (12 sugestões):

- Simulação de vida com ciclos mensais de receitas e despesas
- Sistema de gestão orçamentária com categorias personalizáveis
- Mecânicas de investimento com diferentes perfis de risco e retorno
- Eventos aleatórios que simulam emergências financeiras
- Representação de impactos de longo prazo de decisões financeiras
- Sistema de crédito e empréstimo com taxas de juros realistas
- Simulação de mercado de trabalho com opções de carreira e educação
- Cenários de decisão sobre moradia (alugar vs. comprar)
- Mecânicas de economia colaborativa e decisões de consumo

- Sistema de metas financeiras de curto, médio e longo prazo
- Representação de impostos e obrigações fiscais
- Tutorial integrado com conceitos financeiros básicos

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode tornar visíveis conceitos financeiros abstratos como juros compostos ou inflação?”
- “De que forma o jogo equilibrará realismo financeiro com uma experiência acessível e motivadora?”
- “Como serão representadas as desigualdades financeiras estruturais sem reforçar estereótipos?”
- “Qual abordagem será utilizada para incentivar comportamentos financeiros saudáveis sem ser prescritivo?”
- “Como o jogo pode facilitar a transferência de conhecimentos para situações financeiras reais?”

As respostas geraram os seguintes insights:

- Utilizar visualizações dinâmicas e projeções futuras para conceitos abstratos como juros compostos
- Implementar diferentes níveis de dificuldade e pontos de partida que refletem diversas realidades financeiras
- Focar em princípios e habilidades transferíveis em vez de produtos financeiros específicos
- Criar um sistema de reflexão que incentiva os jogadores a analisar suas próprias decisões financeiras

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de simulação financeira em tempo acelerado onde cada decisão tem consequências de curto e longo prazo, com visualizações dinâmicas de fluxo de caixa e patrimônio líquido.
- **Agente de Narrativa:** Desenvolveu uma estrutura narrativa personalizada onde o jogador define objetivos de vida (carreira, família, moradia) que contextualizam as decisões financeiras em um horizonte de vários anos.
- **Agente de Engajamento:** Sugeriu um sistema de conquistas alinhado com marcos financeiros saudáveis, complementado por desafios situacionais e um componente social onde jogadores podem compartilhar estratégias.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar conceitos financeiros básicos como juros, investimentos e tipos de despesas.
- **Entender:** Explicar como diferentes decisões financeiras afetam o fluxo de caixa e patrimônio ao longo do tempo.
- **Aplicar:** Utilizar ferramentas de planejamento financeiro para estabelecer e seguir um orçamento.
- **Analisar:** Avaliar diferentes opções de investimento, crédito e consumo considerando riscos e benefícios.
- **Avaliar:** Julgar a adequação de produtos financeiros a objetivos pessoais específicos.
- **Criar:** Desenvolver estratégias financeiras personalizadas alinhadas com valores e metas de vida.

Resultado Final - Endo-GDC do Jogo “FinLife”: O Endo-GDC do jogo “FinLife” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de simulação financeira onde conceitos como orçamento, investimento e gestão de crédito são incorporados como mecânicas centrais de jogo, com feedback visual e numérico que torna tangíveis conceitos financeiros abstratos.
- **Narrativa Integrada:** Jornada de vida personalizada onde decisões financeiras são contextualizadas em objetivos pessoais e valores, com eventos narrativos que respondem dinamicamente às escolhas financeiras do jogador.
- **Elementos de Engajamento:** Sistema de conquistas e desafios que recompensa tanto comportamentos financeiros saudáveis quanto a exploração de diferentes estratégias, com um componente de "vida alternativa" que permite experimentar caminhos não escolhidos.
- **Avaliação Incorporada:** Feedback multidimensional que vai além de métricas puramente financeiras para incluir bem-estar, segurança e alinhamento com valores, complementado por ferramentas de reflexão e planejamento.



Figura A.11: Endo-GDC do jogo “FinLife”

A.12 Estudo de Caso 12: Ensino de Geografia e Estudos Culturais

A.12.1 Design de Jogo para Exploração Geográfica e Cultural

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo educacional sobre geografia mundial e diversidade cultural.

Fase de Brainstorming: Uma professora de geografia e estudos sociais inseriu as seguintes ideias iniciais (3 inputs):

- Jogo de exploração geográfica com mapa mundial interativo
- Foco em diversidade cultural e sistemas sociais
- Perspectiva de sustentabilidade e interconexão global

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (5 sugestões):

- Personagens guias locais representando diferentes regiões e culturas
- Sistema de missões baseadas em desafios globais contemporâneos
- Mecânica de curadoria de artefatos culturais e conhecimentos tradicionais
- Representação de ecossistemas e adaptações humanas a diferentes biomas
- Visualização de dados geográficos, demográficos e econômicos em camadas

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode representar culturas de forma autêntica e respeitosa, evitando estereótipos?”
- “De que forma será equilibrada a representação de questões contemporâneas sensíveis sem simplificação excessiva?”
- “Como o jogo pode promover empatia cultural sem ‘exotizar’ comunidades diferentes?”
- “Qual abordagem será utilizada para integrar perspectivas indígenas e conhecimentos tradicionais?”
- “Como serão representadas continuidades históricas e processos de mudança cultural?”

As respostas geraram os seguintes insights:

- Colaborar com consultores culturais de diversas regiões para garantir representações autênticas
- Implementar múltiplas perspectivas sobre questões complexas, evitando visões monolíticas
- Focar em experiências humanas universais e valores compartilhados, destacando diversidade de expressões
- Incorporar narrativas em primeira pessoa e histórias orais de membros das próprias comunidades

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de exploração geográfica com mapa multicamadas que revela informações físicas, ecológicas, culturais e econômicas, complementado por um atlas interativo que se expande com descobertas.

- **Agente de Narrativa:** Desenvolveu uma estrutura narrativa não-linear centrada em "encontros culturais" conduzidos por personagens locais, com histórias pessoais que humanizam questões globais e revelam conexões interculturais.
- **Agente de Engajamento:** Sugeriu um sistema de missões baseadas em Objetivos de Desenvolvimento Sustentável da ONU, complementado por desafios de compreensão cultural e um componente de "embaixador cultural" onde jogadores compartilham aprendizados.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar características geográficas, países, culturas e tradições de diversas regiões do mundo.
- **Entender:** Explicar como fatores geográficos influenciam o desenvolvimento cultural e sistemas sociais.
- **Aplicar:** Relacionar conhecimentos geográficos e culturais para analisar questões contemporâneas globais.
- **Analisar:** Comparar diferentes perspectivas culturais sobre questões comuns e examinar interconexões globais.
- **Avaliar:** Considerar criticamente narrativas sobre diferentes culturas e regiões, reconhecendo vieses e complexidades.
- **Criar:** Desenvolver propostas para desafios globais que respeitem diversidade cultural e conhecimentos locais.

Resultado Final - Endo-GDC do Jogo “GeoVentures”: O Endo-GDC do jogo “GeoVentures” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de exploração geográfica e cultural onde o conhecimento sobre lugares e pessoas se traduz diretamente em capacidade de interação e conexão, com descobertas que expandem possibilidades de navegação no mundo do jogo.
- **Narrativa Integrada:** Encontros culturais conduzidos por personagens locais que compartilham histórias pessoais, tradições e perspectivas, contextualizando questões geográficas e sociais em experiências humanas autênticas.
- **Elementos de Engajamento:** Sistema de missões baseadas em desafios globais reais que incentivam pesquisa, compreensão cultural e pensamento crítico, com um componente de compartilhamento e colaboração entre jogadores.
- **Avaliação Incorporada:** Feedback multidimensional através de um "diário de viagem" que registra descobertas, reflexões e conexões estabelecidas, complementado por um sistema de "compreensão cultural" que evolui com interações respeitosas.

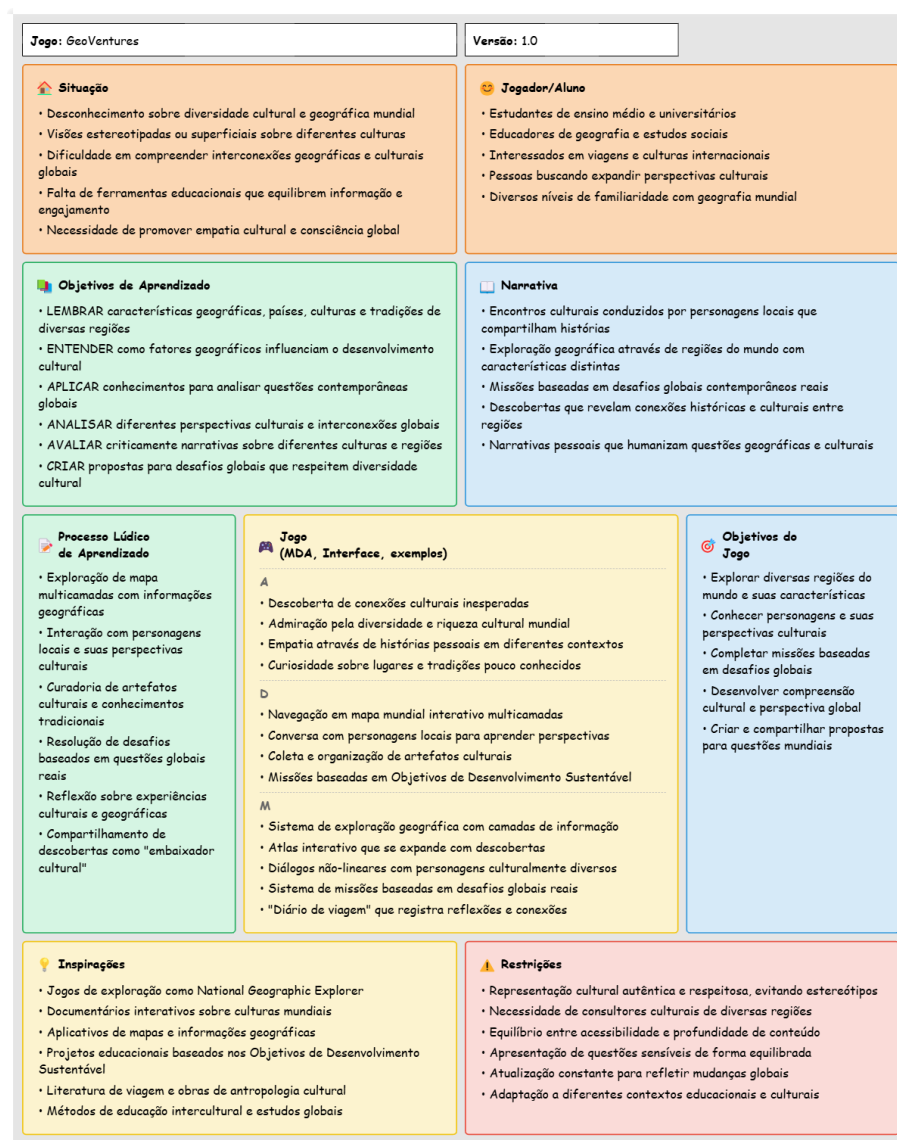


Figura A.12: Endo-GDC do jogo “GeoVentures”

A.13 Estudo de Caso 13: Educação Ambiental para Ensino Fundamental II

A.13.1 Design de Jogo para Conscientização Ambiental e Sustentabilidade

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional sobre ecologia e sustentabilidade para estudantes do Ensino Fundamental II.

Fase de Brainstorming: Uma professora de ciências inseriu as seguintes ideias iniciais (6 inputs):

- Jogo de gestão de ecossistema com simulação de cadeia alimentar
- Representação de impactos ambientais de atividades humanas
- Mecânicas de tomada de decisão com consequências visíveis
- Perspectiva de longo prazo para demonstrar mudanças graduais
- Abordagem baseada em soluções, não apenas problemas
- Contexto local que conecta com questões globais

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (3 sugestões):

- Personagens mentores representando diferentes perspectivas ambientais
- Sistema de "diários de campo" que registram observações e descobertas
- Componente de ação comunitária que conecta o jogo a projetos reais

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode representar complexidades ecológicas de forma acessível sem simplificação excessiva?”
- “De que forma serão equilibradas mensagens sobre urgência ambiental e empoderamento para ação?”
- “Como o jogo evitará política partidária enquanto aborda questões ambientais frequentemente politizadas?”
- “Qual abordagem será utilizada para conectar comportamentos individuais a impactos sistêmicos?”
- “Como o jogo promoverá pensamento crítico sobre soluções, reconhecendo que há compromissos em qualquer intervenção?”

As respostas geraram os seguintes insights:

- Implementar níveis de complexidade que podem ser ativados progressivamente, mantendo sistemas centrais visíveis em todos os níveis
- Equilibrar representações de desafios ambientais com oportunidades concretas para ação e experimentação de soluções
- Focar em evidências científicas e apresentar múltiplas perspectivas sobre questões complexas
- Criar visualizações que conectam ações individuais a impactos coletivos através de "efeito borboleta" ecológico

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de simulação ecossistêmica baseado em relações interconectadas entre espécies, habitat e recursos, com indicadores visuais de saúde ambiental e feedback em múltiplas escalas temporais.

