

UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM

INSTITUTO DE COMPUTAÇÃO - ICOMP

PROGRAMA PÓS-GRADUAÇÃO EM INFORMÁTICA - PPGI

GLORIA: ferramenta interativa integrada a um *Large Language Model* para auxiliar na modelagem de dados e implementação de técnicas de *Game Learning Analytics*

Fabrizio Honda Franzoia

Manaus - AM

Julho de 2026

Fabrizio Honda Franzoia

GLORIA: ferramenta interativa integrada a um *Large Language Model* para auxiliar na modelagem de dados e implementação de técnicas de *Game Learning Analytics*

Dissertação de mestrado submetida à avaliação como requisito para a obtenção do título de Mestre em Informática no Programa de Pós-Graduação em Informática, Instituto de Computação.

Orientadora:

Elaine Harada, D.Sc.

Coorientadoras:

Fernanda Gabriela de Sousa Pires, D.Sc.

Marcela Sávia Picanço Pessoa Bastos, D.Sc.

Universidade Federal do Amazonas - UFAM

Instituto de Computação - IComp

Manaus - AM

Julho de 2026

Ficha Catalográfica

Elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

F837g Franzoia, Fabrizio Honda
GLORIA: ferramenta interativa integrada a um Large Language Model para auxiliar na modelagem de dados e implementação de técnicas de Game Learning Analytics / Fabrizio Honda Franzoia. - 2026.
248 f. : il., color. ; 31 cm.

Orientador(a): Elaine Harada.
Coorientador(a): Fernanda Gabriela de Sousa Pires.
Coorientador(a): Marcela Sávia Picanço Pessoa Bastos.
Dissertação (mestrado) - Universidade Federal do Amazonas, Programa de Pós-Graduação em Informática, Manaus, 2026.

1. Game Learning Analytics. 2. GLBoard. 3. Large Language Models. 4. Modelagem de dados. 5. Ferramenta. I. Harada, Elaine. II. Pires, Fernanda Gabriela de Sousa. III. Bastos, Marcela Sávia Picanço Pessoa. IV. Universidade Federal do Amazonas. Programa de Pós-Graduação em Informática. V. Título

---

Ministério da Educação
Universidade Federal do Amazonas
Coordenação do Programa de Pós-Graduação em Informática

FOLHA DE APROVAÇÃO

"GLORIA: FERRAMENTA INTERATIVA INTEGRADA A UM LARGE LANGUAGE MODEL PARA AUXILIAR NA MODELAGEM DE DADOS E IMPLEMENTAÇÃO DE TÉCNICAS DE GAME LEARNING ANALYTICS"

FABRIZIO HONDA FRANZOIA

Dissertação de Mestrado defendida e aprovada pela banca examinadora constituída pelos Professores:

Prof.ª Dr.ª Elaine Harada - **Presidente**

Prof. Dr. David Braga Fernandes de Oliveira - **Membro Interno**

Prof. Dr. Luiz Rodrigues - **Membro Externo**

gov.br Documento assinado digitalmente
LUIZ ANTONIO LIMA RODRIGUES
Data: 23/06/2026 07:41:35-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Rafael Dias Araújo - **Membro Externo**

gov.br Documento assinado digitalmente
RAFAEL DIAS ARAUJO
Data: 23/06/2026 08:05:15-0300
Verifique em https://validar.iti.gov.br

Manaus, 15 de junho de 2026.

---

sei! assinatura eletrônica

Documento assinado eletronicamente por **Elaine Harada**, **Professor do Magistério Superior**, em 18/06/2026, às 10:57, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **David Braga Fernandes de Oliveira**, **Professor do Magistério Superior**, em 18/06/2026, às 16:53, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015 .

A autenticidade deste documento pode ser conferida no site https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3192953** e o código CRC **0AF253CB**.

Avenida General Rodrigo Octávio, 6200 - Bairro Coroado I Campus Universitário Senador Arthur Virgílio Filho, Setor Norte - Telefone: (92) 3305-1181 / Ramal 1193
CEP 69080-900, Manaus/AM, coordenadorppgi@icomp.ufam.edu.br

Referência: Processo nº 23105.022157/2026-61 SEI nº 3192953

Dedico este trabalho a todos que acreditam em mim e estão comigo nessa jornada, em especial à minha família e ao ThinkTEd Lab.

Agradecimentos

Gratidão. Essa é palavra que define grande parte da minha vida, incluindo minha jornada no mestrado. Primeiramente, agradeço a Deus pela minha saúde, por sempre me abençoar e por ter desenhado os caminhos para que eu pudesse chegar até aqui. Fez-me perceber que sou rico, e não estou falando de dinheiro. Estou cercado por pessoas maravilhosas que me apoiam e que acreditam em mim. Então, essa não é uma conquista só minha: também é de vocês. Nos momentos mais desafiadores, eu sempre lembro que não estou fazendo isso apenas por mim.

Bom... Eu acho que essa é a parte mais complexa da dissertação, parafraseando Fernanda Pires. Aqui devo escrever com o coração e eu geralmente não sou muito bom com essas coisas, mas prometo que tentarei. Provavelmente não conseguirei colocar em palavras, mas saibam que carrego comigo cada um de vocês.

Agradeço ao amor da minha vida, minha companheira de todos os momentos, Jeniffer Macena, pelo amor e suporte incondicionais. Você é inspiradora, até para a escrita desses agradecimentos. Você ter entrado na minha vida foi um acontecimento divino. Eu aprendi tanto com você, que nem caberia escrever aqui. Você me ajuda a acreditar em mim mesmo todos os dias e a continuar seguindo em frente, como fez Rock Lee. Cada momento com você é valioso e me ajuda a me tornar uma pessoa melhor, sejam em viagens, no carro, nos rolês, nos restaurantes, em qualquer lugar! Sou muito feliz com você e muito sortudo em tê-la comigo. É uma honra poder ter graduado, entrado no mestrado, estar se formando e ser aprovado no doutorado junto a você. Obrigado por sempre ficar do meu lado, mesmo quando eu ativo meu “modo chato”. Eu te amo muito, que Deus sempre nos proteja e abençoe!

Sou muito grato à minha família, que me deu um lar, alimento, educação, amor e as oportunidades para eu poder estar aqui. Vó Magali, quem seria eu sem você e a mamãe? Mesmo com tantos "depois eu faço, vó", você nunca deixou de estar ao meu lado. Não mediu esforços para me ajudar, seja me dando dinheiro para ir de Uber até a universidade, fazendo marmitas para o almoço, emprestando dinheiro quando eu ficava sem bolsa, me presenteando com um carro que parcelou de tantas vezes, preparando um chá de erva-doce quando eu ficava mal e inúmeras outras coisas... E tudo isso sem pedir nada em troca. Eu gostaria que todos pudessem ter uma vó tão maravilhosa quanto você. Mãe Márcia, herdei muita coisa de você: a escrita, a cordialidade, a beleza (embora eu seja mais bonito), o respeito, as piadas arianas e a coragem para seguir em frente mesmo quando as coisas parecem complicadas. Obrigado por cuidar das nossas gatinhas quando estamos em eventos científicos, por sempre estar no aeroporto para nos receber junto a vovó e ao vovô, por me incentivar a fazer exercícios físicos (tô tentando!) e por continuar me amando mesmo quando eu bagunço seu cabelo. Mãe e vó, sou eternamente grato e prometo sempre continuar sendo um rapaz respeitoso e que ajuda as pessoas. Aquele menino loirinho cresceu e agora está virando mestre, acreditam? Sei que posso ser difícil às vezes, mas isso não fez vocês me amarem menos. Amo muito vocês!

Vô Antônio, você nunca deixou de me ajudar, mesmo quando o pneu do meu carro furava, ou quando eu precisava trocar o óculos, ou até mesmo quando estávamos atrasados e precisávamos ir para o aeroporto. Eu não sei se você lembra, mas você que me deu carona na primeira vez que saí com a prof.ª Fernanda e companhia. Te amo, você faz parte dessa história! Minhas irmãs, Ludmila e Paola, guardo com muito carinho as memórias de quando éramos crianças, principalmente as viagens. De vez em quando me pego pensando nisso e percebo o quanto o tempo passa rápido. Agora vocês cresceram e estão literalmente maiores que eu. O mano está orgulhoso por tudo o que vocês conquistaram. Amo vocês, obrigado por continuarem do meu lado, ainda que não morando mais perto. Tio Fábio, você é o irmão que eu nunca tive. Gosto muito da sua companhia, da tia Karlinha e do Luke, seja em casa, no sítio ou em qualquer outro lugar. As conversas sempre são boas, assim como os "tira-gosto" que não podem

faltar. Obrigado por tudo, amo vocês! Vó Sônia, tenho um carinho enorme por você. Obrigado por sempre orar por mim, suas orações são poderosas! Gosto muito de visitá-la, conversar sobre a vida e implicar com os passarinhos. Te amo! Vô Roberto, Beto e Gabriel, estendo meus agradecimento à vocês! Pai Lincoln... Embora distantes, aprendi muito com você. Obrigado por ter cuidado de mim e por todo o apoio. Tia Marly, nos conhecemos há pouco tempo mas já lhe considero muito. Agradeço pela receptividade em Juiz de Fora, já estou com saudades e prometo que em breve passamos por aí de novo! Estendo meus agradecimentos para a Kalu e Teo!

ThinkTEd Lab, vocês são minha casa. É uma honra poder fazer parte disso tudo. Vocês me acolheram, cuidaram de mim e me proporcionaram experiências maravilhosas que eu jamais esquecerei. Eu não adquiri só conhecimentos técnicos, como também lições pra vida toda. Eu ganhei uma família que eu sei que sempre poderei contar. Cada viradão, lanches em conjunto, idas ao supermercado, mudanças de espaço, *researchs*, conversas descontraídas, viagens para congressos e diversos outros momentos têm um valor enorme pra mim que dinheiro nenhum paga.

Profª Fernanda, tudo isso começou com a Sra, lá em 2017. Eu não sabia exatamente o que eu queria da minha vida e, mesmo tantas vezes tentando sair do curso, a Sra me guiou como uma bússola. Sua vontade de ajudar as pessoas e o amor que tem pela educação me inspiraram. Aprendi a amar a Licenciatura em Computação e o crédito é todo seu. Eu realmente acredito que a Sra já tenha a começado a mudar o mundo: você me transformou em uma pessoa melhor e carrego comigo seus sonhos. "Sonho que se sonha sozinho é só um sonho. Mas sonho que se sonha junto vira realidade", não é? Profª Marcela, a Sra é uma pessoa iluminada. Não deixou de acreditar em mim mesmo em momentos complicados, sempre fui bem recebido em sua casa e me divirto muito quando viajamos juntos. Fico muito feliz quando recebo seus elogios, me sinto mais confiante e capaz de seguir em frente. A Sra me fez evoluir consideravelmente como pesquisador: colocar crase corretamente (é difícil), incluir as contribuições do trabalho no resumo, não fazer média de escala Likert, adicionar as limitações nas considerações finais, comparar os achados obtidos com resultados da literatura e inúmeras outras coisas. Prometo que ainda vou lhe dar um MSL. Fernanda Pessoa e Marcela Pires, obri-

gado por aguentarem esse menino teimoso. Estou aqui hoje porque vocês acreditam em mim. Vocês são minhas mães e me ensinaram a voar. Digo "ensinar" porque realmente aprendi (aí pode, né Fefa?). Obrigado pelos conselhos, orientações, puxões de orelha e conversas. Espero um dia poder ser para outras pessoas o que vocês são pra mim.

Profª Elaine, é um prazer trabalhar com a Sra. A "orientadora da minha orientadora é minha orientadora" kskssksks. Nossos cafés foram necessários para a evolução da pesquisa e definição dos próximos passos, bem como para conhecer um pouco mais dessa nossa cidade maravilhosa. E chega até ser engraçado, às vezes costumamos nos ver mais em eventos em outros estados do que em Manaus hahaha. Agradeço pela confiança que a Sra tem em mim, pela oportunidade de ser seu orientando, por nossas conversas inspiradoras, pelos cafés, por aquela caneta maravilhosa de Harvard e por fazer parte dessa conquista.

Thinkers, tudo é mais divertido com vocês. As conversas, as brincadeiras, as risadas, as reuniões e as pesquisas são fundamentais para mim. Agarram-se se à essa oportunidade de ouro e não larguem. Poderia agradecer a cada um de vocês de formas distintas, mas essa seção se prolongaria mais ainda. Então, de forma geral, meu muito obrigado a: Adriana, Alexander, Airton, Amora, Antônia, Bena, Carlos, Clei, Cris, Davi, Defala, Douglas, Duda, Dryco, Flávia, Handryo, Henry, Higor, Isabela, JB, Jean, José, Julya, Manu, Nathalia, Nezi, Rafa, Raquel, Samira, Sarah, Lucas, Tiago, Trovão, Valdenei, Victor, Wal, prof.ª Márcia, prof Osvaldo e prof David. Agradecimento também ao Léo Portes e Drica, pelo apoio e excelente recepção toda vez que visito vocês. Aos integrantes que tiveram passagem pelo Think e hoje trilham seus próprios caminhos: vocês também foram importantes nessa jornada, muito obrigado! Também agradeço ao ThinkNexus, coordenado pelos professores Fernanda, Marcela e Choji. O apoio proporcionado foi fundamental, seja para trabalhar em um ambiente confortável e com um computador de qualidade, almoçar na cozinha, deitar nos *puffs* da descompressão ou fazer reuniões. É uma conquista ter uma infraestrutura tão maravilhosa como essa e fico feliz em fazer parte dela.

Essa ferramenta não foi feita sozinha, houveram dias de luta. E agora chegaram os dias de GLORIA (essa não podia faltar, hehe)! Profª Fernanda e prof.ª Marcela,

obrigado por me guiarem em cada passo da pesquisa. Se hoje eu sei fazer ciência, foi porque eu aprendi com vocês. Gratidão imensa à força motriz que me auxiliou no desenvolvimento da ferramenta quando eu precisei: Defala, fico feliz de acompanhar sua evolução. Quantas noites você ficou em claro para que o GLORIA pudesse chegar onde está hoje? Davi, mesmo eu te pedindo esporadicamente por ajustes mínimos nas interfaces, você não me deixou na mão. Clei, ainda que atolada com atividades da especialização você quis me ajudar. Você é uma amiga maravilhosa, obrigado pela sua companhia. Henry, quedas de energia e excesso de cache no servidor ocorreram, mas logo em seguida você resolveu e deixou a aplicação no ar. Valdenei, a refatoração no *back-end* foi fundamental e ficou excelente. José e Manu, o empenho de vocês com a elaboração das telas no *front* é notável. E Victor, seus conhecimentos sobre LLM contribuíram não só para a ferramenta, mas para a evolução do Defala. Todos vocês fizeram um trabalho maravilhoso e fiquei muito feliz com o resultado. Deixo o meu agradecimento também à todos que participaram das avaliações ao longo da pesquisa.

Prof Luiz e prof Rafael, é um prazer tê-los na minha banca de dissertação. Muito obrigado pela disponibilidade em poder contribuir com a nossa pesquisa. Suas contribuições foram fundamentais para refiná-la, assim como para que eu me desenvolvesse mais como pesquisador. Os mimos entregues em Curitiba são apenas uma pequena forma de agradecer. Prof Esteban, seus apontamentos também foram pertinentes e auxiliaram na evolução do trabalho. Prof David, agradeço por suas contribuições, pois tanto os comentários feitos durante a defesa final quanto os apontamentos na dissertação foram valiosos e auxiliaram no aprimoramento do trabalho.

Agradeço à Universidade Federal do Amazonas (UFAM) e ao Instituto de Computação (IComp) pelo suporte e pela oportunidade em ingressar no Programa de Pós-Graduação em Informática (PPGI), cuja excelência e qualidade são inquestionáveis. Gratidão ao corpo docente do programa que lecionou disciplinas primordiais, onde pude aprender e aprimorar a pesquisa. Em especial, gostaria de agradecer ao Frank, da secretaria do PPGI. Você me ajudou tantas vezes e sempre manteve sua paciência, cordialidade e respeito em resposta aos meus e-mails. Além disso, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Finan-

ciamento 001. Este trabalho foi parcialmente financiado pela Fundação de Amparo à Pesquisa do Estado do Amazonas – FAPEAM – por meio do projeto POSGRAD. O apoio recebido, além de possibilitar minha participação em eventos renomados de educação e tecnologia, contribuiu para que eu conhecesse localidades e pessoas maravilhosas. Nesse sentido, também agradeço à Helen da Secretaria de Projetos do IComp, que me deu suporte na solicitação de recursos para eventos científicos.

Gostaria também de agradecer às famílias Macena (Milena, Mildrey, Tinho, Júnior, Mateus, Jennifer e Laurinha) e Souza (Zenilton, Elisa, Neto, Milena e Rebeca) pelo apoio e pelos bons momentos que compartilhamos juntos. Agradeço ao seu Zequinha, seu André e dona Zena, que descansam em paz no reino dos ceús e continuam a olhar pela Jenny. O carinho e amor de vocês nos rodeiam.

Sou grato também aos meus amigos e amigas, que me trouxeram muitos momentos divertidos e contribuições pra vida: “Sushitzu”, “soloCORE”, “Rolês da Massa”, “Academigos”, “Irmandade”, que conheci em eventos, da universidade, da escola e da vida. Agradeço ao curso de Licenciatura em Computação da UEA e ao corpo docente, que contribuíram para minha evolução profissional.

Reservo este espaço também para agradecer àqueles que estão noutro plano e nos acompanham dos ceús: bisa Naomi, bisa Luci, tia Rosa, tio Marco, dona Helena e seu Carlão.

Com certeza eu não podia deixar de agradecer aos meus filhos e filhas: Shiro, a primogênita. Mesmo você sendo briguenta, o papai ama. Ma Fiinha e Xixia, vocês são minhas queridas, sempre motivam o papai a continuar seguindo em frente e sei que a mamãe também sente o mesmo. Também agradeço aos pets: Sukurimo, Tiga, Luvinhas, Nharlie, Último do estoque, Gertrudes, Dalila, Celinho, Pin, Serelepinha e Mils. Meus agradecimentos à CRP, Grey, Bebas, Snoop, Serelepe, Sansão, Xisa, Xis, Jones, Óculos, Nery, Atem, Grilin, Peepeewee, Elisa, Genoveva, Genesino, Vladimir, Felícia, Abel e aos demais pets que agora descansam ou que não estão mais próximos de nós. Todos vocês, em inúmeras ocasiões da minha vida, me deram forças para continuar seguindo em frente.

Somos como o sangue em nossas veias. Devemos manter o fluxo, fornecendo oxigênio para que o cérebro funcione corretamente. Não podemos parar.

Kuroo Tetsurō (Haikyu!!)

GLORIA: ferramenta interativa integrada a um *Large Language Model* para auxiliar na modelagem de dados e implementação de técnicas de *Game Learning Analytics*

Autor: Fabrizio Honda Franzoia

Orientadora: Elaine Harada, D.Sc.

Coorientadoras: Fernanda Gabriela de Sousa Pires, D.Sc.

Marcela Sávia Picanço Pessoa Bastos, D.Sc.

Resumo

Game Learning Analytics (GLA) é um conjunto de técnicas para coletar e analisar dados de interação do jogador (*logs*), permitindo rastrear seu percurso no jogo e identificar evidências de aprendizagem. O GLBoard é um modelo que viabiliza a implementação dessas técnicas, por meio de um *template* de dados (JSON) genérico. Entretanto, *designers* de aprendizagem, sobretudo estudantes que atuam nessa função em disciplinas, podem apresentar dificuldades na modelagem de dados para jogos educacionais: elencar e justificar os dados que serão coletados, conforme os objetivos de aprendizagem do jogo. A definição dos dados orienta a criação e organização das variáveis de GLA em um *template* de dados, sendo a base para uma coleta de dados eficiente. Nesse contexto, a Inteligência Artificial Gerativa (GAI), sobretudo com os *Large Language Models* (LLMs), pode ser uma alternativa. Esses modelos possuem capacidades de gerar textos em linguagem natural e realizar tarefas como seres humanos, aos quais já foram utilizados em contextos de *Learning Analytics* e *Game Analytics*. Diante disso, esta dissertação investiga a modelagem de dados de jogos educacionais e as implicações do uso de LLMs para o campo de GLA, operacionalizando os achados por meio da proposta e avaliação uma

ferramenta integrada a um LLM *open-source*, intitulada GLORIA (*Game, Learning, Observation, Recommendation, Intelligence e Analytics*). Seu propósito é auxiliar na modelagem de dados em GLA, sendo futuramente incorporada ao sistema GLBoard como módulo introdutório. A pesquisa adota a abordagem metodológica-epistemológica de *Design Science Research* (DSR), cujos resultados estão relacionados a seus ciclos: (i) Relevância: revisões da literatura e investigações sobre desafios em modelagem de dados e uso de LLMs para GLA, bem como a realização de um *Design* Participativo; (ii) *Design*: criação e avaliação de um agente inteligente em GLA, processo de construção da ferramenta utilizando uma metodologia de *game design* educacional e avaliação do GLORIA; e (iii) Rigor: composição de uma base de conhecimento a partir dos achados da literatura e das investigações. A ferramenta foi avaliada por meio de uma avaliação heurística (especialistas em UX/UI) e de um estudo empírico (*designers* de aprendizagem). Os resultados demonstram indícios do GLORIA para apoiar a modelagem de dados de jogos educacionais, bem como apontam que LLMs para GLA podem atuar como mediadores de conhecimento, estimulando a reflexão de *designers* de aprendizagem (*scaffolding*). Entretanto, refinamentos na ferramenta ainda são necessários, sobretudo em relação à estética e à interação com o *chatbot*. A principal contribuição desta dissertação é ampliar discussões sobre a modelagem de dados para jogos educacionais e investigar o uso de LLMs para apoiá-la, resultando em achados valiosos e em uma ferramenta que orienta o *designer* de aprendizagem nos primeiros passos em GLA, estimulando-o a refletir sobre essa etapa e, potencialmente, a aprender a modelar dados.

Palavras-chave: *Game Learning Analytics*, GLBoard, *Large Language Models*, modelagem de dados, ferramenta.

GLORIA: ferramenta interativa integrada a um *Large Language Model* para auxiliar na modelagem de dados e implementação de técnicas de *Game Learning Analytics*

Autor: Fabrizio Honda Franzoia

Orientadora: Elaine Harada, D.Sc.

Coorientadoras: Fernanda Gabriela de Sousa Pires, D.Sc.

Marcela Sávia Picanço Pessoa Bastos, D.Sc.

Abstract

Game Learning Analytics (GLA) encompasses a set of techniques for collecting and analyzing player interaction data (*logs*), making it possible to track players' progress throughout a game and identify evidence of learning. GLBoard is a model that enables the implementation of these techniques through a generic JSON data *template*. However, learning *designers*, particularly students who perform this role in academic courses, may face difficulties when modeling data for educational games, such as selecting and justifying the data to be collected according to the game's learning objectives. Data definition guides the creation and organization of GLA variables within a data *template* and provides the foundation for efficient data collection. In this context, Generative Artificial Intelligence (GAI), particularly *Large Language Models* (LLMs), may offer an alternative. These models can generate natural-language texts and perform human-like tasks, and researchers have already used them in *Learning Analytics* and *Game Analytics* contexts. Accordingly, this dissertation investigates data modeling for educational games and the implications of using LLMs in the GLA field. It operationalizes its findings through the proposal and evaluation of a tool integrated with an *open-source* LLM, named GLORIA

(*Game, Learning, Observation, Recommendation, Intelligence, and Analytics*). Its purpose is to support data modeling in GLA, and it may later be incorporated into the GLBoard system as an introductory module. The research adopts the methodological and epistemological approach of *Design Science Research* (DSR), and its results relate to the following cycles: (i) Relevance: literature reviews and investigations into the challenges of data modeling and the use of LLMs for GLA, as well as the implementation of a Participatory *Design* process; (ii) *Design*: the creation and evaluation of an intelligent agent for GLA, the development of the tool using an educational *game design* methodology, and the evaluation of GLORIA; and (iii) Rigor: the development of a knowledge base derived from the literature findings and the investigations. The tool was evaluated through a heuristic evaluation involving UX/UI experts and an empirical study involving learning *designers*. The results provide evidence that GLORIA can support data modeling for educational games. They also indicate that LLMs for GLA can act as knowledge mediators by encouraging learning *designers* to reflect on their decisions through *scaffolding*. However, the tool still requires refinement, particularly regarding its visual design and interaction with the *chatbot*. The main contribution of this dissertation is to broaden discussions on data modeling for educational games and investigate the use of LLMs to support this process. The research produced valuable findings and a tool that guides learning *designers* through their first steps in GLA, encourages them to reflect on this stage, and may help them learn how to model data.

Keywords: *Game Learning Analytics*, GLBoard, *Large Language Models*, data modeling, tool.

LISTA DE ILUSTRAÇÕES

Figura 1 – Relação entre capítulos, objetivos específicos e investigações. 33
Figura 2 – Arquitetura para um sistema de GLA, adaptado de Freire et al. (2016). 41
Figura 3 – *Template* de dados do GLBoard. 44
Figura 4 – Etapas para a busca por trabalhos relacionados. 52
Figura 5 – Diagrama de Venn sobre a localização do trabalho. 55
Figura 6 – Etapas da pesquisa sobre modelagem de dados. 61
Figura 7 – *Boxplot* com a avaliação das questões quantitativas. 64
Figura 8 – Estrutura JSON de captura e exemplo, respectivamente. 73
Figura 9 – Estruturas geradas, respectivamente, com o ChatGPT 3.5, 4 e 4o. . . . 74
Figura 10 – *Boxplot* com as avaliações gerais de cada estrutura. 75
Figura 11 – Etapas da pesquisa sobre *Prompting* em LLM para GLA 80
Figura 12 – Resultados das avaliações realizadas pelos especialistas. 84
Figura 13 – Etapas da investigação sobre o agente Inteligente em GLA. 90
Figura 14 – Agente "*GLA Specialist*". 94
Figura 15 – Etapas da investigação sobre aplicação do Agente Inteligente. 98
Figura 16 – *Boxplot* das questões quantitativas avaliadas. 100
Figura 17 – Comparação do nível de dificuldade na modelagem de dados. 102
Figura 18 – Equipes no *Design* Participativo. 106
Figura 19 – Propostas de sistemas para modelagem de dados. 107
Figura 20 – Metodologia de *game design* educacional (Pires et al., 2021). 109
Figura 21 – *Brainstormings* do GLORIA. 109
Figura 22 – Protótipos: baixa, média fidelidade e POC, respectivamente. 112

Figura 23 – Fluxo de GLA, ilustrando a atuação do GLORIA e do GLBoard. . . . 114
Figura 24 – Variáveis do jogo “Cadê minha pizza?”. 116
Figura 25 – Arquitetura do GLORIA. Modelo C4 - Nível de Contêiner. 120
Figura 26 – Fluxo de telas do GLORIA. 120
Figura 27 – *Landing page* da ferramenta. 121
Figura 28 – Manual de uso. 122
Figura 29 – Página de “Sobre nós”. 123
Figura 30 – Cadastro de usuário. 123
Figura 31 – *Login* da ferramenta. 124
Figura 32 – Tela de “Gestão de Jogos”. 124
Figura 33 – Cadastro de jogo no GLORIA. 125
Figura 34 – Página específica de um jogo, aba de “Dados do Jogo”. 126
Figura 35 – Aba de “Agente Inteligente”, no primeiro acesso. 127
Figura 36 – Tela de interação com o agente inteligente (*chatbot*). 128
Figura 37 – Tabela de modelagem de dados do jogo específico. 128
Figura 38 – Estrutura JSON gerada, conforme a modelagem. 129
Figura 39 – *Heatmap* com as avaliações referentes às heurísticas de Nilsen. 137
Figura 40 – Nível de experiência em modelagem de dados e implementação de GLA, por participante. 142
Figura 41 – Avaliações do instrumento “*Player Level Up!*”. 146
Figura 42 – Variáveis de GLA propostas com o GLORIA. 149
Figura 43 – Proposta de arquitetura de GLA com base em Freire et al. (2016). . . 166
Figura 44 – Telas do jogo “Cadê minha pizza?”. 194
Figura 45 – Diagrama de Caso de Uso do GLORIA. 217
Figura 46 – Diagrama de Entidade-Relacionamento (DER) do GLORIA. 218

LISTA DE TABELAS

Tabela 1 – Ciclos da *Design Science Research* (DSR) e sua aplicação na pesquisa. . 32
Tabela 2 – Critérios de inclusão e exclusão 54
Tabela 3 – Objetivo da pesquisa, conforme a abordagem GQM. 68
Tabela 4 – Instrumento de avaliação das estruturas de captura no *template* GLBoard. 72
Tabela 5 – Estatística descritiva por estrutura. 76
Tabela 6 – Distribuição das Notas por Questão. 95
Tabela 7 – Questões e Respostas sobre o Agente. 95
Tabela 8 – Tabela de Requisitos Funcionais. 118
Tabela 9 – Atividades realizadas durante o teste com a ferramenta. 143
Tabela 10 – Instrumento de avaliação das estruturas de captura no *template* GLBoard. 144
Tabela 11 – Grupos de participantes e duração dos testes. 148
Tabela 12 – Matriz de confusão da classificação das variáveis 151
Tabela 13 – Evidências observadas na interação com o GLORIA (*Think Aloud*). . . 155
Tabela 14 – Lista de publicações associadas. 182

LISTA DE ABREVIATURAS E SIGLAS

API Application Programming Interface

BD Banco de Dados

BNCC Base Nacional Comum Curricular

DER Diagrama Entidade-Relacionamento

DP Design Participativo

DSR Design Science Research

EGDD Educational Game Design Document

FSE Fundamentos de Software Educacional

GA Game Analytics

GAI Generative Artificial Intelligence

GAN Generative Adversarial Networks

GLA Game Learning Analytics

GLORIA Game, Learning, Observation, Recommendation, Intelligence and Analytics

GQM Goal-Question-Metric

IA Inteligência Artificial

ICL In-Context Learning

JSON JavaScript Object Notation

LA Learning Analytics

LAM Learning Analytics Model

LLM Large Language Model

LRS Learning Record Store

MPS Modelagem e Projeto de Sistemas

PCN Parâmetros Curriculares Nacionais

PLMs Pre-Trained Language Models

PLN Processamento de Linguagem Natural

POC Proof of Concept

RAG Retrieval-Augmented Generation

SIMVA SIMple VAlidator

T-MON Teacher-Monitoring Tool

UEA Universidade do Estado do Amazonas

UFAM Universidade Federal do Amazonas

VAE Variational Autoencoders

xAPI Experience API

XASU xAPI Analytics Supplier

SUMÁRIO

1 INTRODUÇÃO 27
1.1 Questões de Pesquisa 30
1.2 Objetivos 31
1.2.1 Objetivo Geral 31
1.2.2 Objetivos Específicos 31
1.3 Metodologia 31
1.4 Organização do Trabalho 34