- **Agente de Narrativa:** Desenvolveu uma estrutura narrativa em que os jogadores assumem o papel de "guardiões ecológicos" responsáveis por uma área natural ao longo de gerações, enfrentando desafios ambientais e buscando equilíbrio entre necessidades humanas e ecológicas.
- **Agente de Engajamento:** Sugeriu um sistema de missões baseado em desafios ecológicos reais, complementado por um "laboratório de soluções" para experimentação e um componente comunitário que conecta ações no jogo a iniciativas locais.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar componentes-chave de ecossistemas e seus papéis na manutenção do equilíbrio ambiental.
- **Entender:** Explicar como diferentes elementos de um ecossistema interagem e como atividades humanas impactam estes sistemas.
- **Aplicar:** Implementar estratégias de gestão ambiental para resolver problemas ecológicos específicos.
- **Analisar:** Examinar relações de causa e efeito em sistemas ecológicos e prever consequências de intervenções.
- **Avaliar:** Avaliar diferentes abordagens para questões ambientais com base em critérios de sustentabilidade, viabilidade e impacto.
- **Criar:** Desenvolver soluções inovadoras para desafios ambientais que considerem necessidades ecológicas e humanas.

Resultado Final - Endo-GDC do Jogo “EcoGuardiões”: O Endo-GDC do jogo “EcoGuardiões” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de simulação ecológica onde princípios de interdependência, ciclos naturais e resiliência são incorporados como mecânicas centrais, com intervenções que produzem efeitos visíveis em diferentes escalas espaciais e temporais.
- **Narrativa Integrada:** Jornada de gerações de guardiões ecológicos que enfrentam desafios ambientais em evolução, contextualizando conceitos científicos em histórias pessoais e comunitárias com significado emocional.
- **Elementos de Engajamento:** Sistema de missões baseadas em problemas ecológicos reais, complementado por um laboratório de experimentação e um componente de ação comunitária que conecta aprendizado virtual com iniciativas locais concretas.
- **Avaliação Incorporada:** Feedback visual através de indicadores de saúde ecossistêmica e bem-estar comunitário, complementado por um "diário de

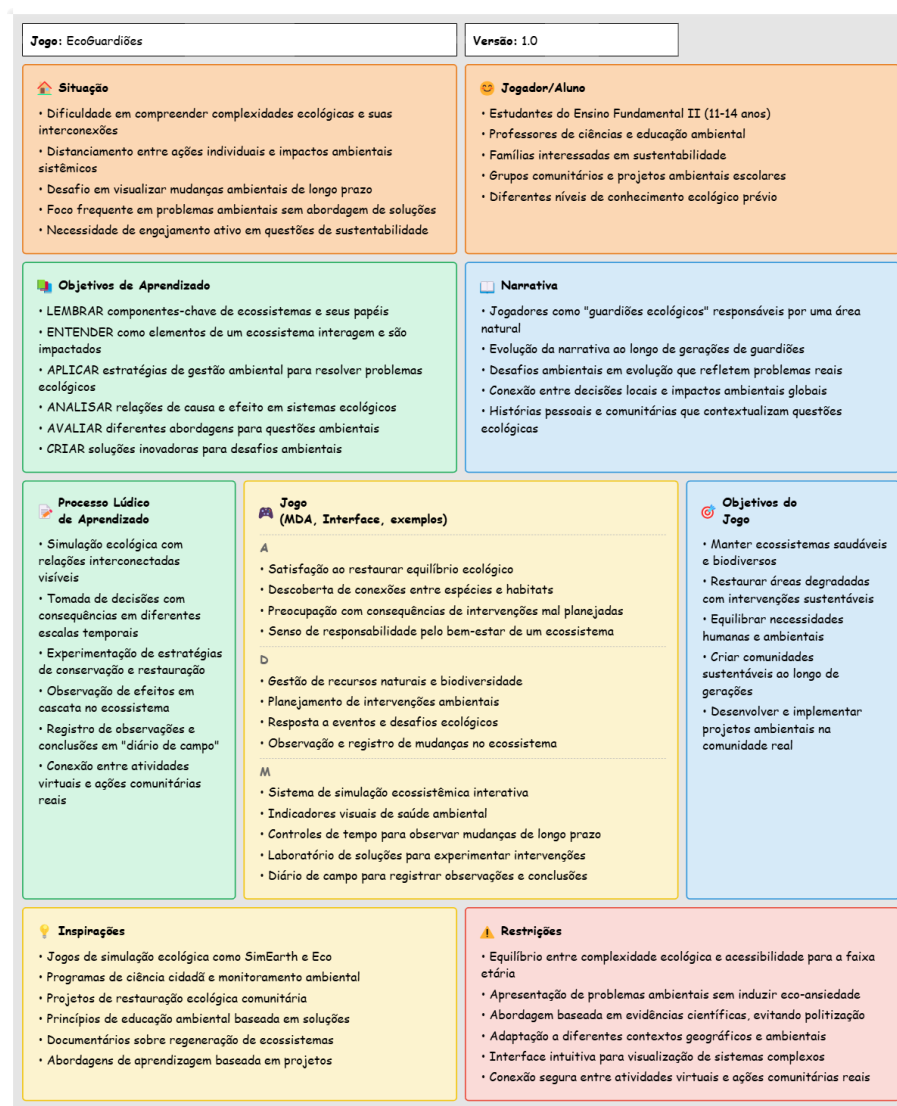


Figura A.13: Endo-GDC do jogo “EcoGuardiões”

campo"que registra observações, hipóteses e conclusões científicas do jogador.

A.14 Estudo de Caso 14: Ensino de Geografia no Ensino Fundamental II

A.14.1 Design de Jogo para Ensino de Geografia no Ensino Fundamental II

Este estudo de caso apresenta a aplicação do sistema para o design de um jogo educativo sobre mapear regiões e entender fenômenos físicos e humanos em Geografia.

Fase de Brainstorming: Uma professora de Geografia inseriu as seguintes ideias iniciais (5 inputs):

- Jogo de exploração de mapas e biomas
- Foco em relevo, hidrografia e clima
- Público-alvo: estudantes de 1114 anos
- Ferramenta para criar e customizar mapas interativos
- Missões relacionadas a desafios de planejamento urbano e meio ambiente

O sistema sugeriu conceitos complementares (5 sugestões):

- Mecânica de coleção de dados usando sondas virtuais em diferentes ambientes
- Eventos baseados em desastres naturais e mudanças climáticas
- Integração de fontes de dados reais (imagens de satélite simplificadas)
- Desafios de logística para conectar cidades por rodovias e ferrovias
- Sistema de indicadores socioeconômicos e ambientais para tomada de decisão

Fase de Método Socrático: O Agente Socrático aplicou as questões reflexivas padronizadas:

- Como será a interface de criação de mapas? Haverá camadas de informação?
- De que forma o jogo incorpora corretamente a escala e proporção geográfica?
- Como garantir que o jogador compreenda a relação entre relevo, clima e vegetação?
- Quais mecanismos permitem comparações entre diferentes regiões do mundo?
- Como evitar simplificações excessivas que deturpem conceitos essenciais?

Fase de Insights do Usuário: As respostas foram refinadas em insights como:

- Incluir tutoriais interativos sobre leitura de curvas de nível e legenda de mapas
- Utilizar missões que exigem planejamento de rotas e análise de riscos de desastres
- Implementar feedback visual imediato em camadas de clima e uso do solo
- Adicionar questionários rápidos entre fases para reforçar conceitos-chave

Fase de Endo-GDC:

- **Agente de Mecânicas:** Sugeriu um sistema de marcação de waypoints e coleta de amostras, onde cada coleta requer aplicação de conceitos de hidrografia e clima.
- **Agente de Narrativa:** Criou uma história de expedição científica em diferentes biomas, com personagens especialistas em geociências.
- **Agente de Engajamento:** Propôs conquistas de explorador vinculadas a descobertas de novos territórios e resolução de enigmas geográficos.

Fase de Taxonomia de Bloom:

- **Lembrar:** Reconhecer tipos de relevo, rios e zonas climáticas.

- **Entender:** Explicar como fatores físicos influenciam a distribuição de biomas.
- **Aplicar:** Utilizar mapas para planejar rotas considerando elevação e condições climáticas.
- **Analisar:** Comparar indicadores socioambientais de diferentes regiões.
- **Avaliar:** Julgar propostas de desenvolvimento urbano frente a riscos naturais.
- **Criar:** Projetar um mapa temático personalizado com legendas e camadas informativas.

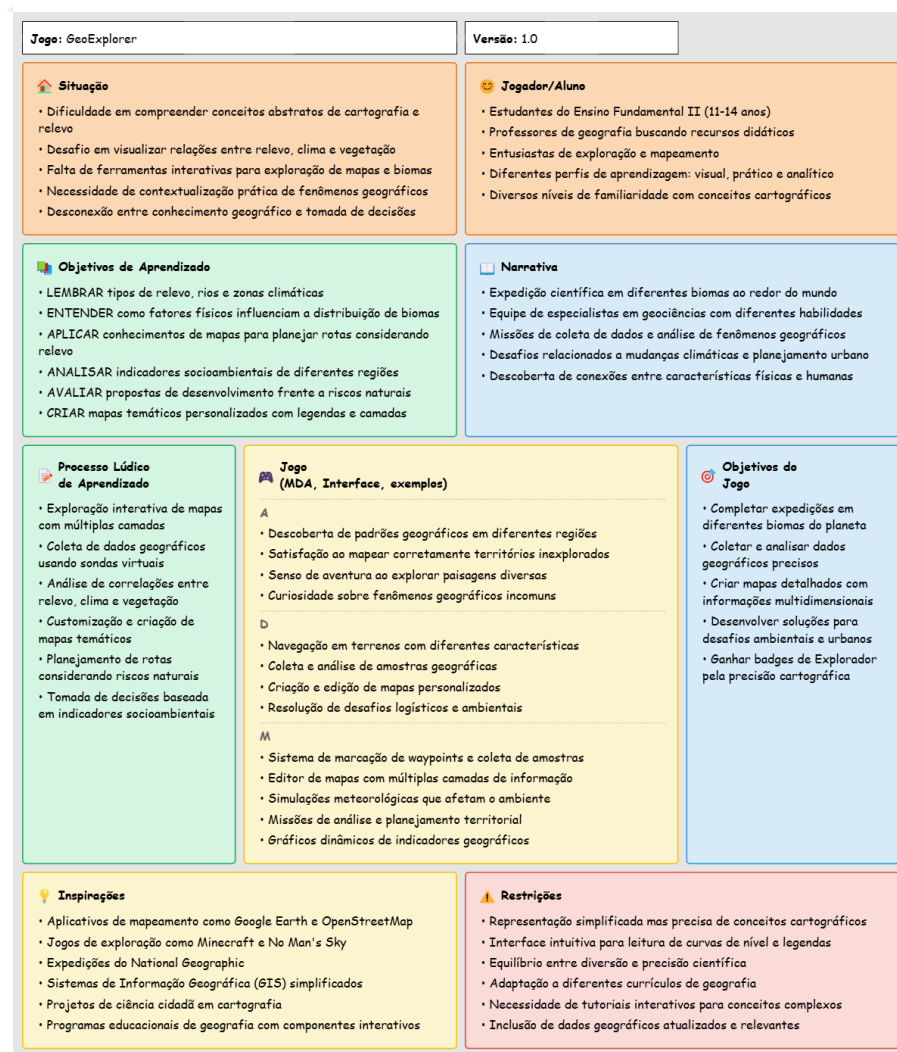


Figura A.14: Endo-GDC do jogo “GeoExplorer”

Resultado Final Endo-GDC do Jogo “GeoExplorer”: O Endo-GDC do jogo “GeoExplorer” integra:

- **Mecânicas Endógenas:** Coleta de dados geográficos, edição de mapas e simulações meteorológicas que afetam o ambiente de jogo.
- **Narrativa Integrada:** Expedição de cientistas em biomas distintos, com missões ligadas à preservação ambiental.

- **Elementos de Engajamento:** Sistema de badges de Explorador e Desafiante que recompensa análise crítica e precisão cartográfica.
- **Avaliação Incorporada:** Gráficos dinâmicos de indicadores e relatórios automáticos que permitem ao aluno autoavaliar sua compreensão.

A.15 Estudo de Caso 15: Treinamento de Atendimento ao Cliente em Ambiente Corporativo

A.15.1 Design de Jogo para Formação de Soft Skills em Empresas

Este estudo de caso ilustra a aplicação do sistema em um jogo para desenvolver habilidades de atendimento ao cliente e comunicação.

Fase de Brainstorming: Um analista em RH inseriu as seguintes ideias iniciais (4 inputs):

- Simulação de cenários de atendimento ao cliente (telefone, chat, presencial)
- Foco em empatia, escuta ativa e resolução de conflitos
- Sistema de feedback em tempo real baseado em escolha de diálogo
- Missões moduladas por diferentes perfis de clientes (satisfeitos, insatisfeitos, técnicos)

O sistema sugeriu conceitos complementares (6 sugestões):

- Módulo de role play com personagens virtuais que expressam emoções faciais
- Indicadores de performance em KPIs como tempo de resposta e satisfação do cliente
- Eventos aleatórios como picos de demanda e crises de imagem
- Ferramenta de análise de sentimento do texto digitado pelo jogador
- Componentes de gamificação: pontos, níveis e rankings internos
- Relatórios de desempenho individual e comparativo por equipe

Fase de Método Socrático: O Agente Socrático aplicou as questões padronizadas:

- Como garantir que as escolhas de diálogo reflitam práticas de atendimento realistas?
- De que forma será medido o equilíbrio entre rapidez e qualidade no atendimento?
- Que mecanismos assegurarão autenticidade nas reações dos personagens virtuais?

- Como o jogo lida com diferentes perfis de aprendizagem e habilidades interpessoais?
- Quais recursos ajudarão na transferência de habilidades do virtual para o mundo real?