2 REFERENCIAL TEÓRICO 36
2.1 Jogos Educacionais 36
2.2 Avaliação de Aprendizagem em jogos 37
2.3 Game Learning Analytics - GLA 39
2.4 GLBoard 43
2.5 Modelagem de dados 45
2.6 IA Gerativa e *Large Language Models* 48
2.7 Trabalhos Relacionados 52
2.8 Síntese do Capítulo 57

3 INVESTIGANDO DIFICULDADES E O USO DE LLMS PARA GLA 59
3.1 Análise das dificuldades de estudantes de computação em modelagem de dados 59
3.1.1 Objetivo e participantes 60
3.1.2 Procedimentos 61
3.1.3 Coleta e análise de dados 62

3.1.4 Resultados . . . 63
3.1.5 Considerações sobre a investigação . . . 66
3.2 Geração de estruturas de captura de GLA: LLM vs especialista . . . 67
3.2.1 Escopo . . . 68
3.2.2 Planejamento . . . 68
3.2.3 Operação . . . 72
3.2.4 Análise e interpretação . . . 75
3.3 Investigação de uma técnica de *prompt* interativa de LLM para GLA . . . 78
3.3.1 Objetivo, contexto e participantes . . . 78
3.3.2 Instrumentos e procedimentos . . . 79
3.3.3 Coleta e análise de dados . . . 82
3.3.4 Resultados . . . 83
3.3.5 Considerações sobre a investigação . . . 85
3.4 Síntese do Capítulo . . . 86

4 CONSTRUINDO E AVALIANDO UM AGENTE ESPECIALISTA EM GLA . . . 88
4.1 Construção e Avaliação de Agente Inteligente em GLA . . . 89
4.1.1 Objetivo, contexto e participantes . . . 89
4.1.2 Procedimentos . . . 90
4.1.3 Coleta e análise de dados . . . 92
4.1.4 Resultados . . . 93
4.1.5 Considerações sobre a investigação . . . 96
4.2 Avaliação do Agente Inteligente em GLA por estudantes de computação . . . 97
4.2.1 Procedimentos . . . 97
4.2.2 Coleta e análise de dados . . . 99
4.2.3 Resultados . . . 99
4.2.4 Considerações sobre a investigação . . . 102

4.3 Síntese do Capítulo . . . 103

5 GLORIA: ANALISANDO JOGOS E ORIENTANDO A INSERÇÃO DE GLA . . . 105

5.1 *Design* Participativo . . . 105

5.2 Proposta de ferramenta . . . 108

5.2.1 Processo de desenvolvimento . . . 108

5.2.2 Ferramenta para modelagem de dados . . . 113

5.3 Síntese do Capítulo . . . 129

6 AVALIANDO O GLORIA PARA MODELAGEM DE DADOS EM GLA . . . 131

6.1 Avaliação heurística com especialistas em UX/UI . . . 131

6.1.1 Procedimentos . . . 133

6.1.2 Coleta e análise de dados . . . 134

6.1.3 Resultados . . . 135

6.1.4 Considerações sobre a avaliação . . . 139

6.2 Avaliação de suporte à modelagem de dados . . . 140

6.2.1 Contexto e participantes . . . 141

6.2.2 Procedimentos . . . 142

6.2.3 Coleta e análise dos dados . . . 144

6.2.4 Resultados . . . 145

6.2.5 Considerações sobre o estudo . . . 156

6.3 Síntese do Capítulo . . . 159

7 DISCUSSÕES DOS ACHADOS . . . 161

7.1 A modelagem de dados como desafio estrutural em *Game Learning Analytics* . . . 162

7.2 A aplicação de LLMs para contextos de aprendizagem . . . 165

7.3 GLORIA: concretizando os pressupostos teóricos . . . 173

8 CONSIDERAÇÕES FINAIS . . . 175

8.1 Contribuições 178
8.2 Limitações 179
8.3 Trabalhos futuros 180
8.4 Produção científica 181

Referências 183

APÊNDICE A DESCRIÇÃO DO JOGO “CADÊ MINHA PIZZA?” . 193

APÊNDICE B FORMULÁRIO DE AUTOAVALIAÇÃO DAS MODELAGENS DE DADOS 195

APÊNDICE C 1º PROMPT AO CHATGPT PARA GERAÇÃO DE ESTRUTURAS DE CAPTURA 197

APÊNDICE D ATRIBUTOS DE JOGOS EDUCACIONAIS MAPEADOS PARA ENVIAR AO CHATGPT 199

APÊNDICE E 2º PROMPT AO CHATGPT PARA GERAÇÃO DE ESTRUTURAS DE CAPTURA 201

APÊNDICE F FORMULÁRIO DE AVALIAÇÃO DA CRIAÇÃO DE ESTRUTURA DE CAPTURA DE FORMA MANUAL 202

APÊNDICE G FORMULÁRIO DE AVALIAÇÃO DA CRIAÇÃO DE ESTRUTURA DE CAPTURA COM O CHAT-GPT . 204

APÊNDICE H FORMULÁRIO DE AVALIAÇÃO DO AGENTE INTELIGENTE POR ESPECIALISTAS EM GLA . . . 207

APÊNDICE I AUTODESCRIÇÃO ATUAL DO AGENTE INTELIGENTE EM GLA 209

APÊNDICE J ROTEIRO DE TESTE PARA AVALIAÇÃO DO AGENTE INTELIGENTE 211

APÊNDICE K FORMULÁRIO DE AVALIAÇÃO DO AGENTE INTELIGENTE POR ESTUDANTES DE COMPUTAÇÃO 213

APÊNDICE L INFORMAÇÕES TÉCNICAS DO GLORIA 216

APÊNDICE M PROMPT BASE DO GLORIA 219

APÊNDICE N PROMPT DA TOOL DE GERAÇÃO DA MODELAGEM 224

APÊNDICE O FORMULÁRIO DE AVALIAÇÃO DO GLORIA: AVALIAÇÃO HEURÍSTICA 227

APÊNDICE P FORMULÁRIO DE AVALIAÇÃO DO GLORIA: ESTUDO EMPÍRICO 230

APÊNDICE Q INFORMAÇÕES DOS JOGOS EDUCACIONAIS . . 234

Q.1 WAlgor 234

Q.2 Optima Corporation 235

Q.3 Animal Park 236

Q.4 Matematicamente 237

Q.5 Two Cats 238

Q.6 O Livro do Conhecimento 239

Q.7 RoboHouse 240

Q.8 My Name 241

Q.9 Pensar e Lavar 242

Q.10 Código Turing 243

Q.11 EscaloMania 244

Q.12 Robô Euroi 245

Q.13 StarDust 246

Q.14 Hello Food 247

1

INTRODUÇÃO

Jogos são elementos lúdicos capazes de proporcionar motivação e engajamento aos jogadores (Plass; Homer; Kinzer, 2015). A experiência de jogar pode favorecer o estado de *flow*, caracterizado pelo envolvimento profundo em uma atividade que se torna intrinsecamente gratificante, entre outros aspectos, como sugerem estudos (Csikszentmihalyi; Csikzentmihaly, 1990; Liu, 2016; Piacentini, 2011). O uso de jogos como objetos de aprendizagem tem se intensificado, pois permitem o desenvolvimento de habilidades cognitivas, socialização, aprendizagem por descoberta, facilitação da aprendizagem, entre outros (Gundersen; Lampropoulos, 2025; Savi; Ulbricht, 2008).

Apesar dos inúmeros benefícios que os jogos proporcionam quando aplicados na educação, avaliá-los ainda é um desafio na literatura, sobretudo quanto aos aspectos de aprendizagem. Os questionários são uma das estratégias de avaliação mais comuns, que geralmente consideram apenas a percepção de aprendizagem dos jogadores (Silva et al., 2021) e costumam ser aplicados somente após a *gameplay*. Embora esses tipos de avaliação sejam necessários, não são suficientes para indicar se houve aprendizagem do jogador, pois não contemplam elementos que expressam com clareza sua evolução (Melo et al., 2020), tais como: tomadas de decisão, registros de tempo, interação com as interfaces, inatividade, entre outros. Esses dados são coletados durante o jogo (*in-game*).

Nesse aspecto, tem-se o campo de *Game Learning Analytics* (GLA) (Freire et al., 2016), também denominado de *Learning Analytics* para *Serious Games*[1], uma intersecção

[1] São os termos mais comuns. Entretanto, observou-se na literatura o uso de nomenclaturas distintas, como *learning behavior analysis*, *gameplay telemetry*, *data-driven techniques*, *analytics for game-based learning*, entre outros.

entre *Game Analytics* e *Learning Analytics*. Trata-se da implementação de técnicas para capturar, analisar e visualizar registros de interação dos jogadores (*logs*), cuja coleta ocorre de forma não intrusiva para não interromper a concentração do jogador e em conformidade com a legislação de proteção e privacidade dos dados. O uso de técnicas de GLA vem se intensificando ao longo dos anos (Banihashem et al., 2024), permitindo rastrear o percurso do jogador durante as fases. Em jogos educacionais (*Educational Serious Games*), GLA é fundamental, pois a análise das interações dos jogadores pode fornecer evidências sobre a aprendizagem e o desempenho dos estudantes com base em ações e eventos registrados no jogo, tempo de jogo, níveis e objetivos alcançados e tomadas de decisão (Cano; Fernández-Manjón; García-Tejedor, 2017; Chung, 2015).

Após a coleta dos dados, geralmente utilizam-se *dashboards* (estáticos ou interativos) para apoiar a visualização e interpretação dos *logs* (Scheneider; Lemos, 2020). Com isso, os educadores podem analisar as informações provenientes de GLA e ter *insights* sobre os jogadores. Por exemplo, podem identificar que os estudantes de uma turma apresentam dificuldade em um determinado conteúdo do jogo e aplicar uma intervenção em sala de aula para exercitá-lo. Além disso, os desenvolvedores podem localizar falhas de *game design* e realizar melhorias em seus jogos educacionais (Melo et al., 2020). Dentre os modelos que viabilizam a implementação de técnicas de GLA, destacam-se o XASU (antigo xAPI-SG) do ecossistema SIMVA (Santilario-Berthilier et al., 2025) e o GLBoard (Silva et al., 2022), descritos no Capítulo 2.

Um dos atores fundamentais para que as técnicas de GLA possibilitem a identificação de evidências de aprendizagem é o *designer* de aprendizagem. Um de seus objetivos é definir as mecânicas de aprendizagem, que consistem na organização do *level design* de modo a proporcionar o exercício do conteúdo educacional em consonância com os objetivos de aprendizagem (Plass; Homer; Kinzer, 2015). Com base nessa articulação entre os aspectos pedagógicos e o *game design*, ele precisa definir os dados a serem coletados para que, após analisados, permitam informar em que medida os objetivos foram alcançados (Perez-Colado et al., 2018; Shoukry; Göbel; Steinmetz, 2014).

Essa etapa denomina-se modelagem de dados[2], que orienta a criação e a inserção das variáveis de GLA em um *template* de dados, sendo a base para o fluxo de GLA.

[2] Termo adotado nesta dissertação. Outras nomenclaturas são discutidas na Seção 2.5.

Ela deve ocorrer desde o início da concepção do jogo e de forma clara (Hauge et al., 2014; Alonso-Fernández et al., 2021; Kitto et al., 2020). Entretanto, embora a modelagem constitua uma etapa fundamental de GLA, observa-se que ela geralmente não é tratada como foco central dos trabalhos da área. A maior parte dos estudos concentra-se no *design* e na avaliação de jogos, na proposta de *frameworks*, na análise de comportamento e na realização de inferências sobre o desempenho e a aprendizagem a partir dos dados coletados (Alonso-Fernández et al., 2019; Tlili et al., 2021).

Além disso, a literatura indica que a modelagem de dados é desafiadora, pois exige traduzir competências em indicadores comportamentais contextualizados, identificar, entre numerosas interações, quais variáveis e padrões podem constituir evidências de aprendizagem e definir quais dados devem ser coletados para informar o alcance dos objetivos, articulando-os ao *design* de aprendizagem do jogo (Petersen; Bedek, 2012; Kerr; Chung, 2012; Cano; Fernández-Manjón; García-Tejedor, 2017; Perez-Colado et al., 2018). Essa questão torna-se relevante em contextos acadêmicos, nos quais estudantes podem atuar como *designers* de aprendizagem na construção de jogos educacionais (Earp, 2015; Weitze, 2014; Mercer et al., 2017).

Nesse contexto, pressupõe-se que a complexidade da modelagem de dados possa se intensificar, visto que esses estudantes ainda se encontram em processo de formação e podem não possuir os conhecimentos necessários para realizá-la. No entanto, não foram localizados trabalhos que investiguem especificamente esse aspecto. Essa lacuna pode estar relacionada ao fato de que, geralmente, os estudos da área de GLA se concentram em contextos nos quais se dispõe de profissionais ou equipes multidisciplinares (Folkestad et al., 2015; Cano; Fernández-Manjón; García-Tejedor, 2017; Perez-Colado et al., 2018), sem considerar estudantes em formação.

Uma área emergente que pode mitigar as dificuldades de modelagem de dados é a Inteligência Artificial Generativa (GAI) (Jovanovic; Campbell, 2022), sobretudo os *Large Language Models* (LLMs), capazes de trabalhar com textos e realizar diversas atividades como seres humanos (Kasneci et al., 2023). Nesses modelos, o usuário geralmente interage com um agente de conversação, ou *chatbot* de IA (Dam et al., 2024), por meio de instruções (*prompts*). Com base nelas, o modelo gera resultados em linguagem natural a

partir de padrões aprendidos durante o treinamento.

Os LLMs vêm sendo utilizados em contextos diversos, cuja aplicação se expandiu para as áreas de *Learning Analytics* (Misiejuk et al., 2025) e *Game Analytics* (Yang; Kleinman; Harteveld, 2025). Entre as principais aplicações nessas áreas estão a análise dos dados coletados durante as interações do usuário com o ambiente de aprendizagem ou jogo, bem como a adaptação de conteúdos e *feedbacks* com base nesses dados (Yang; Kleinman; Harteveld, 2025; Peng et al., 2024; Davalos et al., 2025; Rodríguez-Ortiz; Santana-Mancilla; Anido-Rifón, 2025). Portanto, pressupõe-se que os LLMs possam contribuir para o campo de GLA, que integra conceitos e técnicas de ambos os domínios. Uma das possibilidades, contemplada nesta dissertação, é minimizar a complexidade da modelagem de dados de jogos educacionais, assumindo um papel de "tutor" para orientar o *designer* de aprendizagem a definir as variáveis de GLA e suas justificativas, conforme os objetivos de aprendizagem do jogo (Honda et al., 2025).

Diante disso, considerando as dificuldades de *designers* de aprendizagem (educadores, desenvolvedores e estudantes em disciplinas que assumem essa função) em modelar dados para identificar indícios de aprendizagem em jogos educacionais e o potencial dos LLMs para auxiliar no campo de GLA, esta dissertação investiga o estado da arte a respeito de modelagem de dados e as implicações de uso dos LLMs para esse processo, operacionalizando os achados por meio da construção de uma estratégia para auxiliar *designers* de aprendizagem, denominada GLORIA (*Game, Learning, Observation, Recommendation, Intelligence and Analytics*). Futuramente, pretende-se incorporar essa estratégia ao GLBoard para atuar como módulo introdutório, que foi desenvolvido no grupo de pesquisa do mestrando e não contempla a modelagem de dados para GLA.