Fase de Insights do Usuário: As respostas geraram insights como:

- Incluir um tutorial inicial sobre técnicas de pergunta aberta e fechada
- Utilizar feedback em múltiplos formatos: texto, voz sintetizada e gravações de resposta ideal
- Permitir revisões de cenários com replay e análise de melhores práticas
- Adotar mentor virtual que sugira melhorias após cada interação

Fase de Endo-GDC:

- **Agente de Mecânicas:** Propôs um sistema de árvore de diálogo com métricas de empatia e assertividade que modulam o roteiro.
- **Agente de Narrativa:** Criou uma empresa fictícia de tecnologia que enfrenta um problema crítico de retenção de clientes.
- **Agente de Engajamento:** Sugeriu desafios semanais de melhoria contínua e reconhecimento interno entre colegas.

Fase de Taxonomia de Bloom:

- **Lembrar:** Identificar boas práticas de atendimento e protocolos de comunicação.
- **Entender:** Explicar a importância de escuta ativa e rapport no diálogo com o cliente.
- **Aplicar:** Utilizar técnicas de questionamento e reflexão para resolver reclamações.
- **Analisar:** Avaliar interações para identificar pontos de melhoria em empatia e clareza.
- **Avaliar:** Julgar a eficácia de diferentes estratégias de comunicação em situações-teste.
- **Criar:** Formular roteiros de atendimento personalizados para perfis variados de clientes.

Resultado Final Endo-GDC do Jogo “ClientCare Pro”: O Endo-GDC do jogo “ClientCare Pro” integra:

- **Mecânicas Endógenas:** Árvore de diálogo dinâmica, análise de sentimento e indicadores de empatia que afetam o desfecho de cada caso.

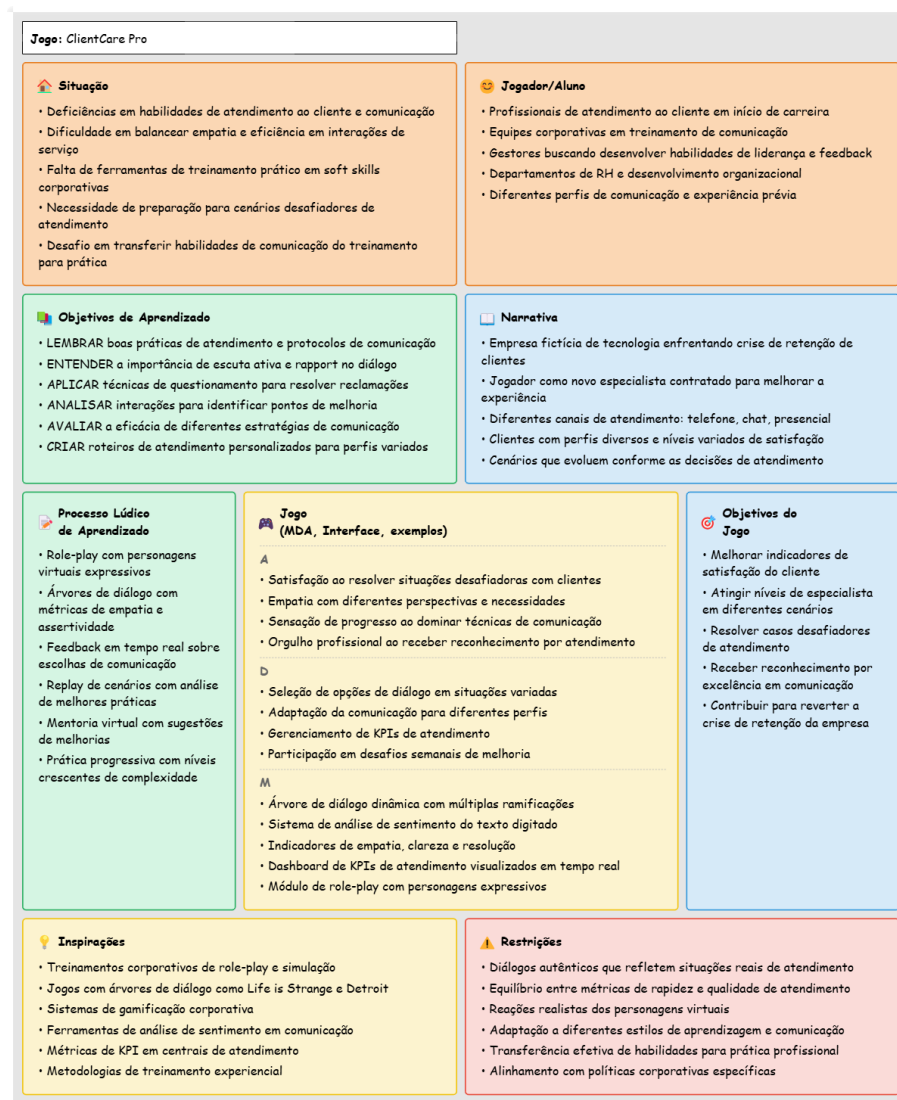


Figura A.15: Endo-GDC do jogo “ClientCare Pro”

- **Narrativa Integrada:** Empresa em crise de reputação, com cenários que evoluem conforme as decisões de atendimento.
- **Elementos de Engajamento:** Gamificação avançada com níveis de especialista, dashboards de equipe e competições internas.
- **Avaliação Incorporada:** Feedback instantâneo e relatórios comparativos que permitem reflexão e melhoria contínua das habilidades.

A.16 Estudo de Caso 16: Ensino de Química no Ensino Médio

A.16.1 Design de Jogo para Aprendizado de Química Interativa

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional voltado para o ensino de química e reações químicas para estudantes do Ensino Médio.

Fase de Brainstorming: Um professor de química inseriu as seguintes ideias iniciais (3 inputs):

- Laboratório virtual para experimentação segura de reações químicas
- Visualização de processos moleculares e atômicos em nível microscópico
- Sistema de desafios baseados em aplicações práticas da química

O sistema, através do serviço GroqService, analisou estas ideias e sugeriu conceitos complementares (6 sugestões):

- Narrativa de um cientista em formação trabalhando em um centro de pesquisa
- Mecânica de combinação de elementos para descobrir novas reações
- Sistema de previsão de resultados baseado em propriedades químicas
- Representação visual de energia e ligações durante reações
- Integração de conceitos teóricos com aplicações na indústria e cotidiano
- Componente de historiografia científica mostrando evolução das descobertas

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como equilibrar o rigor científico com a acessibilidade para estudantes com diferentes níveis de conhecimento?”
- “De que forma o jogo pode demonstrar a relação entre estrutura molecular e propriedades macroscópicas?”
- “Como incentivar a compreensão conceitual em vez da simples memorização de fórmulas e reações?”
- “Qual abordagem será utilizada para tornar visíveis processos químicos abstratos ou difíceis de observar?”
- “Como o jogo pode promover segurança laboratorial sem criar aversão à experimentação prática?”

As respostas geraram os seguintes insights:

- Implementar diferentes níveis de complexidade que introduzem gradualmente novos conceitos

- Utilizar visualizações multi-escala que conectam comportamento molecular a efeitos observáveis
- Criar desafios que requerem aplicação de princípios em vez de replicação de procedimentos
- Desenvolver sistema de animação molecular com representações precisas mas compreensíveis

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de laboratório virtual onde os jogadores manipulam reagentes em condições controladas, com feedback visual imediato mostrando transformações químicas e energéticas.
- **Agente de Narrativa:** Desenvolveu uma estrutura narrativa onde o jogador é um pesquisador em um instituto avançado de ciências, enfrentando desafios que vão desde análise de materiais cotidianos até desenvolvimento de soluções para problemas ambientais e industriais.
- **Agente de Engajamento:** Sugeriu um sistema de descobertas progressivas no estilo "árvore tecnológica", onde novas reações desbloqueiam ferramentas e possibilidades, complementado por um caderno de laboratório digital que registra descobertas.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar elementos, compostos e suas propriedades básicas na tabela periódica.
- **Entender:** Explicar princípios de ligações químicas e como afetam as propriedades dos materiais.
- **Aplicar:** Utilizar conhecimentos de estequiometria e cinética para prever resultados de reações.
- **Analisar:** Examinar reações químicas para identificar padrões e relações entre reagentes e produtos.
- **Avaliar:** Julgar a eficiência e segurança de diferentes processos químicos para objetivos específicos.
- **Criar:** Desenvolver protocolos experimentais para sintetizar compostos ou resolver problemas práticos.

Resultado Final - Endo-GDC do Jogo “ChemLab Ventures”: O Endo-GDC do jogo “ChemLab Ventures” integra os seguintes elementos principais:

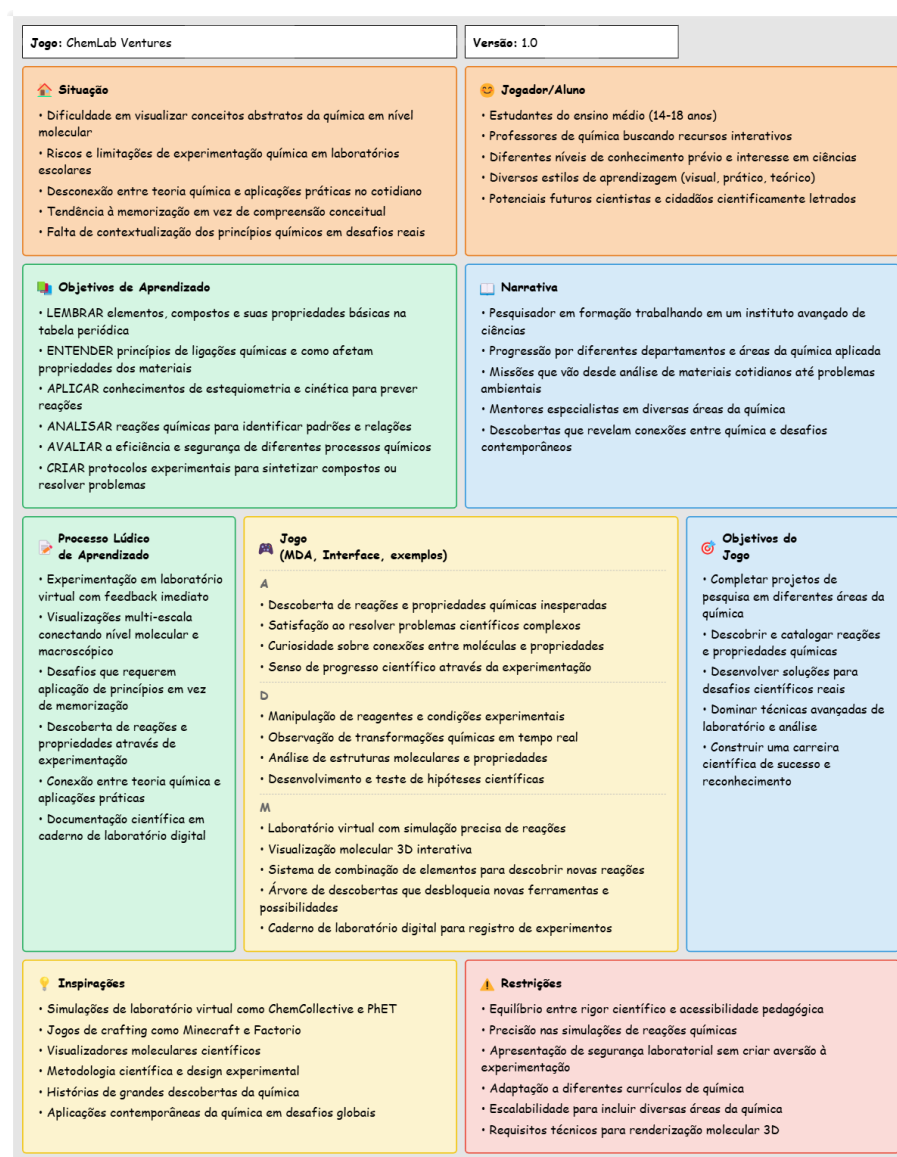


Figura A.16: Endo-GDC do jogo “ChemLab Ventures”

- **Mecânicas Endógenas:** Sistema de simulação química precisa onde princípios atômicos e moleculares determinam o comportamento dos materiais, com visualizações em múltiplas escalas que conectam o mundo microscópico a efeitos observáveis.
- **Narrativa Integrada:** Carreira de um pesquisador atravessando diferentes áreas da química aplicada, com missões que contextualizam conceitos teóricos em problemas práticos socialmente relevantes.
- **Elementos de Engajamento:** Sistema de descobertas que funciona como uma árvore tecnológica, complementado por um laboratório de experimentação livre e um caderno digital que documenta o progresso científico do jogador.
- **Avaliação Incorporada:** Feedback visual e dados quantitativos que revelam o êxito ou falha de experimentos, incentivando análise crítica e iteração

metodológica baseada em princípios científicos.

A.17 Estudo de Caso 17: Educação em Saúde para Adultos

A.17.1 Design de Jogo para Promoção de Bem-Estar e Prevenção

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo focado em educação em saúde, prevenção de doenças e promoção de hábitos saudáveis para adultos.

Fase de Brainstorming: Um professor de educação física inseriu as seguintes ideias iniciais (4 inputs):

- Simulação de impacto de escolhas diárias na saúde a longo prazo
- Sistema de metas personalizadas baseadas em perfil de saúde do jogador
- Visualização dos processos internos do corpo humano
- Abordagem holística integrando saúde física, mental e social

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (5 sugestões):

- Narrativa que acompanha diferentes personagens em jornadas de transformação de saúde
- Mecânica de "construção de hábitos" com sistema de reforço e acompanhamento
- Representação gamificada de biofeedback através de dados de dispositivos vestíveis
- Microsimulações de sistemas corporais respondendo a diferentes intervenções
- Componente social para desafios de saúde coletivos e suporte comunitário

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode promover mudanças de comportamento duradouras sem recorrer a culpabilização?”
- “De que forma serão equilibradas a simplificação necessária e a precisão científica sobre saúde?”
- “Como o sistema pode adaptar-se a diferentes condições de saúde, limitações e contextos culturais?”
- “Qual abordagem será utilizada para manter motivação em objetivos de saúde de longo prazo?”