1.1 Questões de Pesquisa

Esta dissertação apresenta como questão de pesquisa (QP):

QP: Como auxiliar *designers* de aprendizagem a modelar dados de jogos educacionais para que possam implementar técnicas de *Game Learning Analytics* e identificar evidências de aprendizagem?

1.2 Objetivos

A partir da questão de pesquisa, esta dissertação apresenta como objetivos:

1.2.1 Objetivo Geral

Caracterizar o estado da arte sobre modelagem de dados para jogos educacionais e uso de LLMs para GLA, operacionalizando os achados na construção de um artefato de apoio a *designers* de aprendizagem.

1.2.2 Objetivos Específicos

- Consolidar um corpo de conhecimento sobre avaliação de aprendizagem em jogos educacionais, *Game Learning Analytics*, *Large Language Models* e Engenharia de *Prompt*, por meio de uma *scoping review* e de buscas exploratórias na literatura;
- Caracterizar dificuldades e estratégias utilizadas na modelagem de dados e em projetos de *prompts* para LLMs por *designers* de aprendizagem;
- Desenvolver uma proposta (prova de conceito - POC) que dê suporte a *designers* de aprendizagem na modelagem de dados;
- Construir uma ferramenta para auxiliar *designers* de aprendizagem a modelar dados em GLA, a partir dos achados das investigações anteriores.

1.3 Metodologia

A metodologia desta pesquisa está fundamentada em *Design Science Research* (DSR) (Hevner, 2007), uma abordagem epistemológica-metodológica a qual o pesquisador compromete-se a atingir dois objetivos: (i) resolver um problema prático em um contexto específico a partir da construção de um artefato e (ii) produzir novo conhecimento científico (Pimentel; Filippo; Santoro, 2019). Desse modo, os resultados gerados com

a DSR são conhecimentos em *design science* que assumem a forma de artefatos, sejam constructos, modelos, *frameworks*, entre outros (Vaishnavi; Kuechler, 2015).

Nesta dissertação, propõe-se uma ferramenta, classificada como uma *instantiation*: uma implementação situada em um ambiente, que pode ou não operacionalizar artefatos abstratos. Ela pertence à categoria de "melhoria" (*improvement*): novas soluções para problemas já conhecidos (Vaishnavi; Kuechler, 2015). Portanto, alinhada com a DSR, esta pesquisa propõe a criação de uma ferramenta integrada a um LLM (nova solução) para minimizar a complexidade da modelagem de dados de GLA em jogos educacionais (problema já conhecido).

A DSR contém três etapas (ciclos), que se relacionam de maneira cíclica. A Tabela 1 descreve cada um dos ciclos e seus artefatos resultantes em cada etapa da pesquisa desta dissertação (Hevner, 2007; Pimentel; Filippo; Santoro, 2019; Pessoa, 2022).

Tabela 1 – Ciclos da *Design Science Research* (DSR) e sua aplicação na pesquisa.

Ciclo	**Descrição**	**Resultado na pesquisa**
Relevância	Contexto da pesquisa: requisitos a serem alcançados, problema abordado e critérios de aceitação para a avaliação do artefato.	Revisão da literatura sobre GLA e LLM (Capítulo 2), investigações (Capítulo 3) e condução de um *Design* Participativo (Capítulo 5).
Design	O núcleo de qualquer projeto com DSR: projeto e desenvolvimento do artefato para solucionar os problemas identificados no Ciclo de Relevância, bem como avaliações e refinamento que iteram frequentemente.	Realização de investigações para criar e avaliar um agente inteligente em GLA (Capítulo 4), condução de uma metodologia de *game design* educacional (Pires et al., 2021) para proposta de uma ferramenta, consolidação do GLORIA (Capítulo 5) e avaliações (Capítulo 6).
Rigor	A geração da base de conhecimento, que fornece conhecimento prévio à pesquisa para garantir inovação e provém rigor teórico/metodológico aos aspectos identificados no Ciclo de Relevância e construídos no Ciclo de *Design*.	Composição da base de conhecimento: jogos, *educational serious games*, avaliação de aprendizagem em jogos educacionais, GA, LA, GLA, LLM e Engenharia de *Prompt*. Essa base é constantemente atualizada mediante a elaboração do agente especialista em GLA (com o ChatGPT) e da ferramenta GLORIA, incrementando mais achados da literatura. Uma parte desses achados encontra-se no Capítulo 2.

Alinhado com a DSR para sustentar a pesquisa, inclui-se a condução de investigações que resultam em requisitos e subsídios teóricos para a ferramenta GLORIA. A Figura 1 ilustra como essas investigações estão relacionadas com os capítulos desta dissertação e seus objetivos específicos.

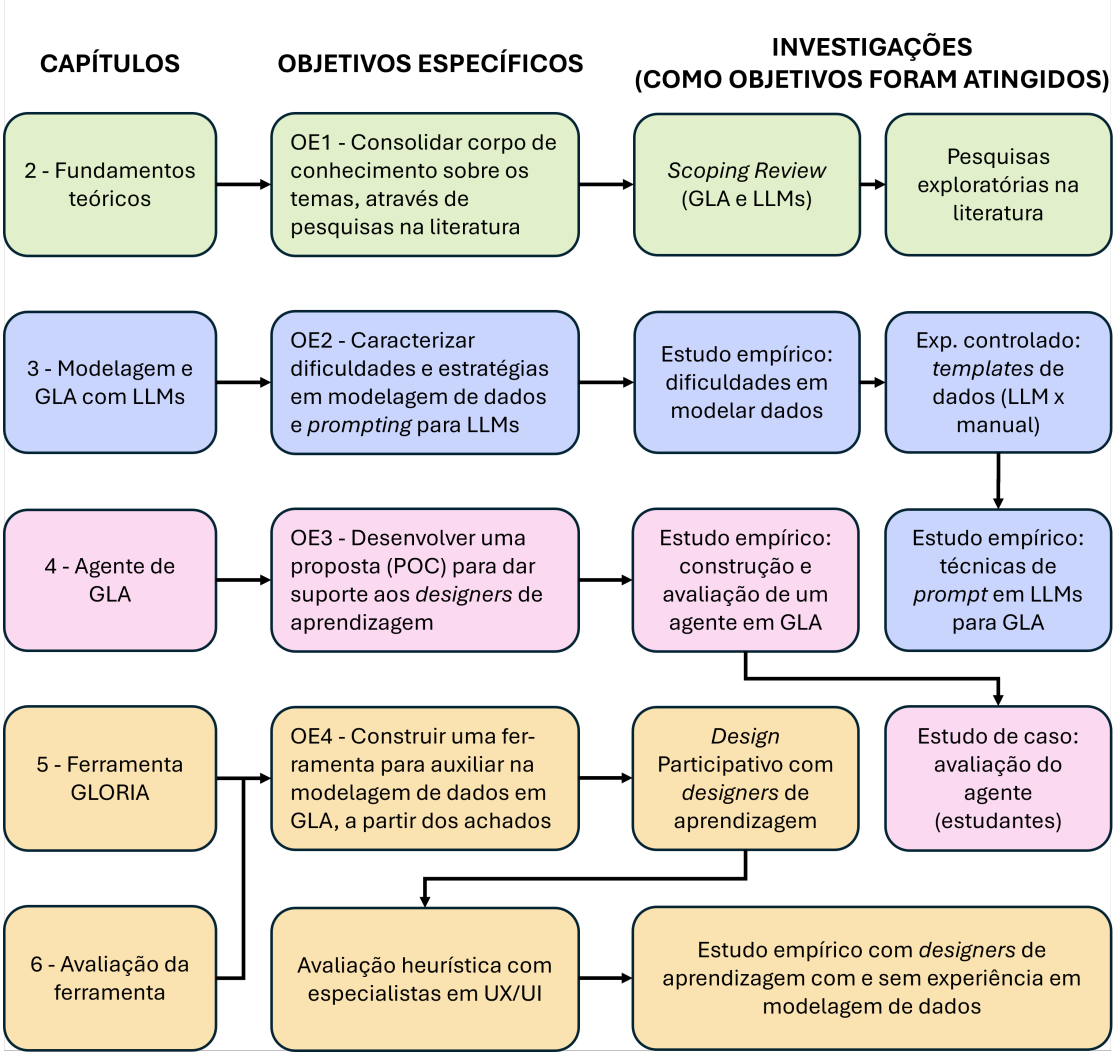


Figura 1 – Relação entre capítulos, objetivos específicos e investigações.

O Capítulo 2 está relacionado ao primeiro objetivo específico, referente à consolidação de um corpo de conhecimento sobre GLA e LLMs. Ele descreve a realização de uma *scoping review* e de buscas exploratórias na literatura. No Capítulo 3, caracterizam-se as dificuldades e estratégias em modelagem de dados e projeção de *prompts* para LLMs (segundo objetivo específico), por meio de dois estudos empíricos e um experimento controlado. Quanto ao Capítulo 4, busca-se uma alternativa para auxiliar na modelagem de dados e na geração de estruturas de GLA, sem a necessidade de *prompts* complexos. Nesse sentido, um agente de conversação no ChatGPT em formato de POC é projetado, desenvolvido e avaliado por *designers* de aprendizagem, contemplando o terceiro objetivo específico. O Capítulo 5 detalha a realização de um *Design* Participativo e descreve a construção da ferramenta GLORIA, atendendo ao quarto objetivo

específico. Por fim, o Capítulo 6, também alinhado ao objetivo específico 4, apresenta a avaliação da ferramenta para investigar:

- Se o artefato está em concordância com seus requisitos: realizado mediante a execução de uma avaliação heurística da ferramenta com especialistas em *design* de interface e experiência do usuário (UX/UI), que já participaram da construção de sistemas;
- Se o problema foi resolvido de modo satisfatório: conduziu-se um estudo empírico com *designers* de aprendizagem (desenvolvedores, educadores e estudantes que atuam nessa função), com o objetivo de verificar se as dificuldades de modelagem de dados foram mitigadas a partir do uso do GLORIA;
- Se as conjecturas teóricas são válidas: também investigado por meio do estudo empírico, analisando se as suposições teóricas que fundamentam o GLORIA (uso de LLMs para apoiar a modelagem de dados) se confirmam empiricamente.

É válido ressaltar que esta dissertação garante a confidencialidade em ética de todas as pesquisas realizadas, sendo aprovada pelo Comitê de Ética em Pesquisa sob o número do CAAE 85800424.8.0000.5020, conforme parecer consubstanciado nº 7.465.500.

1.4 Organização do Trabalho

Nesta dissertação de mestrado, além deste capítulo introdutório, outros capítulos são contemplados, organizados da seguinte forma:

- Capítulo 2 - Referencial Teórico: apresenta a base teórica do trabalho, incluindo jogos educacionais, avaliação de aprendizagem em jogos, *Game Learning Analytics*, GLBoard, modelagem de dados e *Large Language Models*. Além disso, descreve a *scoping review* e as buscas exploratórias na literatura sobre GLA e LLMs, bem como os trabalhos correlatos e as inovações desta pesquisa.
- Capítulo 3 - Investigando dificuldades e o uso de LLMs para GLA: apresenta os resultados obtidos no Ciclo de Relevância da DSR, os quais envolvem a realização

de pesquisas sobre a dificuldade de estudantes de computação, que atuavam como *designers* de aprendizagem, na modelagem de dados para jogos educacionais e a utilização de LLMs para geração de estruturas de captura de dados de GLA.

- Capítulo 4 - Construindo e avaliando um agente especialista em GLA: contempla os resultados obtidos no Ciclo de *Design* da DSR, que envolvem a construção de um agente inteligente em GLA no ChatGPT, em formato de POC, bem como sua avaliação por especialistas na área e estudantes.

- Capítulo 5 - GLORIA: analisando jogos e orientando a inserção de GLA: descreve a realização de *Design* Participativo com o público-alvo para auxiliar no levantamento de requisitos da proposta, detalha o processo de construção da ferramenta e apresenta o artefato GLORIA.

- Capítulo 6 - Avaliando o GLORIA para modelagem de dados em GLA: descreve a avaliação da ferramenta GLORIA, por meio de uma avaliação heurística com especialistas em UX/UI e de um estudo empírico com *designers* de aprendizagem.

- Capítulo 7 - Discussões dos achados: sintetiza e discute os principais achados da pesquisa, relacionados a modelagem de dados como desafio em GLA e a aplicação de LLMs para contextos de aprendizagem. Além disso, discute a operacionalização da ferramenta GLORIA para concretizar os pressupostos teóricos.

- Capítulo 8 - Considerações finais: apresenta as conclusões do trabalho, as principais contribuições, as limitações, as publicações relacionadas com a dissertação e os trabalhos futuros.

2

Referencial Teórico

Neste capítulo, localizam-se os fundamentos teóricos que embasam o trabalho: jogos educacionais, avaliação de aprendizagem em jogos, *Game Learning Analytics*, GLBoard, modelagem de dados e IA Gerativa, incluindo *Large Language Models*. A fundamentação aqui presente foi resultado de consultas na literatura para compor a base de conhecimento da pesquisa, estritamente relacionada aos Ciclos de Relevância e Rigor da DSR, e ao primeiro objetivo específico[1].

2.1 Jogos Educacionais

Os jogos possuem capacidades de engajamento e motivação, proporcionando experiências lúdicas aos jogadores que os estimulam a continuar jogando (Plass; Homer; Kinzer, 2015). Por conta desses fatores, vêm sendo utilizados para propósitos sérios, como treinamento governamental, saúde, política pública e comunicação estratégica (Djaouti et al., 2011; Zyda, 2005; Wang et al., 2022). Denominados de *Serious Games* (Jogos Sérios), possuem como foco não somente a diversão, mas também o desenvolvimento de habilidades relacionadas a conteúdos. Desse modo, as tomadas de decisão estão vinculadas tanto à *gameplay* quanto ao contexto sério (ABT, 1970; Melo; Silva; Pires, 2019; Charsky, 2010). Além disso, estudos sugerem que jogos podem favorecer experiências de *flow* (Csikszentmihalyi; Csikzentmihaly, 1990), nas quais se observam concentração

[1] Consolidar um corpo de conhecimento sobre avaliação de aprendizagem em jogos educacionais, GLA, LLMs e Engenharia de *Prompt*, por meio de uma *scoping review* e de buscas exploratórias na literatura

intensa e alteração da percepção do tempo, especialmente quando há objetivos claros, *feedback* imediato e equilíbrio entre desafios e habilidades (Liu, 2016; Piacentini, 2011).

Diante do potencial dos jogos para promover motivação e engajamento, bem como de sua possível contribuição para experiências de *flow*, seu uso tem sido explorado em abordagens como a *Game-based learning*, na qual são utilizados como objetos de aprendizagem (Plass; Homer; Kinzer, 2015; Qian; Clark, 2016). Nesse aspecto, a educação destaca-se como um dos principais campos onde os *Serious Games* são aplicados e avaliados, denominados de *Educational Serious Games* ou Jogos Educacionais (Hsiao, 2007; Din; Baig; Khan, 2023). Por meio de elementos lúdicos e ambientes imersivos, esses objetos buscam fomentar o desenvolvimento de competências e habilidades relacionadas a conteúdos curriculares, como matemática, língua portuguesa, programação, entre outros (Gasparini; Carvalho; Hounsell, 2015; Macedo, 2024; Santos et al., 2019).

Conforme Savi e Ulbricht (2008), os jogos educacionais trazem diversos benefícios aos processos de aprendizagem, como: efeito motivador, facilitador do aprendizado, desenvolvimento de habilidades cognitivas, aprendizado por descoberta, socialização, entre outros. Nesse aspecto, trazem uma ressignificação de aprendizagem: geralmente é associada a algo monótono, mas quando o indivíduo se conecta a uma atividade que possibilita satisfação, consegue dedicar tempo e esforço para a mesma (Melo; Pires; Melo, 2018; Gros, 2007). Portanto, ao proporcionar motivação e engajamento, podem também facilitar a aprendizagem (Tarouco et al., 2004). No entanto, avaliar jogos educacionais apresenta obstáculos, descritos na seção seguinte.

2.2 Avaliação de Aprendizagem em jogos

Em relação aos jogos educacionais, o processo de avaliação apresenta lacunas na literatura: (i) questionários são a forma de avaliação mais utilizada, geralmente extensos, aplicados em um único momento (pós-jogo) e considerando somente a percepção dos jogadores; (ii) a não utilização de modelos/métodos bem definidos para avaliação, que majoritariamente são relatados de forma informal *ad-hoc*, não incluindo aspectos como objetivo, medidas ou instrumentos para coletar dados (Petri; Wangenheim, 2017); (iii)

ausência de trabalhos que incluem a comparação de resultados com pré e pós-testes (Battistella; Wangenheim, 2016), dificultando a compreensão do impacto do jogo; e (iv) discussões limitadas que consideram o jogo como ferramenta avaliativa, cuja avaliação ocorre somente após o fim da *gameplay* (*post-game*) (Alonso-Fernandez et al., 2017; Silva et al., 2021). Ademais, em jogos educacionais, além dos elementos de entretenimento, é necessário avaliar os aspectos de aprendizagem (Melo et al., 2020).

No entanto, é complexo avaliar a aprendizagem. Esse desafio está correlacionado com a neurociência, visto que os seres humanos aprendem de formas distintas e não há uma fórmula para identificar a aprendizagem. Dentre os tipos de avaliação que buscam mensurá-la, tem-se: (i) diagnóstica - geralmente aplicada antes do início formal do estudo, cujo objetivo é auxiliar no direcionamento dos estudantes conforme suas habilidades; (ii) formativa - realizada de forma contínua, por meio de exercícios, tarefas e testes, visando fornecer *feedback* para estimular o aprendizado dos estudantes; (iii) somativa - tem o propósito de mensurar a progressão do estudante com um julgamento (geralmente irreversível) sobre seu desempenho, como notas, porcentagens ou classificações; e (iv) autoavaliação - o estudante analisa seu progresso a partir da própria percepção (Oliveira et al., 2019; Cox; Imrie; Miller, 2014). Em jogos educacionais, os tipos mais comuns são as avaliações diagnósticas, com pré-testes; somativas, com pós-testes; e autoavaliativas, com questionários sobre percepção de aprendizagem.

É fundamental localizar estratégias para avaliar se os objetivos de aprendizagem do jogo estão sendo atingidos. No entanto, os modos de avaliação citados anteriormente, apesar de importantes, são insuficientes para esse aspecto e não permitem identificar com exatidão o progresso do jogador. Além disso, avaliações somativas são as menos utilizadas em relação aos tipos de avaliação e a aprendizagem contempla vários aspectos de competências e habilidades adquiridas durante o processo educativo (Gee, 2003), que geralmente não são considerados na avaliação dos estudantes com jogos educacionais.

Nesse sentido, os jogos assemelham-se a caixas-pretas: apresentam um estado final com os resultados dos estudantes, como métricas de desempenho (pontuação), mas que não fornecem informações sobre o processo de aprendizagem (Alonso-Fernandez et al., 2017). Por consequência, os educadores não possuem controle ou percepção sobre

o que ocorre enquanto os estudantes estão jogando (Alonso-Fernández et al., 2021). Essa limitação decorre do fato de os traços de aprendizagem estarem relacionados aos elementos de mecânica (Melo et al., 2020; Eck, 2006) e precisarem ser coletados durante o jogo (*in-game*), como tomadas de decisão, tempo para resolução das fases, performance, entre outros. Esses dados permitem analisar o comportamento do jogador e rastrear suas interações durante a *gameplay*, o que não é apenas desejável, mas fundamental (Alonso-Fernández et al., 2021; Alonso-Fernandez et al., 2017). Essa estratégia é explorada na área de *Game Learning Analytics*, descrita a seguir.

2.3 Game Learning Analytics - GLA

GLA, ou *Learning Analytics* para *Serious Games*, é a intersecção de duas áreas: *Game Analytics* (GA) e *Learning Analytics* (LA):

- GA consiste na aplicação de métodos de análise de dados ao desenvolvimento, à avaliação e à pesquisa em jogos, visando transformar dados coletados em informações úteis para apoiar diferentes tomadas de decisão. Esses dados podem incluir telemetria, métricas de comportamento dos jogadores, dados de desempenho, informações de mercado e outras fontes relacionadas ao ciclo de vida do jogo. Dessa forma, GA pode apoiar decisões de *game design*, fornecendo *feedback* sobre o *design*, indícios sobre a experiência do jogador, anomalias de comportamento, problemas técnicos, entre outros aspectos (Drachen; El-Nasr; Canossa, 2013).

- LA trata-se da medição, coleta, análise e relato de dados em ambientes de aprendizagem (como *Learning Management Systems* e *Massive Open Online Courses*) para extrair informações sobre o desempenho, o comportamento e o progresso dos estudantes dos estudantes e, assim, compreender e otimizar a aprendizagem (Alonso-Fernández et al., 2019; Siemens; Long, 2011; Sclater, 2017). Nesse campo utilizam-se práticas baseadas em evidências que derivam da avaliação de progresso, motivação, atitudes e satisfação dos estudantes, promovendo análises de dados com técnicas distintas como estatística, análise de texto e mineração de

dados (Mangaroska; Giannakos, 2018).

Combinando as tecnologias de GA e os objetivos educacionais de LA, tem-se a área de *Game Learning Analytics* (GLA), uma abordagem baseada em evidências para a coleta, análise e visualização de dados oriundos de jogos sérios, que tem se intensificado nos últimos anos (Freire et al., 2016; Banihashem et al., 2024). Essa estratégia compreende: (i) o armazenamento das interações dos jogadores, comumente denominadas de rastros (Alonso-Fernandez et al., 2017), traços ou registros de dados (*logs*), por meio de uma coleta não intrusiva para evitar interromper o estado de *flow*; (ii) a visualização dos dados coletados, comumente através de *visual analytics*, como *dashboards* (Alonso-Fernández et al., 2022; Scheneider; Lemos, 2020); e (iii) a análise dessas interações, que pode trazer contribuições diversas, como: entender como os estudantes aprendem com jogos (em *Educational Serious Games*), validar o *game design*, investigar dificuldades, comparar comportamentos e, principalmente, replicar a trajetória (percurso) do jogador (Alonso-Fernández et al., 2019).

Esse último tópico chama a atenção de pesquisadores, sobretudo em jogos educacionais, visto que os jogos geralmente oferecem diferentes estratégias e caminhos para os jogadores progredirem. Dessa forma, podem fornecer uma avaliação mais detalhada de como pode ocorrer a aprendizagem (Silva et al., 2021; Wallner; Kriglstein, 2013), mapeando elementos-chave como registros de tempo, tomadas de decisão, interação com objetos, dentre outros. Portanto, GLA consiste em uma estratégia baseada em dados que traz diversas contribuições, cujos principais *stakeholders* são educadores, desenvolvedores de jogos e estudantes (Banihashem et al., 2024).

Cada um desses perfis pode se beneficiar de GLA de formas distintas: educadores podem monitorar o progresso dos estudantes em tempo real ou de forma assíncrona, possibilitando localizar lacunas na aprendizagem durante o jogo e realizar intervenções para aprimorá-la; desenvolvedores podem analisar o *game design* para identificar inconsistências e aplicar melhorias, além de adquirir experiências enriquecedoras que contribuem para seu crescimento profissional, visto que a implementação de GLA é um processo multidisciplinar; e estudantes podem analisar seu próprio progresso e estatísticas, permitindo que tenham consciência de como estão se saindo (Alonso-Fernández

et al., 2019; Banihashem et al., 2024; Melo et al., 2020; Honda et al., 2023). Além disso, ao cruzar os dados de GLA com avaliações heurísticas e utilizar técnicas de ciência de dados mais robustas para as análises, pode-se obter *insights* mais valiosos sobre a aprendizagem e o comportamento dos jogadores (Alonso-Fernández et al., 2021; Silva et al., 2021; Alonso-Fernández et al., 2022).

Para a implementação de um sistema de GLA, Freire et al. (2016) preconizam um conjunto de módulos (Figura 2). O processo inicia-se no jogo que, através da *engine*, envia os dados para o "Coletor" (*tracker*). Em seguida, o "Agregador" classifica e organiza os dados para gerar informações, que posteriormente são utilizadas pelo "Relator" para fornecer relatórios e visualizações, geralmente em *dashboards*, que podem ser atualizados em tempo real. No "Avaliador", as informações gráficas são analisadas para extrair *insights* sobre a evolução do jogador e aspectos de sua aprendizagem. Por fim, o "Adaptador", um módulo opcional, retorna as instruções para o jogo para que se adapte ao jogador com base no seu desempenho (Freire et al., 2016).

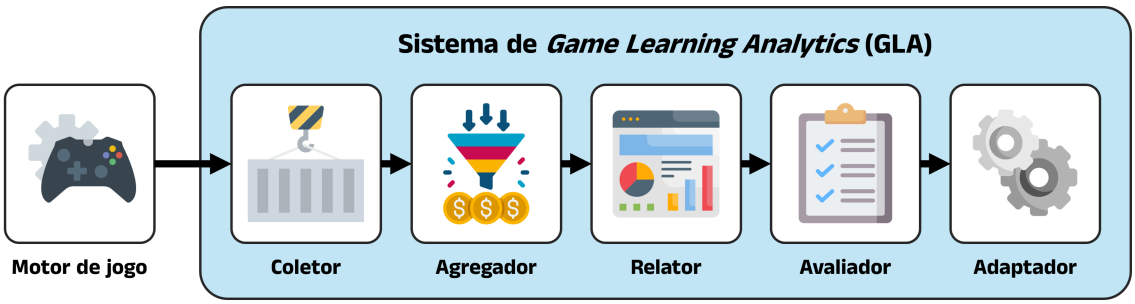


Figura 2 – Arquitetura para um sistema de GLA, adaptado de Freire et al. (2016).

Apesar dos benefícios na utilização de GLA e a arquitetura conceitual de Freire et al. (2016) para auxiliar na implementação de sistemas relacionados, prevalecem desafios no campo: (i) a implementação das técnicas não é simples, tornando-se um impedimento para os desenvolvedores incluirem-nas (Saveski et al., 2016). Esse fator é derivado de LA, cujo processo é complexo e custoso (Alonso-Fernández et al., 2022); e (ii) a ausência de uma padronização, de modo que sejam implementadas coletas de dados a contextos específicos, que não são replicáveis (Alonso-Fernandez et al., 2017). Esse desafio decorre do fato de que os jogos educacionais possuem objetivos específicos de aprendizagem e estruturas próprias, gerando dados distintos.

Apesar desses desafios, a adoção de um modelo sistematizado mostra-se necessária, pois, conforme Melo et al. (2020), traz benefícios: (i) possibilita replicar o componente de coleta de dados, sem reinventá-lo; (ii) inclui um *template* de dados que se flexibiliza a qualquer jogo para atender a novos dados específicos; e (iii) proporciona um ambiente de visualização para que os professores e desenvolvedores façam análises do desempenho dos jogadores a partir dos dados. Dentre os modelos que têm essa proposta, destacam-se o xAPI-SG, atual XASU do ecossistema SIMVA (Santilario-Berthilier et al., 2025), e o GLBoard (Silva et al., 2022).

O xAPI-SG é um perfil da *Experience API* (xAPI) para *Serious Games*, cujo objetivo é padronizar a coleta de interações dos jogadores. Nesse perfil, as interações são representadas por meio de declarações que descrevem atividades de aprendizagem. Essas declarações possuem formato JSON e são compostas por três campos principais: ator (quem fez a ação), verbo (ação em si) e objeto (alvo da ação), podendo incluir atributos adicionais, como *timestamp*, resultado e contexto. O perfil define um conjunto de verbos e atividades, que representam as interações comuns a jogos sérios, como completáveis (tarefas ou fases), alternativas (escolhas do jogador) e variáveis gerais, além de extensões *ad-hoc* que podem ser incluídas para descrever interações específicas. Por exemplo, um jogador (ator) pode ter progredido (verbo) na questão 1 (objeto), obtendo um desempenho de 0.33 (resultado, campo estendido) (Serrano-Laguna et al., 2017; Alonso-Fernandez et al., 2017; Alonso-Fernández et al., 2021).

Posteriormente, o "*eUCM E-Learning Research Group*", da Universidade de Madrid[2], conceberam o SIMVA (*Smart Integration for Monitoring, Visualization and Assessment*), uma plataforma completa para suporte à GLA. O xAPI-SG foi integrado a esse ecossistema e não atua mais diretamente como *tracker*: passou a ser implementado por meio do XASU (*xAPI Serious Games Analytics Support Unit)*, o responsável pela coleta, transformação e envio dos dados ao banco principal do sistema. Desse modo, o xAPI-SG continua atuando como padrão semântico das interações, mas o *tracker* é o XASU, oferecendo recursos adicionais e uma implementação automatizada dentro do SIMVA (Santilario-Berthilier et al., 2025).

Embora o SIMVA represente um avanço para a área de GLA, antes de sua

[2] Um grupo de pesquisa predominante na Espanha sobre GLA. São os criadores do xAPI-SG.

concepção, o xAPI-SG apresentava limitações à sua adoção como modelo de GLA, tais como: (i) a alta curva de aprendizagem, que exige um entendimento aprofundado da estrutura do xAPI e do vocabulário para a coleta dos dados; (ii) a necessidade de adaptação decorrente da flexibilidade do modelo, de forma que, apesar de o perfil focar interações em *Serious Games*, o xAPI foi inicialmente elaborado para ambientes educacionais genéricos; (iii) o conhecimento técnico exigido em *back-end* e integração com *game engines*, visto que era necessária a configuração de um *Learning Record Store* (LRS); e (iv) a ausência de uma ferramenta nativa para analisar e visualizar os dados coletados, como um *dashboard*, exigindo integração com uma ferramenta externa, como o T-MON (*Teacher-Monitoring Tool*) (Perez-Colado et al., 2018), desenvolvido pelos próprios criadores do xAPI-SG. Ademais, embora o SIMVA mitigue as limitações técnicas, a curva de aprendizagem e a necessidade de adaptação do xAPI-SG ainda persistem.

2.4 GLBoard

O GLBoard (Silva et al., 2022) foi concebido para contornar os desafios de GLA e as limitações do xAPI-SG. Ele consiste em um modelo para captura e análise de dados de jogos educacionais, cuja premissa é a de que existem dados genéricos nesses objetos, como registros de tempo, mudanças de nível, tentativas e interações do jogador (Cano; Fernández-Manjón; García-Tejedor, 2017). O modelo é composto por quatro módulos: (i) Pacote *Unity*: uma biblioteca da *game engine* que inclui um *template* de captura de dados, auxiliando na padronização dos dados genéricos a serem extraídos; (ii) API: responsável pelo controle do sistema e comunicação com os demais módulos, além do gerenciamento de dados e condução de análises sobre eles; (iii) Banco de Dados: armazena todos os rastros (*logs* de interação) dos jogadores; e (iv) *Dashboard*: exibe as análises feitas pela API em forma de relatórios e visualizações em gráficos.