- “Como o jogo lidará com questões sensíveis como imagem corporal e condições crônicas?”

As respostas geraram os seguintes insights:

- Focar em celebração de progressos e capacitação em vez de abordagens negativas
- Implementar sistema de personalização profunda que adapta conteúdo a necessidades específicas
- Utilizar narrativas inspiradoras e autênticas de diversas jornadas de saúde
- Criar mecanismos de micro-recompensas e visualização progressiva de melhorias

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "gêmeo digital de saúde" que simula o impacto de escolhas cotidianas em múltiplos sistemas corporais, com calibração gradual baseada em dados reais do usuário quando disponíveis.
- **Agente de Narrativa:** Desenvolveu uma estrutura narrativa não-linear centrada em um "conselho de bem-estar" com mentores de diferentes especialidades, cada um guiando uma dimensão da saúde através de histórias pessoais e desafios contextualizados.
- **Agente de Engajamento:** Sugeriu um sistema de "jardim de hábitos" visual que floresce com consistência e mostra conexões entre diferentes práticas, complementado por celebrações de marcos significativos e um componente de "aliados de jornada".

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar componentes essenciais de um estilo de vida saudável e sinais de alerta para condições comuns.
- **Entender:** Explicar como diferentes sistemas corporais interagem e respondem a hábitos e intervenções.
- **Aplicar:** Implementar estratégias de saúde preventiva adaptadas a necessidades e circunstâncias pessoais.
- **Analisar:** Examinar padrões pessoais de comportamento e seus impactos na saúde geral.
- **Avaliar:** Julgar a qualidade de informações de saúde e a adequação de diferentes abordagens ao bem-estar.
- **Criar:** Desenvolver planos de saúde personalizados que integram múltiplas dimensões de bem-estar de forma sustentável.

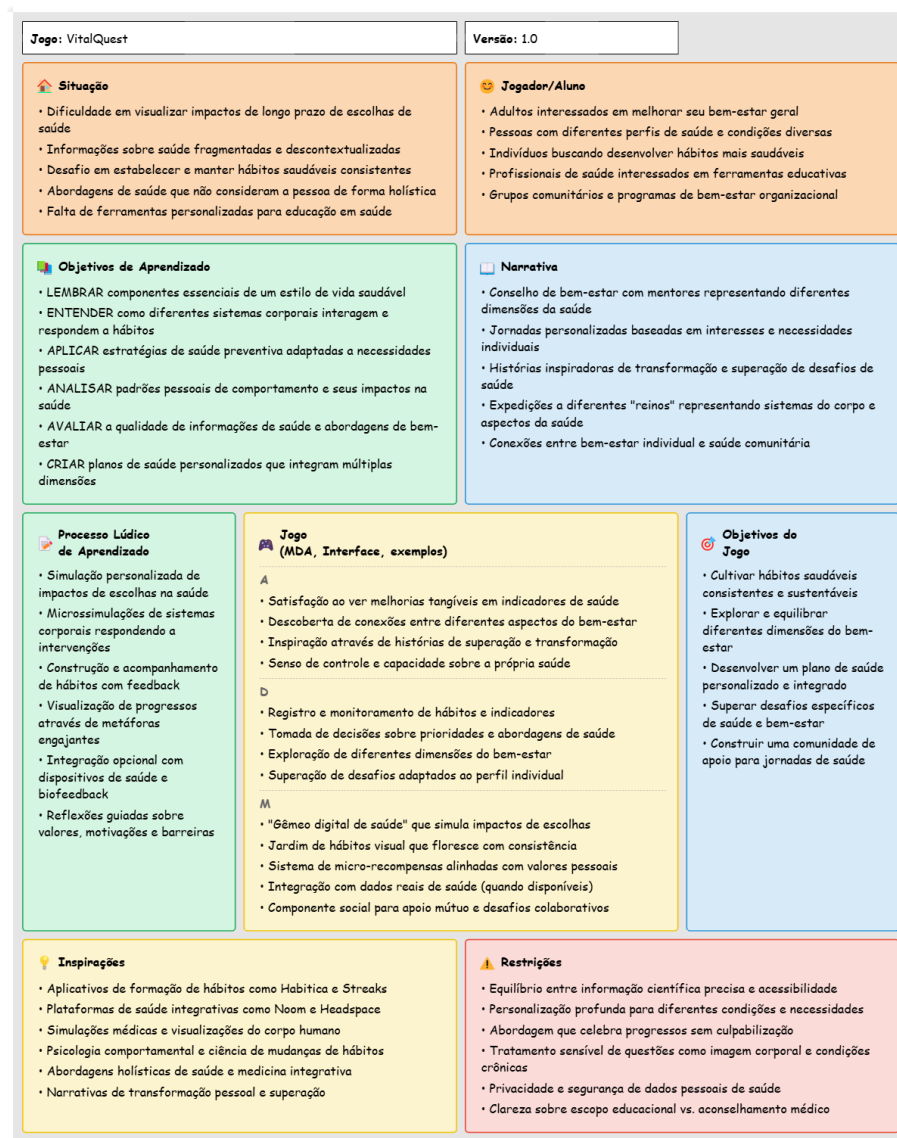


Figura A.17: Endo-GDC do jogo “VitalQuest”

Resultado Final - Endo-GDC do Jogo “VitalQuest”: O Endo-GDC do jogo “VitalQuest” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de simulação personalizada de saúde onde princípios de fisiologia, psicologia e ciência comportamental se manifestam em um "gêmeo digital" que evolui em resposta às escolhas do jogador e dados do mundo real.
- **Narrativa Integrada:** Conselho de bem-estar com mentores representando diferentes dimensões da saúde, oferecendo jornadas de aprendizado contextualizadas em histórias autênticas e desafios relevantes para o perfil do jogador.
- **Elementos de Engajamento:** Jardim de hábitos visual que prospera com práticas consistentes, sistema de micro-recompensas alinhadas com valores pessoais, e componente social para apoio mútuo e desafios colaborativos.
- **Avaliação Incorporada:** Visualizações dinâmicas de progresso em indicado-

res de saúde significativos, feedback contextualizado sobre padrões de comportamento, e reflexões guiadas que conectam aprendizado a experiência pessoal.

A.18 Estudo de Caso 18: Educação Cívica para Jovens

A.18.1 Design de Jogo para Desenvolvimento de Cidadania Ativa

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional sobre cidadania, participação democrática e engajamento comunitário para jovens entre 14-18 anos.

Fase de Brainstorming: Um educador cívico inseriu as seguintes ideias iniciais (3 inputs):

- Simulação de processos democráticos e tomada de decisão coletiva
- Exploração de direitos e responsabilidades cidadãos em contextos diversos
- Sistema para compreender impactos de políticas públicas em comunidades

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (7 sugestões):

- Cidade virtual onde jogadores assumem diferentes papéis sociais e cívicos
- Mecânica de debate com sistema de argumentação baseada em evidências
- Campanhas de mobilização comunitária para enfrentar desafios locais
- Representação de perspectivas diversas sobre questões sociais complexas
- Sistema de deliberação coletiva com mecanismos de construção de consenso
- Ciclos de feedback mostrando consequências de longo prazo de decisões políticas
- Elementos de historiografia mostrando evolução de direitos civis e participação

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode abordar temas políticos potencialmente divisivos de forma educativa e imparcial?”
- “De que forma será equilibrada a apresentação de complexidades dos sistemas democráticos sem causar cinismo?”
- “Como incentivar transferência de habilidades cívicas do ambiente virtual para o real?”
- “Qual abordagem será utilizada para tornar processos burocráticos e institucionais acessíveis e engajantes?”

- “Como o jogo pode representar diversidade de perspectivas e experiências políticas de forma respeitosa?”

As respostas geraram os seguintes insights:

- Focar em processos e habilidades cívicas em vez de posições políticas específicas
- Implementar sistema de múltiplas perspectivas com base em valores compartilhados
- Criar conexões explícitas com oportunidades de participação cívica real
- Utilizar narrativas pessoais e impactos concretos para humanizar processos abstratos

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "laboratório democrático" onde jogadores podem simular processos de governança em diferentes escalas, com mecânicas de deliberação, negociação e implementação de políticas que revelam complexidades e trade-offs.
- **Agente de Narrativa:** Desenvolveu uma estrutura narrativa centrada em uma cidade em transformação enfrentando múltiplos desafios, onde jogadores podem assumir diferentes papéis (cidadãos, representantes eleitos, servidores públicos, ativistas) para experimentar perspectivas complementares.
- **Agente de Engajamento:** Sugeriu um sistema de "impacto cívico" que visualiza mudanças comunitárias resultantes de ações coletivas, complementado por desafios baseados em casos reais e um componente de projetos que pode se estender para ações no mundo real.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar instituições democráticas, seus papéis e princípios fundamentais da cidadania.
- **Entender:** Explicar como diferentes mecanismos democráticos funcionam e a interação entre direitos e responsabilidades.
- **Aplicar:** Utilizar processos deliberativos estruturados para abordar questões comunitárias concretas.
- **Analisar:** Examinar políticas públicas e suas implicações para diferentes grupos e interesses sociais.
- **Avaliar:** Julgar a eficácia e equidade de diferentes abordagens para resolver desafios cívicos.
- **Criar:** Desenvolver propostas e iniciativas cívicas que mobilizem recursos coletivos para necessidades identificadas.

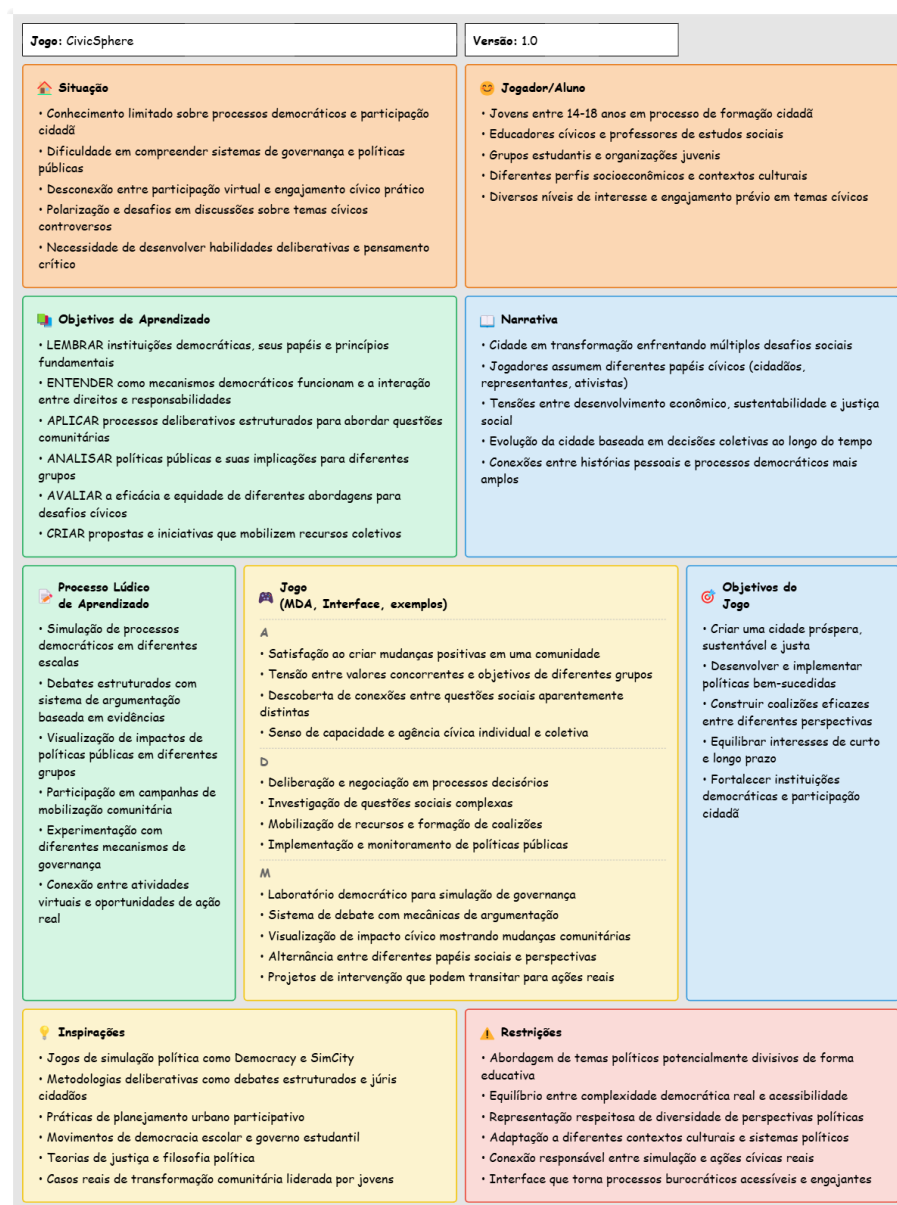


Figura A.18: Endo-GDC do jogo “CivicSphere”

Resultado Final - Endo-GDC do Jogo “CivicSphere”: O Endo-GDC do jogo “CivicSphere” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de laboratório democrático onde princípios de governança, deliberação e participação cidadã são incorporados em simulações multinível que revelam complexidades e interdependências dos processos cívicos.
- **Narrativa Integrada:** Cidade em transformação com múltiplos desafios sociais onde jogadores experimentam diferentes papéis cívicos, vivenciando perspectivas complementares sobre o funcionamento da democracia e os impactos de decisões coletivas.
- **Elementos de Engajamento:** Sistema de visualização de impacto cívico,

desafios baseados em casos reais, e projetos de intervenção que podem transitar do jogo para comunidade, criando pontes entre aprendizado simulado e ação concreta.

- **Avaliação Incorporada:** Feedback multidimensional sobre impactos de decisões políticas em diferentes grupos e setores, indicadores de saúde democrática, e reflexões guiadas sobre valores cívicos e responsabilidade coletiva.

A.19 Estudo de Caso 19: Desenvolvimento de Habilidades Socioemocionais para Crianças

A.19.1 Design de Jogo para Inteligência Emocional e Relacionamentos Saudáveis

Este estudo de caso apresenta a aplicação do sistema para o design de um jogo educacional sobre desenvolvimento socioemocional para crianças de 7-11 anos.

Fase de Brainstorming: Uma psicóloga educacional inseriu as seguintes ideias iniciais (4 inputs):

- Jogo sobre reconhecimento e expressão de emoções
- Cenários que exigem resolução colaborativa de problemas
- Personagens diversos com diferentes perspectivas e necessidades
- Sistema de feedback sobre comportamentos pró-sociais

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (5 sugestões):

- Mundo fantástico onde emoções têm manifestações visuais e poderes especiais
- Mecânica de "leitura social" para interpretar pistas não-verbais
- Sistema de diálogo com opções baseadas em comunicação assertiva
- Desafios que requerem regulação emocional e perspectiva empática
- Personagem companheiro que modela estratégias socioemocionais saudáveis

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode fazer complexidades emocionais acessíveis sem simplificação excessiva?”
- “De que forma serão abordadas diferenças culturais na expressão e interpretação emocional?”
- “Como equilibrar desafios socioemocionais com diversão e engajamento apropriados para a idade?”

- “Qual abordagem incentivará transferência de habilidades para situações reais do cotidiano?”
- “Como o jogo pode ser inclusivo para crianças neurodivergentes ou com diferentes capacidades sociais?”

As respostas geraram os seguintes insights:

- Utilizar metáforas visuais para tornar conceitos emocionais tangíveis e acessíveis
- Implementar sistema de personalização para diferentes estilos de comunicação e processamento
- Criar pontes explícitas entre cenários fantásticos e situações do mundo real
- Desenvolver sistema de suporte adaptativo que ajusta orientação conforme necessidades individuais

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "emojiverso" onde emoções são elementos tangíveis que podem ser reconhecidos, coletados e combinados para resolver desafios, com mecânicas de exploração social e colaboração baseadas em complementaridade emocional.
- **Agente de Narrativa:** Desenvolveu um mundo fantástico onde um "desequilíbrio emocional" está causando problemas em diferentes reinos, cada um representando facetas da experiência socioemocional, com personagens que personificam diferentes temperamentos e estilos de interação.
- **Agente de Engajamento:** Sugeriu um sistema de "amigos emocionais" que podem ser ajudados e que evoluem através de interações significativas, complementado por um diário reflexivo gamificado e desafios semanais que podem ser praticados em casa.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar diferentes emoções, suas manifestações físicas e gatilhos comuns.
- **Entender:** Explicar como emoções influenciam comportamentos e como diferentes perspectivas afetam interações sociais.
- **Aplicar:** Utilizar estratégias apropriadas de comunicação e regulação emocional em diferentes contextos.
- **Analisar:** Interpretar pistas sociais verbais e não-verbais para compreender necessidades e sentimentos alheios.
- **Avaliar:** Julgar a eficácia de diferentes abordagens para resolução de conflitos

e construção de relações positivas.

- **Criar:** Desenvolver soluções colaborativas para desafios interpessoais que considerem necessidades e sentimentos diversos.

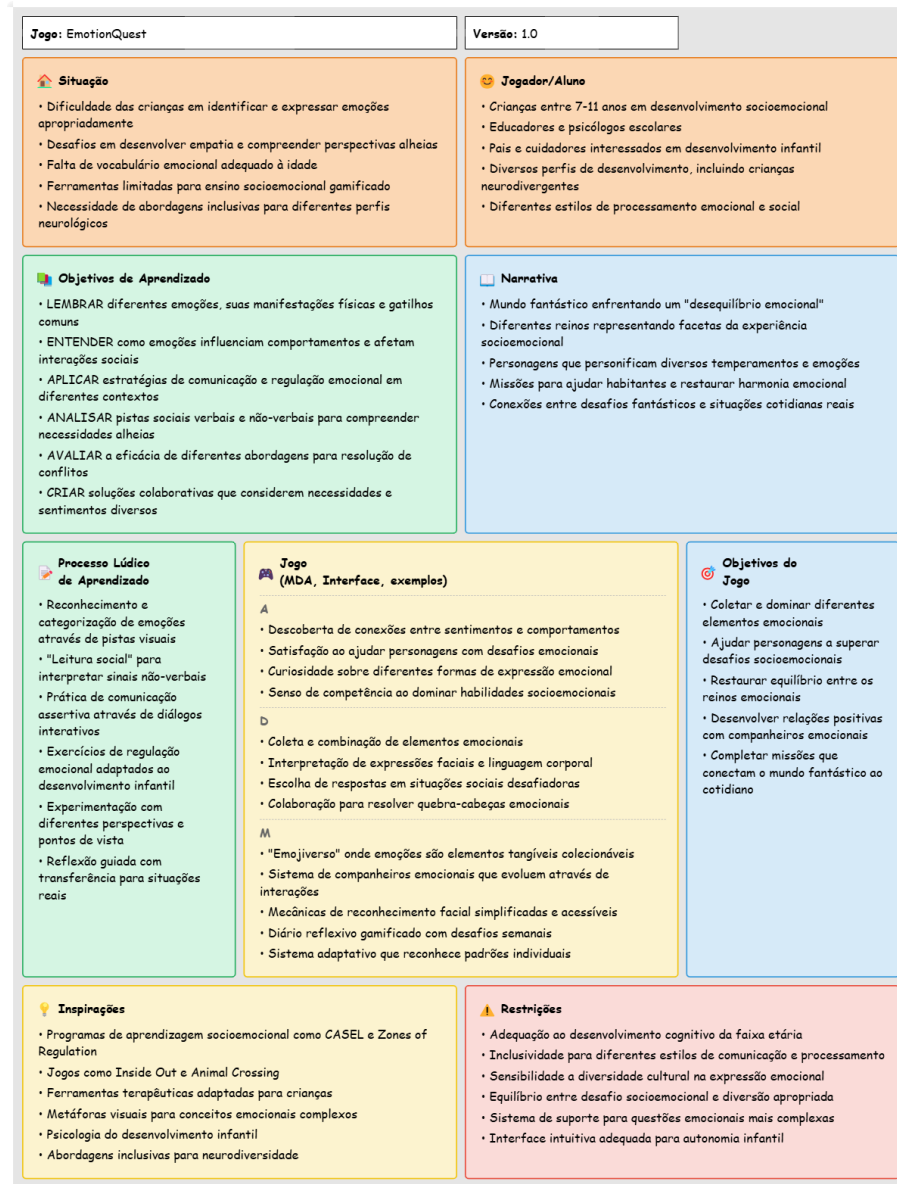


Figura A.19: Endo-GDC do jogo "EmotionQuest"

Resultado Final - Endo-GDC do Jogo "EmotionQuest": O Endo-GDC do jogo "EmotionQuest" integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de emojiverso onde emoções são elementos tangíveis com propriedades únicas, criando uma gramática interativa que torna conceitos socioemocionais acessíveis e manipuláveis através de mecânicas intuitivas e divertidas.
- **Narrativa Integrada:** Mundo fantástico enfrentando desequilíbrio emoci-

onal, com reinos temáticos representando diferentes facetas da experiência socioemocional e personagens que encarnam diversos temperamentos e estilos relacionais.

- **Elementos de Engajamento:** Sistema de companheiros emocionais que evoluem através de interações significativas, diário reflexivo gamificado, e desafios que criam pontes entre mundo fantástico e aplicações cotidianas das habilidades socioemocionais.
- **Avaliação Incorporada:** Feedback visual e narrativo que revela impactos de escolhas socioemocionais, sistema adaptativo que reconhece padrões individuais, e reflexões guiadas adequadas ao desenvolvimento cognitivo das crianças.

A.20 Estudo de Caso 20: Educação Empreendedora para Jovens Adultos

A.20.1 Design de Jogo para Desenvolvimento de Mentalidade Empreendedora

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo educacional sobre empreendedorismo e inovação para jovens adultos (18-25 anos).

Fase de Brainstorming: Um educador em empreendedorismo inseriu as seguintes ideias iniciais (2 inputs):

- Simulação de criação e gestão de startups com diferentes modelos de negócio
- Sistema que reflete dinâmicas reais de mercado e comportamento de consumidores

O sistema, demonstrando sua capacidade de expansão a partir de inputs mínimos, gerou as seguintes sugestões (8 sugestões):

- Mecânica de identificação de problemas e validação de soluções com usuários
- Ciclos de iteração de produtos baseados em feedback do mercado
- Sistema de gestão de recursos limitados com decisões de investimento
- Representação de diferentes ecossistemas de inovação e contextos culturais
- Desafios inspirados em ODS (Objetivos de Desenvolvimento Sustentável)
- Simulação de pitch e negociação com diferentes tipos de investidores
- Componente de construção de equipes e gestão de talentos diversos
- Eventos aleatórios que simulam mudanças de mercado e disrupções tecnológicas

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode equilibrar aspectos financeiros do empreendedorismo com impacto social e satisfação pessoal?”
- “De que forma o sistema refletirá realidades de empreendedorismo em diferentes contextos socioeconômicos?”
- “Como serão simulados fracassos e resiliência sem desestimular os jogadores?”
- “Qual abordagem será utilizada para representar dilemas éticos e responsabilidade empresarial?”
- “Como o jogo pode incentivar pensamento inovador e assumir riscos calculados?”

As respostas geraram os seguintes insights:

- Implementar múltiplas métricas de sucesso além do retorno financeiro
- Criar diversos cenários de empreendedorismo adaptados a contextos globais
- Desenvolver mecânicas que transformam fracassos em aprendizados e novas oportunidades
- Utilizar sistema de tomada de decisão ética com consequências multidimensionais

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "laboratório de inovação" onde jogadores passam por ciclos completos de empreendedorismo desde identificação de problemas até escala, com simulação de mercado dinâmica que responde a decisões e inovações.
- **Agente de Narrativa:** Desenvolveu um framework narrativo baseado em "jornadas empreendedoras" diversas, com mentores inspirados em casos reais e cenários de empreendedorismo em diferentes setores e contextos globais.
- **Agente de Engajamento:** Sugeriu um sistema de portfólio de empreendedor que documenta aprendizados e conquistas, complementado por uma comunidade virtual de inovação e desafios inspirados em hackathons reais.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar elementos-chave do processo empreendedor e princípios fundamentais de modelos de negócio.
- **Entender:** Explicar dinâmicas de mercado, comportamento de consumidores e fatores que influenciam sucesso empreendedor.
- **Aplicar:** Utilizar métodos estruturados para validar problemas, testar soluções e iterar com base em feedback.
- **Analisar:** Examinar modelos de negócio para identificar forças, fraquezas e

oportunidades de inovação.

- **Avaliar:** Julgar viabilidade, impacto potencial e alinhamento estratégico de diferentes oportunidades de negócio.
- **Criar:** Desenvolver propostas inovadoras que enderecem necessidades reais com modelos sustentáveis e escaláveis.