A estrutura padrão do GLBoard para captura de dados é um *template* genérico representado em JSON, composto por quatro classes principais (Figura 3): (i) *PlayerData*, com dados sobre o jogador (nome, gênero, data de aniversário, id); (ii) *GameData*, com dados provenientes da *gameplay* durante as fases; (iii) *Phase*, os níveis do jogo que

possuem um nome, um status (finalizada ou não) e um conjunto de seções; (iv) *Session*, as sessões do jogo, que podem ser consideradas como "tentativas". Cada sessão inclui: data de início e finalização, conclusão (vitória, derrota ou desistência), performance (pontuação, estrelas coletadas, entre outros), uma imagem representativa da rota percorrida ($route\_image\_b64$), o percurso do jogador ($path\_player$) e os resultados finais ($finalized\_challenges$), com indicadores derivados e agregados do percurso.

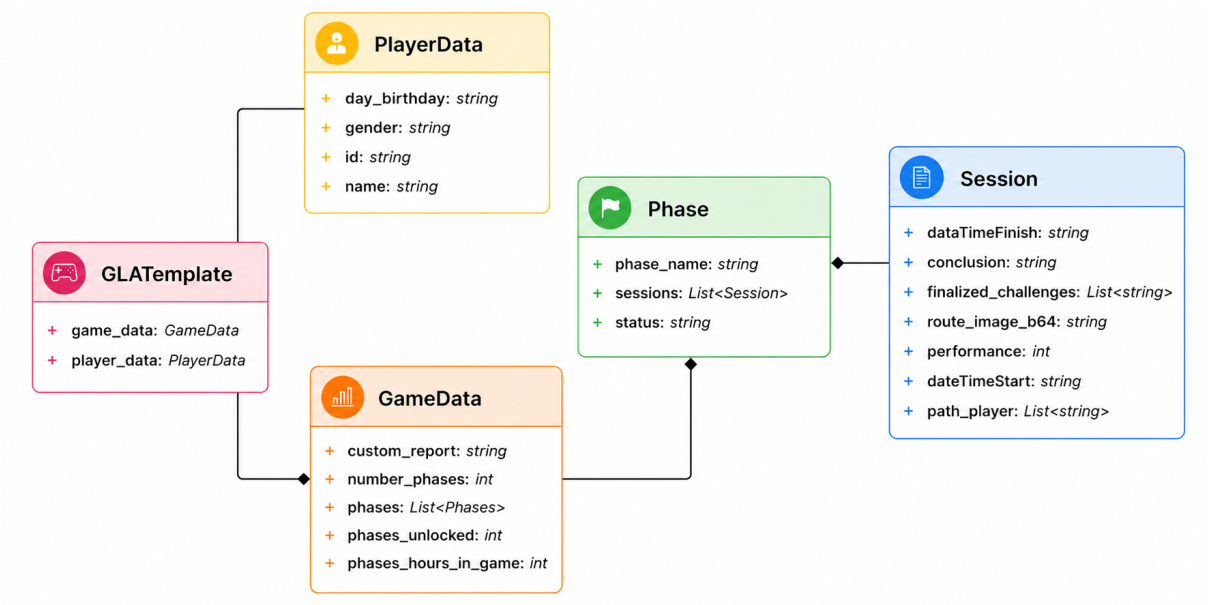


Figura 3 – *Template* de dados do GLBoard.

A variável $path\_player$ armazena o percurso do jogador durante as fases, fundamental para identificar as evidências de aprendizagem, contemplando: interações com objetos, registros de tempo, tomadas de decisão, entre outros. Por ser um objeto do tipo $List<string>$, possibilita que o desenvolvedor inclua outras variáveis de GLA específicas (posteriormente convertidas para $string$), tornando o GLBoard flexível para qualquer jogo educacional. Por exemplo, no jogo "Cadê minha pizza?" (Honda et al., 2023) sobre caminho mínimo (Teoria dos Grafos), descrito no Apêndice A, variáveis do $path\_player$ incluem: seleção do entregador (objeto que irá percorrer o grafo), residência escolhida (vértice pretendido para visita), *feedback* da ação (aresta percorrida ou não) e o tempo para cada tomada de decisão (Honda et al., 2023).

A partir dos dados registrados no $path\_player$, também podem ser calculados indicadores agregados e informações relacionadas aos desafios alcançados, armazenados

na variável $finalized\_challenges$. Esses indicadores sintetizam resultados obtidos ao longo da fase, como a quantidade de decisões corretas, o número de tentativas, o tempo total empregado ou a conclusão de objetivos específicos. Para compor o $path\_player$ e o $finalized\_challenges$, é necessário definir as variáveis de GLA que representarão o percurso do jogador e os resultados alcançados, e inseri-las em um *template* de dados. Essa etapa, denominada modelagem de dados e descrita na Seção 2.5, é fundamental para qualquer ferramenta de GLA.

2.5 Modelagem de dados

Antes da implementação de GLA, deve-se realizar a modelagem de dados. Essa etapa consiste em definir quais dados de interação (rastros) serão coletados, articulando *game design* e *learning design* (Shoukry; Göbel; Steinmetz, 2014), com o objetivo de identificar evidências de aprendizagem. Esses dados geralmente são definidos pelo *designer* de aprendizagem e devem ser suficientes para expressar se os objetivos de aprendizagem estão sendo atingidos (Perez-Colado et al., 2018). Além disso, essa definição deve ocorrer desde a concepção do jogo, para que os dados coletados possibilitem análises relevantes e interpretáveis sobre o desempenho, o progresso e a aprendizagem do jogador (Hauge et al., 2014; Alonso-Fernández et al., 2021; Kitto et al., 2020).

A modelagem de dados é fundamental para a avaliação da aprendizagem no jogo, visto que dados de interação "superficiais" (cliques, acertos e erros) não fornecem *insights* suficientes para identificar as evidências de aprendizagem ou auxiliar na tomada de decisões de *game design*. Esperar que esses dados, por si sós, forneçam tais informações constitui uma expectativa irrealista, que não corresponde à GLA, mas à chamada "informágica" (Perez-Colado et al., 2018). A partir da modelagem, as variáveis de GLA são criadas e organizadas em um *template* de dados específico (como os do xAPI-SG e do GLBoard), e posteriormente implementadas no projeto do jogo. Recomenda-se que essa etapa e o desenvolvimento do jogo sejam realizados em paralelo, visto que jogos parcial ou completamente desenvolvidos apresentam diversos desafios para a implementação da captura de dados, mesmo para profissionais com experiência em

jogos educacionais e GLA (Macena et al., 2024).

Na literatura, não há uma nomenclatura unificada para a modelagem de dados. Os estudos utilizam termos distintos para designar atividades que a compõem total ou parcialmente, como *Choosing Data* (Shoukry; Göbel; Steinmetz, 2014), *competence modelling* (Petersen; Bedek, 2012), *evidence identification* (Kerr; Chung, 2012) e *design of game telemetry* (Chung, 2015). Hauge et al. (2014), por exemplo, utilizam a expressão *Data selection/capturing*. Entretanto, considera-se que a seleção e a captura dos dados correspondem apenas a parte desse processo. Nesse sentido, adota-se, nesta dissertação, o termo modelagem de dados por abranger não apenas a definição dos dados a serem coletados, mas também a especificação das variáveis e de seus tipos, o estabelecimento de relações com as ações do jogo e os objetivos de aprendizagem, a justificativa de sua relevância e sua organização em um *template* de dados.

Ainda que aspectos da modelagem de dados sejam discutidos em estudos distintos de GLA, observou-se uma quantidade limitada de pesquisas que a tratem como foco central ou apresentem orientações operacionais detalhadas para realizá-la[3]. Essa limitação pode estar relacionada à diversidade de nomenclaturas utilizadas para se referir à modelagem, dificultando a identificação de trabalhos relacionados. Também é possível que esteja associada ao fato de, em algumas abordagens, essa etapa envolver educadores, *game designers* e desenvolvedores, o que pode levar ao pressuposto de que esses atores já disponham dos conhecimentos necessários para realizá-la (Cano; Fernández-Manjón; García-Tejedor, 2017; Perez-Colado et al., 2018).

Por exemplo, nos trabalhos do grupo e-UCM, a modelagem é realizada em conjunto por educadores e *game designers*. Nesse caso, ela não recebe uma nomenclatura específica, mas é contemplada como parte de um *Learning Analytics Model* (LAM). O LAM é um modelo conceitual que envolve a definição de: (i) objetivos de aprendizagem; (ii) objetivos do jogo; (iii) dados a serem coletados, componente diretamente relacionado à modelagem de dados; e (iv) como os dados serão interpretados e analisados (Perez-Colado et al., 2018). Portanto, com base no *learning design* e no *game design*, o LAM orienta a definição e a análise dos dados (Alonso-Fernández et al., 2021). Quando não há um LAM, os *game designers* podem sugerir quais informações obter do jogo. Os

[3] Observada durante as pesquisas exploratórias desta dissertação e na *Scoping Review* do Capítulo 2.

trabalhos com maior detalhamento sobre a modelagem de dados envolvendo o LAM e o modelo xAPI-SG são os de Perez-Colado et al. (2018) e Alonso-Fernández et al. (2021), em que se especificam os dados coletados, suas relações com os objetivos do jogo e de aprendizagem, os tipos de eventos utilizados no *template* do xAPI e a forma como são interpretados para apoiar a avaliação dos jogadores.

No entanto, essa estratégia apresenta limitações: (i) os componentes e as etapas do LAM são definidos, mas não é apresentado um guia operacional detalhado para orientar sua construção; (ii) embora seja discutida a articulação entre *learning design* e *game design* para modelar os dados, não é disponibilizado um documento-base para apoiar essa etapa; (iii) há uma quantidade limitada de trabalhos que apresentam exemplos concretos e sistematizados do LAM. Perez-Colado et al. (2018) apresentam um dos exemplos mais detalhados nesse sentido, mas restrito a parte de seus componentes; e (iv) o LAM está fortemente vinculado ao padrão xAPI, o que pode dificultar sua adaptação a outros modelos de GLA.

Somado a essas limitações, observa-se que a modelagem de dados constitui uma etapa desafiadora em razão: (i) do trabalho conceitual necessário para decompor competências e traduzi-las em indicadores comportamentais observáveis, considerando fatores contextuais (Petersen; Bedek, 2012); (ii) da grande quantidade de variáveis observáveis e das diferentes relações que cada ação pode manter com o desempenho, dificultando identificar quais delas representam evidencias de aprendizagem. Esse desafio geralmente está associado a abordagens nas quais as interações são coletadas de forma ampla e a seleção dos observáveis que podem fornecer evidências de aprendizagem ocorre posteriormente (Kerr; Chung, 2012; Cano; Fernández-Manjón; García-Tejedor, 2017); e (iii) da necessidade de definir dados necessários às análises pretendidas e suficientes para informar o grau de alcance dos objetivos do jogo, articulando-os ao *design* de aprendizagem, pois uma definição mais específica e detalhada pode aumentar a complexidade de sua elaboração (Perez-Colado et al., 2018).

Além desses desafios, é comum que estudantes atuem como *designers* de aprendizagem em disciplinas, oficinas ou contextos acadêmicos voltados à construção de jogos educacionais (Earp, 2015; Weitze, 2014; Mercer et al., 2017). Essa é uma realidade

no cenário brasileiro, onde cerca de 70% dos jogos educacionais são projetados em universidades (Oliveira et al., 2019). Nesses casos, os estudantes podem assumir a responsabilidade pela modelagem de dados para a implementação das técnicas de GLA (Honda et al., 2025). Entretanto, por ainda estarem em processo de formação e, frequentemente, não contarem com o apoio de equipes multidisciplinares, esses estudantes podem enfrentar dificuldades nessa etapa, conforme discutido no Capítulo 3. Desse modo, embora a modelagem seja fundamental no fluxo de GLA, sua realização apresenta obstáculos diversos, especialmente para estudantes que desempenham o papel de *designers* de aprendizagem.

No fluxo de GLA, etapas subsequentes à modelagem de dados, como a coleta e a análise, podem ser automatizadas, inclusive por meio de técnicas de aprendizado de máquina. A definição dos dados e sua interpretação como evidências de aprendizagem, porém, permanecem dependentes do julgamento humano (Kerr; Chung, 2012; Cano; Fernández-Manjón; García-Tejedor, 2017; Perez-Colado et al., 2018). Nesse sentido, nesta dissertação, denomina-se modelagem manual o processo em que o próprio *designer* de aprendizagem define as variáveis de GLA a partir da análise do jogo.

Considerando as dificuldades da modelagem de dados, sobretudo para estudantes que atuam como *designers* de aprendizagem, o uso de IA Generativa por meio de LLMs pode auxiliar a mitigá-las. Esses modelos, discutidos na Seção 2.6, têm apresentado contribuições recentes para as áreas de *Learning Analytics* (Misiejuk et al., 2025) e *Game Analytics* (Yang; Kleinman; Harteveld, 2025).

2.6 IA Gerativa e *Large Language Models*

A Inteligência Artificial Gerativa, do inglês *Generative Artificial Intelligence* (GAI), vem se expandindo. Essa área utiliza modelos generativos e técnicas de aprendizagem profunda para gerar conteúdos, como textos, imagens, áudios e vídeos (Jovanovic; Campbell, 2022; Sengar et al., 2024). Dentre as técnicas de GAI, destacam-se as *Generative Adversarial Networks* (GAN), os *Variational Autoencoders* (VAE) e as que são baseadas em *Transformers* (Aydın; Karaarslan, 2023; Jovanovic; Campbell, 2022). Essas últimas técnicas referem-

se a modelos cujo núcleo são os *Transformers*, uma arquitetura de rede baseada em mecanismos de atenção, cuja aplicação tem sido destinada à vários domínios, como o de Processamento de Linguagem Natural (PLN) (Lin et al., 2022; Vaswani, 2017).

PLN é um subcampo da Ciência da Computação e Inteligência Artificial, cujo objetivo é construir um sistema de computador capaz de compreender e responder à fala humana ou texto escrito. Dentre suas atividades, destacam-se conversão de fala em texto, desambiguação do sentido das palavras, análise de sentimentos, geração de linguagem natural, entre outros (Aydın; Karaarslan, 2023; IBM, 2023b). Portanto, trata-se da construção de modelos de linguagem que realizam tarefas como a predição da palavra seguinte em uma sequência (*tokens*). Tais modelos evoluíram da modelagem de linguagem estatística para a neural, depois para modelos de linguagem pré-treinados (PLMs) e, recentemente, para os LLMs (*Large Language Models*), aumentando de modo significativo os parâmetros do modelo e o conjunto de dados de treinamento (Naveed et al., 2023; Raffel et al., 2020; Xue et al., 2020).

Os LLMs são modelos de linguagem neural baseados em *Transformers*, que são pré-treinados em *corpus* extensos e possuem bilhões de parâmetros (Minaee et al., 2024). Esses modelos são capazes de gerar textos semelhantes aos produzidos por seres humanos, responder a questões e realizar diversas tarefas de linguagem natural (Kasneci et al., 2023; Ouyang et al., 2022). O uso de LLMs tem atraído atenção desde o surgimento do GPT-3 em 2020 (Brown et al., 2020) e intensificou-se após o lançamento do ChatGPT em 2022, que possui mais de 900 milhões de usuários ativos semanais (dados de fevereiro de 2026) (DemandSage, 2026). Em relação à categorização, pode-se organizar os LLMs em quatro principais categorias, conforme Minaee et al. (2024):

- **Originalidade:** podem ser originais, isto é, que não foram baseados ou ajustados em outro modelo pré-treinado, como o LLaMA; ou "tunados", que se basearam em um modelo original e foram ajustados com conjuntos de dados distintos ou outras abordagens de treinamento. Esse processo denomina-se *fine-tuning* (ajuste fino) e possibilita adaptar os modelos a tarefas específicas e melhorar sua performance, através da modificação dos pesos. Um exemplo é o Alpaca, que é baseado no LLaMA. Nesta dissertação, busca-se realizar o treinamento em modelos específicos

para adaptá-los ao contexto de GLA;

- **Parâmetros:** em relação à quantidade de parâmetros, podendo ser pequenos, com menos de 1 bilhão (DistilBERT, 66M); médios, possuindo entre 1 e 10 bilhões (LLaMA 1, 7B); extensos, possuindo entre 10 e 100 bilhões (Claude, 93B); e muito extensos, com mais de 100 bilhões (GPT-3, 175B). No âmbito desta pesquisa, considerando que a quantidade de parâmetros influencia os recursos computacionais necessários, o foco está na utilização de modelos médios ou extensos.

- **Tipo:** podendo ser um modelo de fundação, ao qual foi pré-treinado em larga escala antes da realização de *fine-tuning*; um modelo de instrução, ao qual se realizou *fine-tuning* supervisionado, cujo LLM foi treinado com conjuntos de dados em pares (instrução, saída) (Zhang et al., 2023); ou um modelo de *chat*, que consiste em um modelo de instrução que posteriormente foi ajustado conforme preferências humanas. O modelo com maior conformidade com esta dissertação é o de *chat*, pois a ferramenta proposta busca orientar o processo de modelagem de dados, envolvendo interações contínuas com o usuário. Portanto, faz-se necessário alinhá-lo às preferências humanas, de modo a tornar os diálogos e a interação adequados ao contexto de GLA.

- **Disponibilidade:** relacionada a se os pesos do modelo são públicos, isto é, estão disponíveis para acesso, como o LLaMA (Meta); ou privados, cujos pesos estão indisponíveis publicamente, como o GPT-4 (OpenAI) e Gemini (Google). Investigações conduzidas nesta dissertação (ver Capítulos 3 e 4) revelaram limitações dos modelos fechados, optando-se pelo uso dos modelos públicos, também denominados de *open-source*.

Além disso, os LLMs podem apresentar capacidades multimodais, analisando, raciocinando e respondendo com informações em múltiplas modalidades, como textos, imagens, áudios, vídeos (Yin et al., 2023). A personalização é fundamental nos LLMs, pois possibilita a melhoria do desempenho em domínios específicos, aumentando sua eficiência e precisão (Nexocode, 2023). Entre as técnicas de personalização mais comuns, destacam-se o *fine-tuning*, que consiste no treinamento do modelo para um contexto

específico, com o ajuste de aspectos como camadas, número de épocas e taxa de aprendizado (Parthasarathy et al., 2024); e a engenharia de *prompt*, tratando-se construção de instruções (*prompts*) efetivas para que o modelo gere resultados adequados, seguindo princípios como claridade e precisão, informação contextual, formato desejado e controle de verbosidade (Lo, 2023). Existem diferentes técnicas de *prompt* na literatura, mas, dado o caráter emergente da área, não é possível determinar com precisão qual delas é a mais apropriada, pois sua adequação depende do contexto de aplicação.

Uma das características da engenharia de *prompt* é denominada de *In-context-learning* (ICL), que se refere ao aprendizado do modelo a partir do contexto fornecido na entrada. Através de ICL, é possível que o modelo apresente um bom desempenho somente com instruções em linguagem natural e/ou alguns exemplos (Li et al., 2023; Dong et al., 2022). Tal característica pode ser classificada em: (i) *zero-shot learning*, onde somente a instrução, sem exemplos, é informada; (ii) *one-shot learning*, que inclui um exemplo na entrada; e (iii) *few-shot learning*, cujo *prompt* inclui dois ou mais exemplos. Estudos apontam que, em diversas tarefas, o *few-shot learning* alcança resultados superiores nas métricas de avaliação adotadas em comparação às demais configurações (Brown et al., 2020; Mann et al., 2020), sugerindo que a inclusão de exemplos pode favorecer os resultados do LLM.

Além da engenharia de *prompt* e do *fine-tuning*, existem inúmeras outras estratégias para a personalização de modelos. Cada abordagem possui suas vantagens e desvantagens, cabendo ao pesquisador determinar qual é a mais adequada a partir de seu objetivo e contexto específico ao qual está investigando. O uso de algumas dessas estratégias é contemplado nos Capítulos 3 e 4, como uma técnica interativa de questionamentos e a criação de uma agente inteligente.

O rápido avanço dos LLMs, incluindo estratégias de treinamento e técnicas de *prompt*, tem impulsionado pesquisas em diferentes áreas do conhecimento, como a de *Learning Analytics*. Dentre as principais contribuições que esses modelos trazem para LA, destacam-se: (i) a análise de dados complexos e não estruturados; (ii) a integração e síntese de dados multimodais; e (iii) o aprimoramento da interpretação dos dados, facilitando a geração de *feedbacks* e relatórios (Clavié; Gal, 2019; Mazzullo et al., 2023;

Misiejuk et al., 2025; Rodríguez-Ortiz; Santana-Mancilla; Anido-Rifón, 2025).

Além disso, o uso de LLMs também impactou a área de *Game Analytics*, tanto na análise de dados de interação dos jogadores quanto no apoio à adaptação de conteúdos e *feedbacks* no jogo (Peng et al., 2024; Yang; Kleinman; Harteveld, 2025). Nesse sentido, os LLMs podem ser poderosos aliados ao campo de GLA (uma intersecção de LA e GA), auxiliando a reduzir a complexidade de modelagem de dados de jogos educacionais. Esse é o foco desta dissertação, cujos trabalhos correlatos são descritos a seguir.

2.7 Trabalhos Relacionados

Esta pesquisa contempla aspectos de duas áreas: *Game Learning Analytics*, uma junção de *Learning Analytics* com *Game Analytics*, e *Large Language Models*. Em pesquisas preliminares, notou-se a ausência de trabalhos que exploram concomitantemente ambos os tópicos. Diante disso, para localizar trabalhos correlatos, optou-se por uma busca na literatura por meio de uma *scoping review*. As etapas metodológicas seguiram o protocolo de Arksey e O'Malley (2005), ilustradas na Figura 4 e descritas a seguir.



Figura 4 – Etapas para a busca por trabalhos relacionados.

Identificação da questão de pesquisa: essa revisão foi conduzida com o objetivo de mapear e entender como os emergentes LLMs estão sendo utilizados no campo de GLA. Nesse intuito, definiu-se como questão de pesquisa: "Como pesquisadores estão aplicando LLMs para contribuir com o campo de GLA?".

Identificação dos estudos relevantes: compõe de três subetapas: (i) seleção das bases de dados, em que foram escolhidas Scopus, ACM Digital Library, IEEE Xplore Digital Library, SpringerLink e ERIC. A seleção baseou-se na relevância que possuem em âmbito acadêmico, com destaque à Scopus, que é a coleção que lidera o setor de resumos e citações acadêmicas, disponibilizando diversas métricas de pesquisa (Elsevier, 2024); (ii) elaboração da *string* de busca, na qual se optou pela utilização de duas *strings*: uma mais abrangente (S1) e outra mais restrita (S2), a fim de analisar a relevância do tema. As *strings* elaboradas foram: S1 - *("Game Learning Analytics" OR "GLA" OR "Educational Game Analytics" OR "Learning Analytics for Games") AND ("Large Language Models" OR "LLM" OR "GPT" OR "Transformer-based Models" OR "Generative AI" OR "GenAI" OR "IA Gerativa" OR "IA Generativa")* e S2 - *"Game Learning Analytics" AND "Large Language Model"*; e (iii) aplicação das *strings* nas bases de dados, que retornaram 26 trabalhos: 12 da SpringerLink, 10 da IEEE Xplore e 4 da Scopus, enquanto ACM Digital Library e ERIC não retornaram nenhum trabalho.

Seleção dos estudos: essa etapa se iniciou com a definição dos critérios de inclusão e exclusão[4], que são ilustrados na Tabela 2. Em seguida, com os trabalhos retornados após a aplicação das *strings* de busca, realizou-se a leitura e análise dos *abstracts*, ao qual se utilizou o ChatGPT (versão 4o) como ferramenta auxiliar para verificar se o trabalho atendia à pesquisa. Cada *abstract* foi enviado individualmente ao modelo, que o analisava e retornava uma resposta com a sugestão de inclusão ou exclusão e os principais argumentos. Em seguida, a resposta era avaliada para determinar se o trabalho seria ou não incluído. Apesar de o uso do ChatGPT poder representar uma ameaça à validade, seu uso destinou-se apenas a servir de subsídio para a análise, visando otimizar o tempo de leitura. Após a análise, constatou-se que nenhum trabalho atendia à pesquisa: todos focavam-se em apenas uma das áreas (GLA ou LLM). Embora não tenha sido incluído, localizou-se um trabalho sobre GLA que citava IA Gerativa, mas apenas como sugestão, não sendo seu foco. Considerando que nenhum trabalho atendeu aos critérios de inclusão, as etapas posteriores de "organização e categorização dos dados" e "síntese e relato dos resultados" não foram executadas.

[4] Ao contrário de revisões sistemáticas, o método de Arksey e O'Malley (2005) para *scoping reviews* permite definir ou ajustar critérios de inclusão e exclusão *post-hoc*, conforme a familiaridade do pesquisador com a leitura dos estudos na triagem inicial.

Tabela 2 – Critérios de inclusão e exclusão

Tipo	**Id**	**Descrição**
Inclusão	[Inc1]	O artigo deve abordar sobre o uso de LLMs para contextos que envolvam GLA
Exclusão	[Exc1]	Não atender nenhum dos critérios de inclusão
	[Exc2]	A versão completa do artigo não estar disponível entre as fontes selecionadas
	[Exc3]	O idioma do artigo não estar em inglês
	[Exc4]	A publicação do artigo não estar em conferências ou periódicos
	[Exc5]	O artigo não estar completo (considerando para isso no mínimo 6 páginas)
	[Exc6]	O artigo estar duplicado
	[Exc7]	O artigo ser livro, tese ou dissertação
	[Exc8]	O estudo não ser um estudo primário, MSLs, revisões sistemáticas da literatura ou mesmo estudos apresentando revisões informais da literatura; ou ser sumário de anais de eventos científicos, descrição de curso, editorial, resumo de palestra, workshop ou tutorial, entre outros
	[Exc9]	O estudo ser uma versão mais antiga de outro estudo já considerado

A pesquisa realizada evidencia que, apesar da emergência da área de IA Gerativa com o uso de LLMs e do crescente interesse sobre o uso de técnicas de GLA (Banihashem et al., 2024), esse ponto de convergência ainda é limitado. Nesse sentido, esta pesquisa localiza-se na intersecção entre essas áreas, como ilustra a Figura 5. Visando não se limitar somente à *scoping review*, também foram realizadas buscas exploratórias na literatura para localizar trabalhos semelhantes: (i) de modo manual, em bases de dados como *Scopus* e Google Acadêmico, utilizando sinônimos de GLA; e (ii) por meio da função de "Investigar" (*Deep Research*) do ChatGPT, que realiza pesquisas avançadas e autônomas na *web*, analisando múltiplas fontes (textos, imagens, PDFs) e sintetizando os resultados em forma de relatórios, geralmente incluindo citações (OpenAI, 2025). Essa estratégia exploratória permitiu localizar trabalhos similares, descritos a seguir.

Wang et al. (2025) descrevem um jogo de realidade alternada (LLM-ARG), mediado por um agente virtual baseado em LLM (ERNIE-Bot-4). O estudo buscou investigar o impacto do *feedback* adaptativo e metacognitivo na aprendizagem dos estudantes, integrando o modelo via plataforma Qianfan. Um desenho quase-experimental foi adotado, com 57 estudantes de 12-13 anos de uma escola primária da China. Os participantes

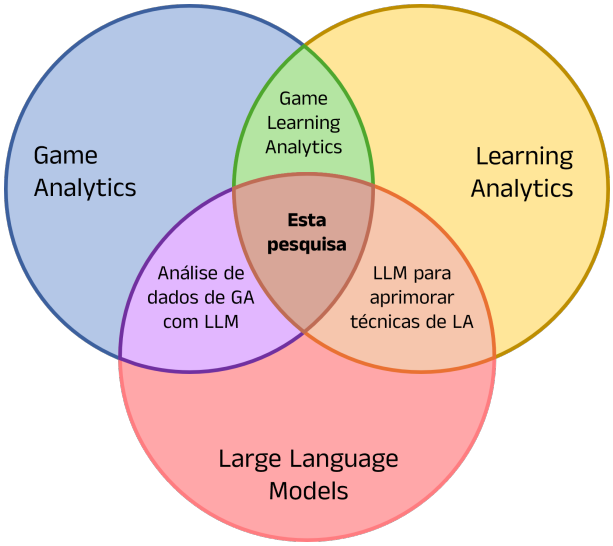


Figura 5 – Diagrama de Venn sobre a localização do trabalho.

foram organizados em um grupo experimental (28), com o uso do agente e um grupo de controle (27), que interagiu com um ARG baseado em regras (C-ARG). O jogo realizou a coleta automática de *logs* dos jogadores, categorizados em códigos comportamentais (E, R, S, entre outros), cujos dados foram analisados por meio de *lag sequential analysis* e testes estatísticos, como o ANCOVA e correlação de Pearson. Além disso, também foram aplicados testes de desempenho e um questionário de consciência metacognitiva, em escala Likert. Os resultados indicam um efeito pequeno, mas significativo, favorecendo o grupo experimental e sugerindo o potencial dos LLMs para apoiar processos metacognitivos e de autorregulação da aprendizagem em jogos educacionais.

Em Liu et al. (2025), os autores apresentam uma abordagem híbrida que integra LLMs e *machine learning*. O estudo visou detectar comportamentos de dificuldade (*struggle*) em estudantes, com o uso do jogo educacional *Wake: Tales from the Aqualab*. O sistema é integrado ao *Open Game Data*, responsável pela captura automática de *logs*, registrando ações, eventos e progresso dos jogadores. Esses dados são convertidos em *text replays* (narrativas legíveis derivadas dos *logs*) e analisados pelo GPT-4o para identificar indícios de *struggle*. O estudo utilizou um conjunto de dados de 42.889 sessões, provenientes de 19.186 jogadores, extraídas entre janeiro e maio de 2024, organizadas em três rodadas de teste com amostras progressivas (200, 100 e 100 casos) para validar e refinar o método. O desempenho do GPT-4o foi comparado ao do XGBoost, um modelo treinado com 58 variáveis numéricas derivadas das interações do jogo. Os modelos demonstraram

desempenho semelhantes, sendo que o modelo híbrido (GPT-4o + XGBoost) apresentou melhores resultados, superando ambas as abordagens isoladas. O estudo contribuiu para o avanço da análise automatizada de comportamentos de aprendizagem em jogos educacionais, utilizando técnicas de *machine learning* e LLM.