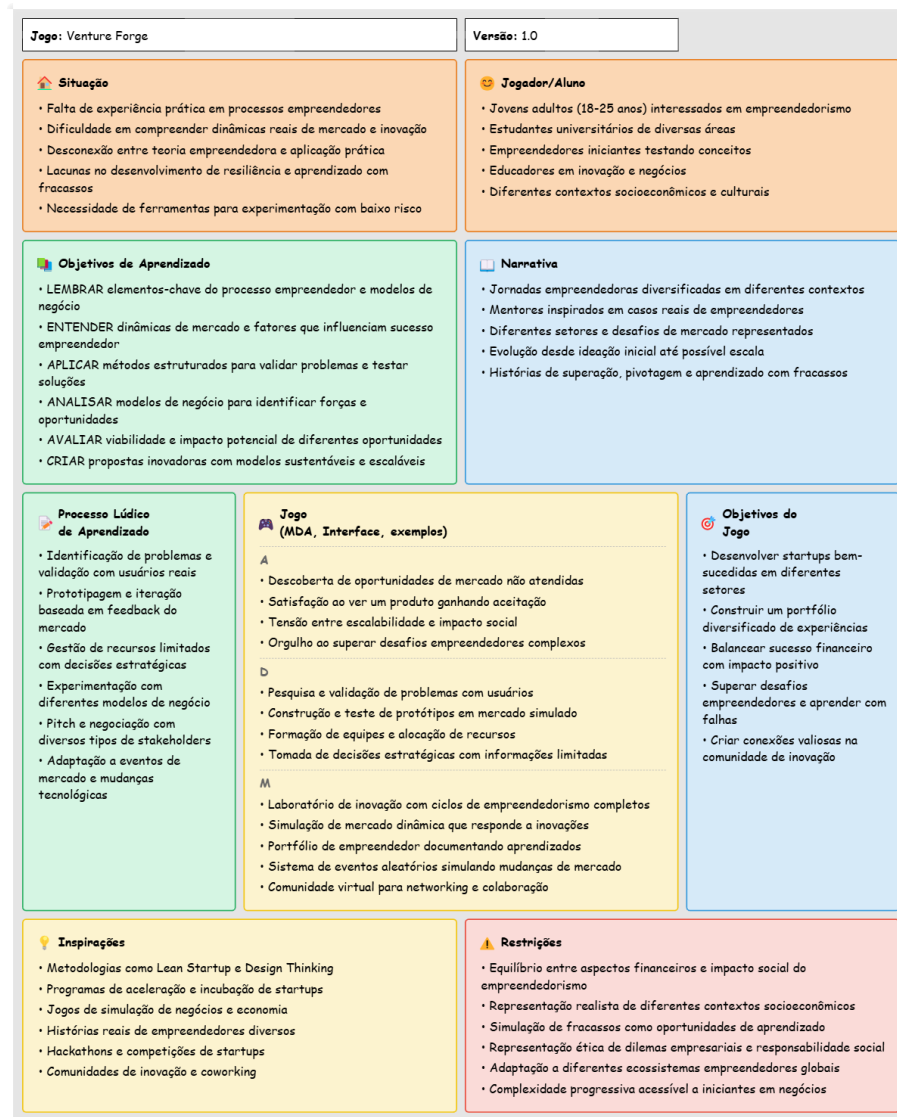


Figura A.20: Endo-GDC do jogo “Venture Forge”

Resultado Final - Endo-GDC do Jogo “Venture Forge”: O Endo-GDC do jogo “Venture Forge” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de laboratório de inovação que simula ciclos completos do processo empreendedor, com mecânicas de validação, iteração e crescimento baseadas em metodologias reais como design thinking e lean startup.

- **Narrativa Integrada:** Jornadas empreendedoras diversificadas com mentores inspirados em casos reais e cenários contextualizados em diferentes setores, regiões e desafios sociais, proporcionando perspectivas complementares sobre o empreendedorismo global.
- **Elementos de Engajamento:** Portfólio de empreendedor que documenta trajetória e aprendizados, comunidade virtual para networking e colaboração, e desafios baseados em hackathons que conectam a experiência de jogo com oportunidades do mundo real.
- **Avaliação Incorporada:** Dashboard multidimensional que visualiza impacto de decisões em diferentes stakeholders, feedback de mercado realista, e sistema de reflexão que incentiva mentalidade de crescimento e aprendizado a partir de tentativas e erros.

A.21 Estudo de Caso 21: Aprendizado de Línguas para Crianças

A.21.1 Design de Jogo para Ensino de Idiomas no Ensino Fundamental I

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional voltado para o ensino de línguas estrangeiras para crianças de 6-10 anos.

Fase de Brainstorming: Uma professora de línguas inseriu as seguintes ideias iniciais (4 inputs):

- Jogo imersivo com personagens que falam apenas o idioma-alvo
- Foco em vocabulário cotidiano e frases simples contextualizadas
- Sistema de pronúncia com feedback visual intuitivo
- Progressão que acompanha ritmo individual de aprendizado

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (5 sugestões):

- Mundo fantástico onde cada região representa um tema vocabular diferente
- Criaturas mágicas que só podem ser compreendidas no idioma-alvo
- Mecanismo de "construção de frases" que desenvolve intuição gramatical
- Mini-jogos específicos para treinar diferentes habilidades linguísticas
- Sistema de "amigos falantes" virtuais que evoluem com a interação

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode oferecer uma abordagem naturalística de aprendizado, similar à aquisição da língua materna?”
- “De que forma o sistema proporcionará repetição necessária sem se tornar monótono?”
- “Como equilibrar o uso exclusivo do idioma-alvo com a necessidade de compreensão clara das instruções?”
- “Qual abordagem será utilizada para desenvolver compreensão gramatical sem regras explícitas?”
- “Como o jogo pode incluir elementos culturais associados ao idioma?”

As respostas geraram os seguintes insights:

- Implementar contextos visuais ricos que facilitam compreensão sem tradução
- Criar variações temáticas de atividades fundamentalmente similares
- Utilizar um personagem companheiro bilíngue que oferece suporte inicial
- Desenvolver padrões gramáticos através de blocos de construção frasais
- Incorporar celebrações, comidas e tradições culturais como elementos de jogo

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "tesouros linguísticos" onde palavras e frases são colecionáveis e utilizáveis, com feedback imediato de pronúncia através de análise de áudio simplificada adaptada para vozes infantis.
- **Agente de Narrativa:** Desenvolveu uma história sobre um mundo mágico onde um "enevoamento linguístico" causou confusão entre diferentes regiões, e o jogador, com ajuda de um companheiro tradutor, precisa restaurar a comunicação aprendendo o idioma dos diferentes habitantes.
- **Agente de Engajamento:** Sugeriu um sistema de amigos linguísticos customizáveis que reagem positivamente quando compreendidos no idioma-alvo, complementado por um livro de aventuras pessoal que registra conquistas e "fotos" de momentos especiais.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Reconhecer e reproduzir vocabulário básico e expressões cotidianas no idioma-alvo.
- **Entender:** Compreender instruções simples e conversas contextualizadas no idioma-alvo.
- **Aplicar:** Utilizar palavras e frases aprendidas para comunicar necessidades e ideias simples.
- **Analisar:** Identificar padrões de construção frasal e relações entre palavras e

contextos.

- **Avaliar:** Determinar a adequação de expressões específicas para diferentes situações comunicativas.
- **Criar:** Formular novas combinações de palavras e expressões para comunicar ideias próprias.



Figura A.21: Endo-GDC do jogo “LinguaMagica”

Resultado Final - Endo-GDC do Jogo “LinguaMagica”: O Endo-GDC do jogo “LinguaMagica” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de tesouros linguísticos colecionáveis e utilizáveis que funcionam como ferramentas de comunicação, com mecânicas de reconhecimento de fala e pronúncia adaptadas para crianças, e feedback visual intuitivo.
- **Narrativa Integrada:** Mundo mágico afetado por "enevoamento linguístico" onde restaurar comunicação entre regiões temáticas requer aprendizado

do idioma, com personagem companheiro bilíngue que oferece suporte adaptativo decrescente.

- **Elementos de Engajamento:** Criaturas e amigos linguísticos que respondem positivamente quando compreendidos, livro de aventuras personalizado, e sistema de recompensas baseado em desbloqueio de novas áreas e habilidades mágicas.
- **Avaliação Incorporada:** Feedback contextual através de reações dos personagens, sistema de compreensão que adapta complexidade a partir do desempenho, e visualização de progresso no livro de aventuras.

A.22 Estudo de Caso 22: Exploração Espacial e Astronomia

A.22.1 Design de Jogo para Ensino de Conceitos Astronômicos

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo educacional sobre astronomia, exploração espacial e física básica para estudantes entre 11-15 anos.

Fase de Brainstorming: Um professor de ciências inseriu as seguintes ideias iniciais (3 inputs):

- Simulação de missões espaciais com física realista simplificada
- Exploração do sistema solar com dados astronômicos precisos
- Sistema de construção e lançamento de foguetes

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (6 sugestões):

- Progressão histórica das tecnologias e descobertas espaciais
- Multiplayer colaborativo para missões complexas exigindo especialidades distintas
- Centro de comando com planejamento de missões e gestão de recursos
- Visualização de fenômenos astronômicos como buracos negros e formação estelar
- Sistema de pesquisa científica desbloqueando novas tecnologias e capacidades
- Representação de desafios reais da exploração espacial (radiação, microgravidade)

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como equilibrar precisão científica com acessibilidade e diversão?”
- “De que forma o jogo pode representar escalas astronômicas de forma compreensível?”
- “Como incorporar princípios de física sem exigir conhecimento matemático avançado?”
- “Qual abordagem será utilizada para conectar exploração espacial com questões terrestres?”
- “Como o jogo pode inspirar interesse em carreiras científicas e tecnológicas?”

As respostas geraram os seguintes insights:

- Utilizar níveis configuráveis de realismo físico com explicações contextuais
- Implementar sistema de "zoom cosmológico" com comparações relativas
- Criar visualizações intuitivas de forças e trajetórias com previsões visuais
- Desenvolver missões que conectam dados espaciais com aplicações na Terra
- Incorporar perfis inspiradores de cientistas reais e caminhos de carreira

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de planejamento e execução de missões com física orbital simplificada mas precisa, complementado por ferramentas de design de espaçonaves com componentes modulares e simulação de condições espaciais.
- **Agente de Narrativa:** Desenvolveu uma estrutura narrativa centrada em uma agência espacial internacional do futuro próximo, com programas de exploração progressivos desde orbitais terrestres até exploração interplanetária e missões científicas especializadas.
- **Agente de Engajamento:** Sugeriu um sistema de conquistas baseadas em marcos históricos reais da exploração espacial, complementado por um enciclopédia astronômica que se expande com descobertas e um componente de compartilhamento de missões customizadas.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar corpos celestes do sistema solar e suas características principais.
- **Entender:** Explicar princípios básicos de física orbital, propulsão e condições espaciais.
- **Aplicar:** Utilizar conhecimentos astronômicos para planejar e executar missões espaciais viáveis.
- **Analisar:** Interpretar dados científicos coletados durante missões e observa-

ções astronômicas.

- **Avaliar:** Julgar a eficiência e viabilidade de diferentes abordagens para desafios da exploração espacial.
- **Criar:** Desenvolver missões originais e veículos espaciais adaptados a objetivos científicos específicos.

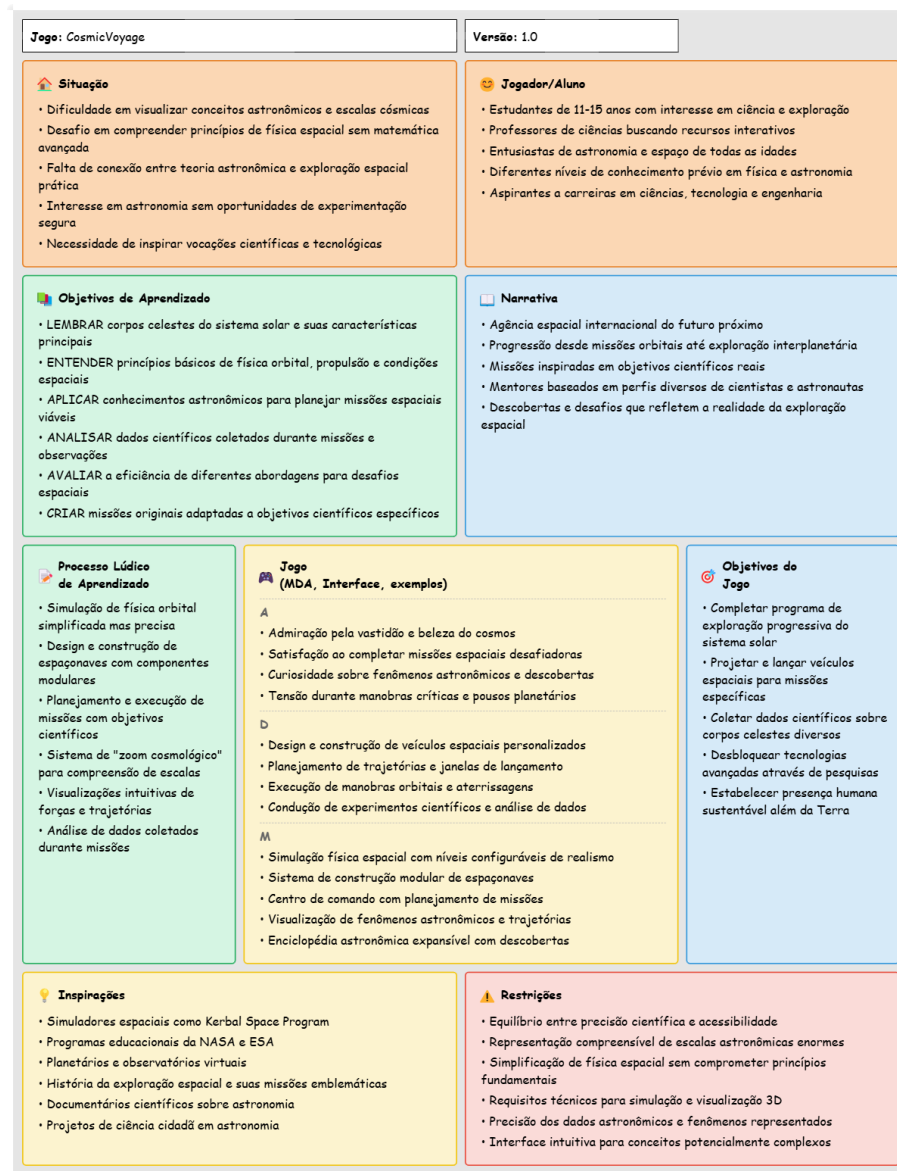


Figura A.22: Endo-GDC do jogo “CosmicVoyage”

Resultado Final - Endo-GDC do Jogo “CosmicVoyage”: O Endo-GDC do jogo “CosmicVoyage” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de física orbital simplificada mas precisa que governa todas as operações espaciais, com mecânicas de design de espaçonaves modulares e simulação de condições extraterrestres que refletem princípios científicos reais.

- **Narrativa Integrada:** Agência espacial internacional do futuro próximo com programas de exploração progressivos, missões inspiradas em objetivos científicos reais, e personagens mentores baseados em perfis diversos de cientistas e astronautas.
- **Elementos de Engajamento:** Conquistas baseadas em marcos históricos da exploração espacial, enciclopédia astronômica expansível, e plataforma de compartilhamento de missões e designs personalizados.
- **Avaliação Incorporada:** Simulações preditivas de trajetórias e resultados de missão, análise científica de dados coletados, e recapitulação de decisões técnicas com explicações contextuais.

A.23 Estudo de Caso 23: Pensamento Crítico e Lógica

A.23.1 Design de Jogo para Desenvolvimento de Raciocínio Lógico

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional voltado para o desenvolvimento de pensamento crítico, lógica e habilidades de argumentação para estudantes do Ensino Médio.

Fase de Brainstorming: Um professor de filosofia inseriu as seguintes ideias iniciais (2 inputs):

- Jogo sobre identificação de falácias lógicas e construção de argumentos
- Simulação de debates com análise estruturada de argumentação

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (7 sugestões):

- Mundo narrativo onde o pensamento lógico afeta a realidade física
- Sistema de "desconstrução" de afirmações em seus componentes lógicos
- Desafios de investigação baseados em evidências e inferência
- Personagens representando diferentes métodos de raciocínio e vieses cognitivos
- Mecânica de debate competitivo com estrutura formal de argumentação
- Visualização dinâmica de conexões lógicas e relações causais
- Dilemas éticos exigindo aplicação de diferentes sistemas filosóficos

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como tornar conceitos abstratos de lógica formal acessíveis e engajantes?”

- “De que forma o jogo pode equilibrar o desenvolvimento de pensamento crítico com respeito a diferentes perspectivas?”
- “Como simular a complexidade de argumentos do mundo real sem simplificação excessiva?”
- “Qual abordagem será utilizada para incentivar transferência de habilidades lógicas para contextos cotidianos?”
- “Como o jogo pode evitar reforçar polarização ao abordar tópicos controversos?”

As respostas geraram os seguintes insights:

- Utilizar metáforas visuais e puzzles para representar estruturas lógicas
- Implementar sistema que separa valor de verdade de técnicas de argumentação
- Criar um espectro de complexidade gradual com múltiplas variáveis
- Desenvolver cenários paralelos entre mundo fantástico e aplicações cotidianas
- Focar em temas históricos ou hipotéticos para praticar com distanciamento emocional

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "construção lógica" onde premissas e conclusões são representadas como blocos interconectáveis, com ferramentas de análise para identificar inconsistências e falácias, e um mecanismo de debate estruturado em turnos.
- **Agente de Narrativa:** Desenvolveu um mundo dividido entre reinos representando diferentes abordagens de pensamento (empírico, dedutivo, dialético), com uma academia de lógica onde o jogador treina para se tornar um "arquiteto de argumentos" capaz de resolver conflitos através do raciocínio.
- **Agente de Engajamento:** Sugeriu um sistema de progressão baseado no domínio de diferentes estruturas lógicas e técnicas de argumentação, complementado por torneios de debate com personagens históricos e um diário de reflexão sobre aplicações no mundo real.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar estruturas lógicas básicas, falácias comuns e tipos de argumentos.
- **Entender:** Explicar relações entre premissas e conclusões e como diferentes estruturas argumentativas funcionam.
- **Aplicar:** Utilizar ferramentas lógicas para analisar afirmações e construir argumentos válidos.

- **Analisar:** Decompor argumentos complexos em seus componentes e identificar pressupostos implícitos.
- **Avaliar:** Julgar a solidez de diferentes argumentos considerando validade lógica e evidência factual.
- **Criar:** Desenvolver linhas argumentativas originais para diferentes contextos e audiências.



Figura A.23: Endo-GDC do jogo “LogicCraft”

Resultado Final - Endo-GDC do Jogo “LogicCraft”: O Endo-GDC do jogo “LogicCraft” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de construção lógica com blocos representando premissas, inferências e conclusões, com ferramentas de análise para identificar relações, inconsistências e falácias, complementado por um mecanismo de debate estruturado.

- **Narrativa Integrada:** Mundo dividido entre reinos de pensamento com uma academia de lógica central, apresentando desafios que exigem diferentes formas de raciocínio e personagens históricos que encarnam tradições filosóficas distintas.
- **Elementos de Engajamento:** Progressão através do domínio de estruturas lógicas e técnicas argumentativas, torneios de debate com dificuldade crescente, e diário reflexivo que conecta aprendizados abstratos com aplicações concretas.
- **Avaliação Incorporada:** Visualização dinâmica da estrutura lógica de argumentos, feedback imediato sobre validade e solidez, e sistema de análise pós-debate que identifica forças e pontos de melhoria.

A.24 Estudo de Caso 24: Sustentabilidade e Gestão de Recursos

A.24.1 Design de Jogo para Educação Ambiental e Econômica

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo educacional sobre gestão sustentável de recursos, economia circular e desafios ambientais para público jovem adulto.

Fase de Brainstorming: Um engenheiro ambiental inseriu as seguintes ideias iniciais (3 inputs):

- Simulação de cidade com gestão de recursos e impactos ambientais
- Sistema de ciclo de vida de produtos e materiais
- Visualização de externalidades e efeitos de longo prazo

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (6 sugestões):

- Múltiplas escalas de gestão desde individual até global
- Eventos climáticos e crises ambientais que intensificam ao longo do tempo
- Sistema de "fluxo de materiais" visível mostrando ciclos e perdas
- Representação de diferentes stakeholders com interesses e perspectivas diversos
- Integração de inovações tecnológicas e mudanças comportamentais como soluções
- Métricas multidimensionais além do crescimento econômico tradicional

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode representar complexidades de sistemas socioecológicos sem simplificação excessiva?”
- “De que forma serão equilibradas preocupações ambientais, sociais e econômicas?”
- “Como abordar incertezas e complexidades inerentes à gestão ambiental?”
- “Qual abordagem será utilizada para evitar mensagens prescritivas simplistas?”
- “Como o jogo pode inspirar ação real sem sobrecarregar com ansiedade climática?”

As respostas geraram os seguintes insights:

- Implementar múltiplas camadas de complexidade que podem ser reveladas gradualmente
- Utilizar um sistema de dashboard com indicadores interdependentes e trade-offs visíveis
- Incorporar elementos estocásticos e feedback não-linear para representar incertezas
- Apresentar múltiplas soluções viáveis com diferentes compensações
- Focar em capacidade de agência e exemplos de sucesso incremental

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de simulação integrada com fluxos visíveis de materiais, energia e resíduos atravessando fronteiras espaciais e temporais, complementado por ferramentas de design de processos e políticas com feedback em múltiplas dimensões.
- **Agente de Narrativa:** Desenvolveu uma estrutura centrada em uma região que evolui ao longo de várias gerações respondendo às decisões dos jogadores, com personagens representando diferentes setores e visões sobre sustentabilidade, e histórias pessoais que humanizam desafios abstratos.
- **Agente de Engajamento:** Sugeriu um sistema de cenários desafiadores inspirados em casos reais, complementado por um laboratório de inovação para testar soluções e uma plataforma para compartilhar estratégias bem-sucedidas e discutir alternativas.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar componentes essenciais de sistemas econômicos e ecológicos e suas interconexões.
- **Entender:** Explicar princípios de economia circular, serviços ecossistêmicos

e externalidades ambientais.

- **Aplicar:** Implementar estratégias de gestão sustentável adaptadas a diferentes contextos e limitações.
- **Analisar:** Examinar fluxos de recursos e impactos através de múltiplas escalas espaciais e temporais.
- **Avaliar:** Julgar diferentes abordagens considerando trade-offs e impactos em diversos stakeholders.
- **Criar:** Desenvolver inovações e políticas que promovam sustentabilidade econômica e ambiental de longo prazo.

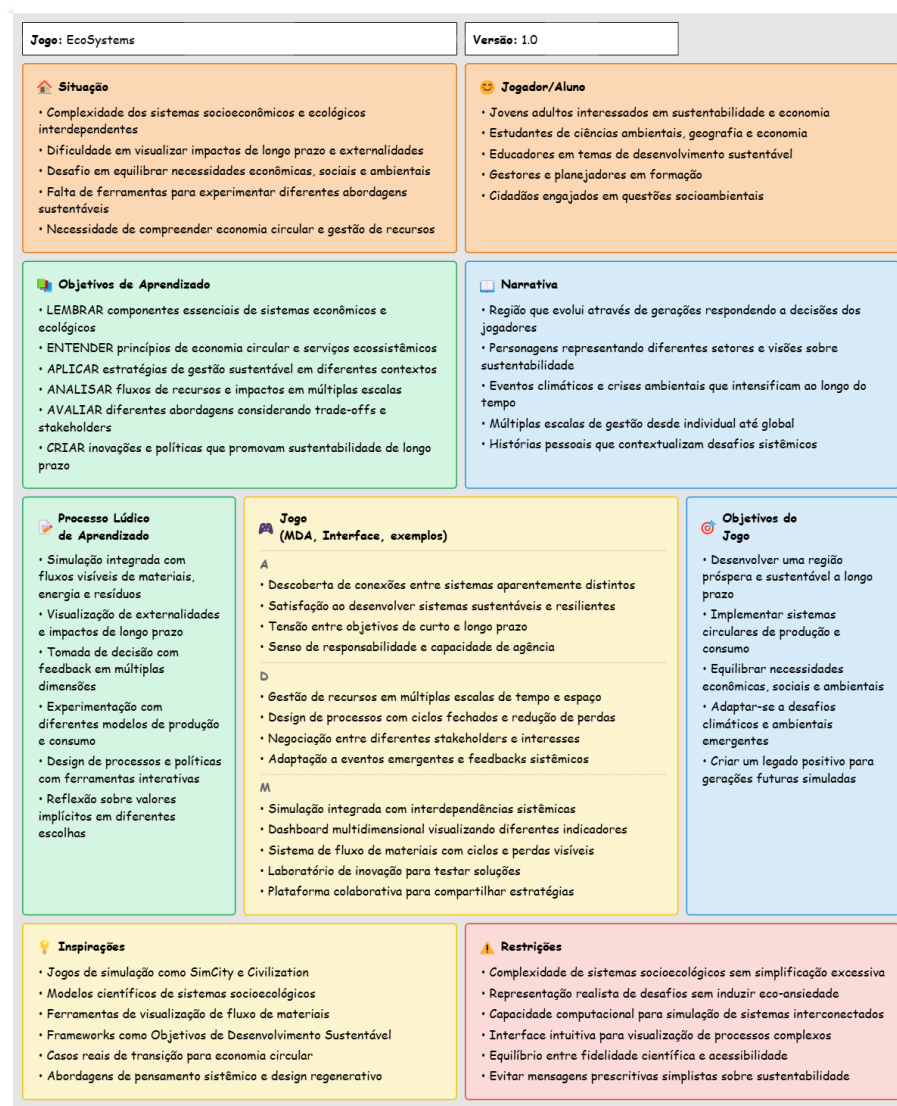


Figura A.24: Endo-GDC do jogo “EcoSystems”

Resultado Final - Endo-GDC do Jogo “EcoSystems”: O Endo-GDC do jogo “EcoSystems” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de simulação integrada com fluxos visíveis

de materiais, energia e resíduos que conecta dimensões econômicas e ecológicas em múltiplas escalas temporais e espaciais, com feedback em cascata e interdependências sistêmicas.

- **Narrativa Integrada:** Região que evolui através de gerações respondendo a decisões dos jogadores, com personagens representando diferentes perspectivas e setores, e histórias pessoais que contextualizam desafios sistemáticos em experiências humanas.
- **Elementos de Engajamento:** Cenários desafiadores baseados em casos reais, laboratório de inovação para experimentação livre, e plataforma colaborativa para compartilhamento de estratégias e discussão de alternativas.
- **Avaliação Incorporada:** Dashboard multidimensional visualizando impactos em diferentes indicadores ao longo do tempo, sistema de previsão mostrando trajetórias prováveis, e reflexões periódicas sobre valores implícitos em diferentes escolhas.

A.25 Estudo de Caso 25: Expressão Artística e Criatividade

A.25.1 Design de Jogo para Desenvolvimento de Habilidades Artísticas

Este estudo de caso ilustra a aplicação do sistema para o design de um jogo educacional voltado para desenvolvimento de criatividade, técnicas artísticas e expressão visual.

Fase de Brainstorming: Uma professora de artes inseriu as seguintes ideias iniciais (3 inputs):

- Jogo focado em exploração de técnicas e meios artísticos diversos
- Sistema de missões inspiradas em movimentos e estilos históricos
- Galeria para compartilhamento e apreciação de criações

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (5 sugestões):

- Mundo onde arte influencia o ambiente e resolve desafios
- Personagens mentores baseados em artistas históricos
- Ferramentas de composição com princípios de design incorporados
- Sistema de desafios criativos com restrições inspiradoras
- Mecânicas de colaboração artística entre jogadores

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode desenvolver habilidades técnicas sem limitar expressão pessoal?”
- “De que forma serão equilibradas diretrizes estruturadas e experimentação livre?”
- “Como incorporar contexto histórico e cultural sem reduzir arte a fórmulas?”
- “Qual abordagem será utilizada para desenvolver apreciação crítica respeitosa?”
- “Como o jogo pode atender diferentes níveis de habilidade e confiança artística?”

As respostas geraram os seguintes insights:

- Implementar ferramentas que facilitam técnicas mas não restringem resultados
- Alternar entre desafios estruturados e espaços de criação livre
- Apresentar contexto como inspiração e diálogo em vez de regras prescritivas
- Desenvolver vocabulário visual e frameworks para discussão sem julgamento binário
- Criar múltiplos pontos de entrada e sistemas de suporte adaptáveis

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de ferramentas artísticas que incorporam princípios fundamentais de design como paletas harmônicas, proporções áureas e teoria das cores, complementado por missões de resolução criativa de problemas.
- **Agente de Narrativa:** Desenvolveu um mundo onde a "criatividade primordial" se fragmentou em diferentes tradições artísticas, com mentores inspirados em artistas históricos guiando o jogador em uma jornada para reunir e revitalizar expressão artística.
- **Agente de Engajamento:** Sugeriu um sistema de exposições temáticas com desafios criativos variados, complementado por um atelier social para colaboração e apreciação, e um diário visual que documenta a evolução artística do jogador.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar elementos básicos de design visual, técnicas artísticas e movimentos históricos.
- **Entender:** Explicar princípios de composição, teoria das cores e contextos

culturais de diferentes estilos.

- **Aplicar:** Utilizar técnicas artísticas e elementos de design para expressar ideias visuais intencionais.
- **Analisar:** Examinar como diferentes elementos formais e escolhas técnicas contribuem para o impacto de uma obra.
- **Avaliar:** Apreciar criticamente trabalhos artísticos considerando aspectos técnicos, expressivos e contextuais.
- **Criar:** Desenvolver obras originais que expressem visão pessoal através de meios visuais.

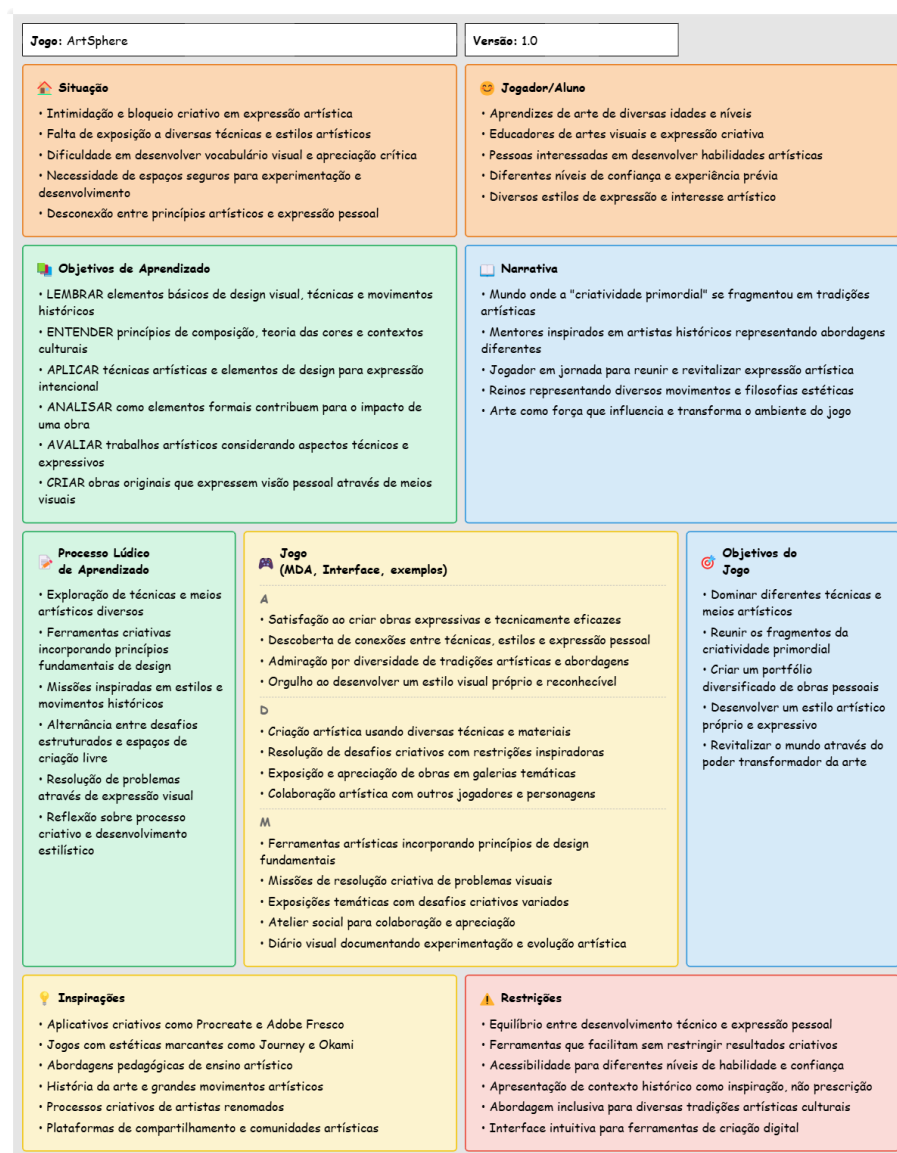


Figura A.25: Endo-GDC do jogo “ArtSphere”

Resultado Final - Endo-GDC do Jogo “ArtSphere”: O Endo-GDC do jogo “ArtSphere” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de ferramentas artísticas que incorporam princípios fundamentais de design como suporte mas não como restrição, com missões onde expressão visual funciona como mecanismo principal de interação com o mundo e resolução de desafios.
- **Narrativa Integrada:** Mundo onde a "criatividade primordial" fragmentou-se em tradições artísticas diversas, com mentores inspirados em artistas históricos representando diferentes abordagens, técnicas e filosofias estéticas.
- **Elementos de Engajamento:** Exposições temáticas com desafios criativos variados, atelier social para colaboração e apreciação, e diário visual documentando experimentação e evolução artística individual.
- **Avaliação Incorporada:** Feedback contextualizado que considera intenção e processo além do resultado, sistema de reflexão que desenvolve vocabulário visual e pensamento crítico, e galerias que celebram diversidade de expressão.

A.26 Estudo de Caso 26: Programação e Pensamento Computacional

A.26.1 Design de Jogo para Ensino de Conceitos de Programação

Este estudo de caso demonstra a aplicação do sistema para o design de um jogo educacional sobre programação e pensamento computacional para adolescentes (12-16 anos).

Fase de Brainstorming: Um educador de tecnologia inseriu as seguintes ideias iniciais (2 inputs):

- Jogo que ensina fundamentos de programação sem depender de sintaxe específica
- Ambiente sandbox para criação de soluções para problemas incrementais

O sistema analisou estas ideias e sugeriu os seguintes conceitos complementares (7 sugestões):

- Mundo onde código afeta o ambiente físico de forma visual
- Personagens controlados através de sequências lógicas programáveis
- Progressão desde sequências simples até estruturas algorítmicas complexas
- Sistema de "depuração" visual que torna erros compreensíveis
- Mecânicas de colaboração onde diferentes funções de código interagem
- Desafios baseados em problemas computacionais do mundo real
- Interface dual que conecta blocos visuais a código textual

Fase de Método Socrático: O Agente Socrático aplicou as seguintes questões reflexivas:

- “Como o jogo pode desenvolver pensamento computacional sem focar em uma linguagem específica?”
- “De que forma a progressão pode atender diferentes ritmos de aprendizado?”
- “Como representar conceitos abstratos de programação de maneira tangível e visual?”
- “Qual abordagem será utilizada para tornar a depuração um processo positivo em vez de frustrante?”
- “Como o jogo pode desenvolver criatividade além da aplicação mecânica de conceitos?”

As respostas geraram os seguintes insights:

- Focar em padrões lógicos e estruturas algorítmicas fundamentais
- Implementar sistema adaptativo com múltiplos caminhos de progressão
- Utilizar metáforas visuais consistentes para representar conceitos
- Transformar depuração em mecânica de investigação com feedback construtivo
- Criar espaços de experimentação livre com desafios de final aberto

Fase de Endo-GDC: No preenchimento do Endo-GDC, os agentes especializados contribuíram com sugestões específicas:

- **Agente de Mecânicas:** Propôs um sistema de "programação visual tangível" onde blocos lógicos afetam um mundo responsivo, com transição gradual para sintaxe textual, complementado por ferramentas de simulação que permitem testar e refinar soluções incrementalmente.
- **Agente de Narrativa:** Desenvolveu um mundo digital que perdeu sua estrutura lógica, onde o jogador como "arquiteto de código" deve restaurar funcionalidade através de sequências lógicas, com áreas temáticas representando diferentes conceitos computacionais.
- **Agente de Engajamento:** Sugeriu um sistema de desafios progressivos inspirados em problemas computacionais reais, complementado por um repositório de soluções compartilháveis e um modo criativo para aplicação livre dos conceitos aprendidos.

Fase de Taxonomia de Bloom: O sistema gerou os seguintes objetivos educacionais:

- **Lembrar:** Identificar estruturas fundamentais de programação como sequências, loops e condicionais.
- **Entender:** Explicar como diferentes estruturas algorítmicas funcionam e interagem para resolver problemas.

- **Aplicar:** Utilizar conceitos de programação para criar soluções funcionais para problemas específicos.
- **Analisar:** Decompor problemas complexos em componentes menores e identificar padrões algorítmicos.
- **Avaliar:** Comparar diferentes soluções de código considerando eficiência, legibilidade e escalabilidade.
- **Criar:** Desenvolver programas originais que apliquem conceitos computacionais para resolver problemas diversos.

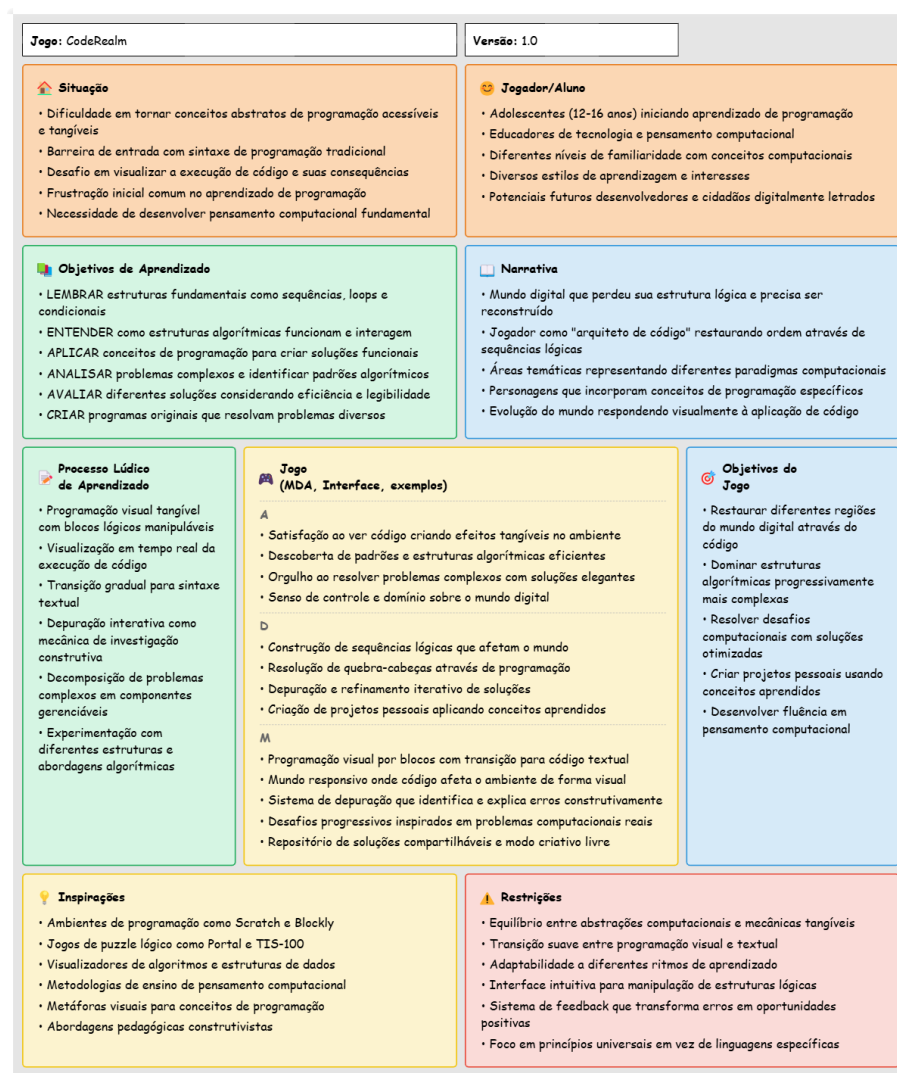


Figura A.26: Endo-GDC do jogo “CodeRealm”

Resultado Final - Endo-GDC do Jogo “CodeRealm”: O Endo-GDC do jogo “CodeRealm” integra os seguintes elementos principais:

- **Mecânicas Endógenas:** Sistema de programação visual tangível onde blocos lógicos manipulam um mundo responsivo, com transição gradual para sintaxe textual, e ferramentas de simulação que permitem testar e refinar soluções de

forma iterativa e visual.

- **Narrativa Integrada:** Mundo digital que perdeu sua estrutura lógica e precisa ser reconstruído através de código, com áreas temáticas representando diferentes paradigmas computacionais e personagens que incorporam conceitos de programação específicos.
- **Elementos de Engajamento:** Desafios progressivos inspirados em problemas computacionais reais, repositório de soluções compartilháveis, e modo criativo onde conceitos aprendidos podem ser aplicados em projetos pessoais sem restrições.
- **Avaliação Incorporada:** Visualização em tempo real da execução de código, sistema de depuração que identifica e explica erros de forma construtiva, e análise automática que sugere otimizações e abordagens alternativas.