No trabalho de Goslen et al. (2025), é proposto um *framework* integrado a um LLM, com o objetivo de gerar planos (sequências de estratégias para resolver desafios) para o jogo educacional *Crystal Island*. Essa abordagem busca oferecer *scaffolding* adaptativo e apoiar a aprendizagem autorregulada. Dados reais de *gameplay* foram utilizados no estudo, referentes a 144 estudantes de ensino fundamental, totalizando 11.610 sequencias de eventos e 510 planos manuais elaborados pelos estudantes. Esses *logs* foram enviados a dois modelos, junto com metas e objetivos do jogo, que geraram planos automáticos de baixo nível. Os modelos empregados foram o T5-small, ajustado por *fine-tuning* e o GPT-3.5, treinado via *few-shot learning*. Os planos produzidos por LLM foram comparados aos planos elaborados de forma manual pelos estudantes, com base em métricas de correspondência e análise qualitativa. Como resultado, o T5-small gerou planos mais semelhantes aos dos estudantes, enquanto o GPT-3.5 produziu planos mais variados e genéricos. Ambos os resultados evidenciam a contribuição de LLMs para auxiliar no planejamento e *scaffolding* adaptativo em jogos educacionais.

Embora nenhum dos trabalhos mencione GLA de forma explícita, empregam técnicas relacionadas à área, como a coleta, processamento e análise de dados de interação dos jogadores. A análise dos trabalhos correlatos evidencia que: (i) as abordagens são limitadas a contextos específicos, não apresentando evidências de replicação em outros jogos; (ii) nenhum contempla a etapa de modelagem de dados. Nesse sentido, os LLMs são aplicados em outra etapa de GLA; (iii) não há evidências que consideram o estudante como um *stakeholder* (atuando como *designer* de aprendizagem), cujo foco está nos educadores e *game designers*; (iv) as abordagens não permitem a personalização das variáveis que serão coletadas; e (v) não há menções que as abordagens estejam disponíveis para comunidade externa de forma *on-line*. Em Liu et al. (2024), apesar da infraestrutura ser um repositório público, o modelo híbrido não foi disponibilizado.

Esses aspectos são os diferenciais desta dissertação, que propõe a ferramenta

GLORIA, que: (i) permite o cadastro de jogos distintos; (ii) é focada majoritariamente na modelagem de dados para jogos educacionais, empregando um LLM que orienta esta etapa; (iii) considera o estudante como *designer* de aprendizagem, focando em interações que o permitam refletir sobre como o *game design* está vinculado aos elementos de aprendizagem; (iv) possibilita a customização das variáveis de GLA; e (v) será disponibilizada abertamente e de forma gratuita para a comunidade, após integração com o modelo GLBoard. A ferramenta é descrita no Capítulo 5.

2.8 Síntese do Capítulo

Este capítulo apresentou os fundamentos teóricos que sustentam a pesquisa: jogos educacionais, avaliação de aprendizagem em jogos, *Game Learning Analytics*, GLBoard, modelagem de dados e IA Gerativa com *Large Language Models*. Observou-se que o uso de LLMs trouxe contribuições para as áreas de *Learning Analytics* e *Game Analytics*, podendo auxiliar em GLA. Nesse aspecto, conduziu-se uma *scoping review* para localizar trabalhos correlatos que envolvam o uso de LLMs para GLA, realizando as seguintes etapas: (i) identificação da questão de pesquisa, sendo "Como pesquisadores estão aplicando LLMs para contribuir com o campo de GLA?"; (ii) identificação dos estudos relevantes, envolvendo a seleção das bases de dados (Scopus, ACM, IEEE Explore, SpringerLink e ERIC), a elaboração da *string* de busca (optou-se por duas) e a aplicação das *strings* nas bases, resultando em 26 trabalhos; e (iii) seleção dos estudos, onde os critérios de inclusão e exclusão foram definidos, bem como a leitura e análise dos *abstracts* com auxílio do ChatGPT. Nenhum dos trabalhos atendeu aos critérios, sugerindo que o ponto de convergência entre LLMs e GLA ainda é escasso.

Conforme o objetivo de prosseguir com a revisão da literatura, conduziram-se buscas exploratórias de forma manual e com o auxílio da função *Deep Research* do ChatGPT. Essa estratégia permitiu a identificação de três trabalhos correlatos, que não mencionam GLA explicitamente, mas que empregam técnicas da área. Após a análise dos trabalhos, constatou-se que: (i) são aplicados em contextos específicos; (ii) não contemplam a modelagem de dados. Assim, o LLM não é utilizado para propor variá-

veis de coleta; (iii) não consideram o estudante como *stakeholder*, que pode atuar como *designer* de aprendizagem; (iv) não envolvem a personalização de variáveis de GLA; e (v) não estão disponíveis para acesso público. Essas limitações são mitigadas nesta pesquisa, que propõe a ferramenta GLORIA para auxiliar *designers* de aprendizagem a modelar dados e implementar técnicas de GLA.

3

INVESTIGANDO DIFICULDADES E O USO DE LLMS PARA GLA

As revisões da literatura conduzidas no capítulo anterior evidenciam que a intersecção entre GLA e LLMs ainda não está consolidada e que a modelagem de dados ainda recebe atenção limitada nos estudos de GLA. Nesse sentido, este capítulo apresenta investigações sobre: (i) a dificuldade de modelagem de dados por estudantes de computação que atuavam como *designers* de aprendizagem (Seção 3.1); (ii) uma comparação de estruturas de captura geradas por um especialista e por um modelo de IA Gerativa, que evidencia o potencial dos LLMs no campo de GLA (Seção 3.2); e (iii) o uso de uma técnica interativa de *prompt* para aprimorar a qualidade de estruturas de captura, que demonstra contribuições em GLA (Seção 3.3). O capítulo pertence ao Ciclo de Relevância da DSR e visa atingir o segundo objetivo específico[1].

3.1 Análise das dificuldades de estudantes de computação em modelagem de dados

A motivação para esta primeira investigação surgiu em campo, a partir de observações de duas turmas em um período aproximado de um ano, durante a realização de monitoria não obrigatória na UEA. As disciplinas, denominadas "Oficina de Desenvolvimento

[1] Caracterizar dificuldades e estratégias utilizadas na modelagem de dados e em projetos de *prompts* para LLMs por *designers* de aprendizagem.

de *Software* Educacional I" e "Fundamentos de *Software* Educacional" (FSE), eram ministradas majoritariamente para o curso de Licenciatura em Computação, contendo 24 e cinco estudantes, respectivamente. Ambas as disciplinas envolviam a construção de artefatos educacionais, principalmente jogos. Os estudantes atuavam como *designers* de aprendizagem e deveriam implementar GLA em seus jogos educacionais. No entanto, observou-se que, além da dificuldade de uso do modelo GLBoard, apresentavam desafios em modelagem de dados: definir e justificar as variáveis de GLA que permitam a identificação de evidências de aprendizagem.

Portanto, a investigação apresentou como questão de pesquisa: QP3.1 - "Quais as dificuldades de estudantes de computação, que atuam como *designers* de aprendizagem, na modelagem de dados para jogos educacionais?". Para respondê-la, realizou-se um estudo empírico, cujos resultados foram publicados nos anais do WEI - CSBC (2025)[2].

3.1.1 Objetivo e participantes

Considerando a motivação da pesquisa, têm-se como objetivo identificar as dificuldades de estudantes que atuam como *designers* de aprendizagem em modelagem de dados de jogos educacionais, através da perspectiva desse público (autoavaliação).

Para a seleção de participantes, considerou-se dois critérios: (i) estudantes de Licenciatura em Computação da UEA que estavam cursando a disciplina de "Fundamentos de Software Educacional" (FSE), cujo objetivo era construir um jogo educacional e implementar técnicas de GLA, devendo realizar a modelagem de dados; e (ii) participantes de um laboratório de pesquisa e desenvolvimento de tecnologias educacionais, que estavam refatorando jogos já desenvolvidos, ao qual retornaram ao processo de modelagem de dados. Desse modo, todos os participantes atuavam como *designers* de aprendizagem, com o objetivo de construir/refatorar jogos educacionais. Dezesseis pessoas participaram da pesquisa, sendo que: onze (69%) se identificaram como gênero masculino e cinco (31%) como feminino, com uma faixa etária de 19 a 26 anos.

Em relação ao perfil e situação acadêmica dos participantes: 100% (n=16) esta-

[2] https://sol.sbc.org.br/index.php/wei/article/view/36257

vam em cursos de graduação em computação em andamento, dos quais doze (75%) de Licenciatura, três (19%) de Sistemas de Informação e um (6%) de Engenharia. A maioria estava no quarto ou sexto períodos do curso (n=12, 75%), enquanto os demais no oitavo ou superior. Todos (100%) já haviam cursado as três disciplinas introdutórias de programação, enquanto que apenas 63% (n=10) cursaram Banco de Dados I, 56% (n=9) cursaram Algoritmos e Estruturas de Dados II, e 50% (n=8) cursaram Modelagem e Projeto de Sistemas. Nessas disciplinas, nove (56%) apontaram não ter reprovado nenhuma vez, enquanto sete (44%) reprovaram ao menos uma vez em alguma delas. Quanto à experiência com jogos educacionais, dez participantes (63%) já haviam desenvolvido pelo menos um jogo. Apenas dois participantes (12%) haviam modelado dados e implementado técnicas de GLA previamente.

3.1.2 Procedimentos

A aplicação da pesquisa ocorreu em uma data combinada com os participantes, organizada em quatro momentos principais, como ilustra a Figura 6 e detalhada seguidamente.

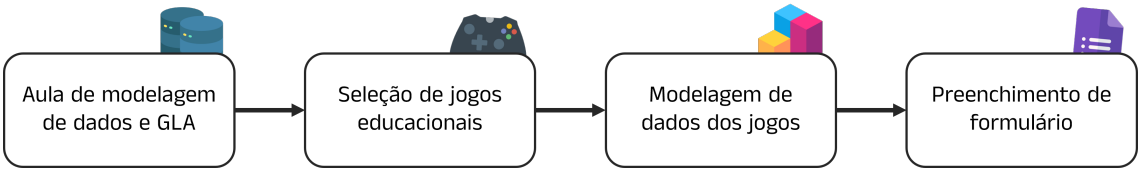


Figura 6 – Etapas da pesquisa sobre modelagem de dados.

Aula de modelagem de dados e GLA: visando introduzir o conteúdo teórico sobre modelagem de dados aos participantes, uma aula foi conduzida de forma presencial na universidade, que abordou: (i) jogos educacionais; (ii) lacunas em avaliação de aprendizagem em jogos; (iii) LA, GA e GLA; (iv) modelo GLBoard, incluindo a definição, arquitetura, *template* de dados, $path\_player$ e exemplos de captura; e (v) modelagem de dados. A aula estendeu-se por uma hora, na qual se utilizou uma apresentação multimídia com informações e figuras dos tópicos mencionados, com exemplos dinâmicos para os estudantes interagirem e praticarem os conteúdos.

Seleção de jogos educacionais: foi informado aos participantes que eles deve-

riam realizar, individualmente, a modelagem de dados de um jogo educacional. Para os estudantes da disciplina, a modelagem deveria ser referente aos próprios jogos que estavam desenvolvendo e, para os participantes do laboratório, um jogo educacional de livre escolha, desde que não houvesse duplicidade. Em seguida, uma planilha foi compartilhada com os participantes, que continha uma tabela com colunas de: nº, nome da variável (dado), tipo de dado, exemplo e justificativa de coleta, solicitando-os que fizessem uma cópia e iniciassem o preenchimento. Ademais, pediu-se aos participantes que realizassem a modelagem de forma manual, evitando o uso de ferramentas de GAI, como o ChatGPT, que poderia comprometer a validade da pesquisa.

Elaboração da modelagem de dados: nesse momento, os participantes deram início à modelagem dos dados com o preenchimento da planilha. Esse processo consistiu tanto do uso e análise dos jogos, quanto da proposta e inserção das variáveis na planilha, que se estendeu por aproximadamente uma hora e meia.

Preenchimento de formulário: a última etapa consistiu no preenchimento de um formulário em *Google Forms* para inserir as planilhas com as modelagens realizadas e avaliar a experiência da pesquisa. Ao passo em que os estudantes finalizavam suas modelagens, preenchiam o formulário e encerravam sua participação no estudo.

3.1.3 Coleta e análise de dados

Para a coleta dos dados, construiu-se um formulário autoral. Essa decisão foi fundamentada pela ausência de trabalhos localizados na literatura (até o presente momento da pesquisa) sobre dificuldades emergentes do processo de modelagem de dados para jogos educacionais. Diante disso, elaborou-se um *Google Forms*, contendo: (i) uma questão de consentimento para autorizar o uso dos dados de forma anônima com finalidade única de pesquisa; (ii) um campo para submissão das modelagens realizadas, em forma de planilha; (iii) questões quantitativas em escala Likert de 5 pontos; e (iv) questões qualitativas (perguntas abertas), ambas relacionadas à percepção dos participantes quanto à experiência realizada, definidas pelo mestrando e orientadoras conforme as necessidades da pesquisa (compreender o nível de dificuldade em modelar dados, a

influência do tempo e de conhecimentos prévios de disciplinas de computação nesse processo, entre outros). Desse modo, coletou-se dos participantes as respostas ao questionário de avaliação da pesquisa e seus artefatos (planilhas de modelagem de dados). O formulário completo pode ser consultado no Apêndice B.

Em relação à análise dos dados coletados, utilizou-se: (i) análise visual para as questões quantitativas e notas do formulário avaliativo das modelagens, por meio de *boxplot*; e (ii) análise temática sobre os dados qualitativos informados pelos participantes no formulário, no qual categorias foram criadas conforme leitura preliminar das respostas. Em seguida, as respostas foram organizadas nas categorias e analisadas em conjunto, para identificar padrões e interpretar os achados de cada categoria. Essa forma de análise foi escolhida em virtude da experiência do mestrando e por se tratar de um método adequado para examinar percepções e evidências qualitativas, que não segue necessariamente um processo rigoroso e sistemático[3].

3.1.4 Resultados

Os resultados deste estudo envolvem a análise das questões quali-quanti, respondidas pelos estudantes de computação, a respeito da modelagem de dados para jogos educacionais. A Figura 7 contém um *boxplot* relacionado à avaliação das questões quantitativas, descritas a seguir e enumeradas conforme sua ordem no formulário (Apêndice B).

Compreensão do objetivo (questão 5): onze participantes (69%) atribuíram notas "4" ou "5". Esse resultado evidencia que a maioria dos participantes considera que compreendeu o objetivo da modelagem de dados, que consiste em definir quais dados devem ser capturados e justificar sua relevância para orientar a implementação de GLA. Em relação à questão "O que você considera ter feito bem na modelagem de dados?" (questão 6), as respostas foram diversas: (i) incoerentes com a pergunta (n=3, 19%). Por exemplo: P15 – "não"; (ii) genéricas, sem entrar em detalhes (n=5, 31%). Por exemplo: P2 – "ter feito escolhas bem objetivas"; (iii) entendimento nulo ou com dificuldade (n=5, 31%). Por exemplo: P9 – "poderia melhorar mais"; e (iv) detalhadas,

[3] Esse mesmo critério orientou a escolha da análise temática nas investigações subsequentes.

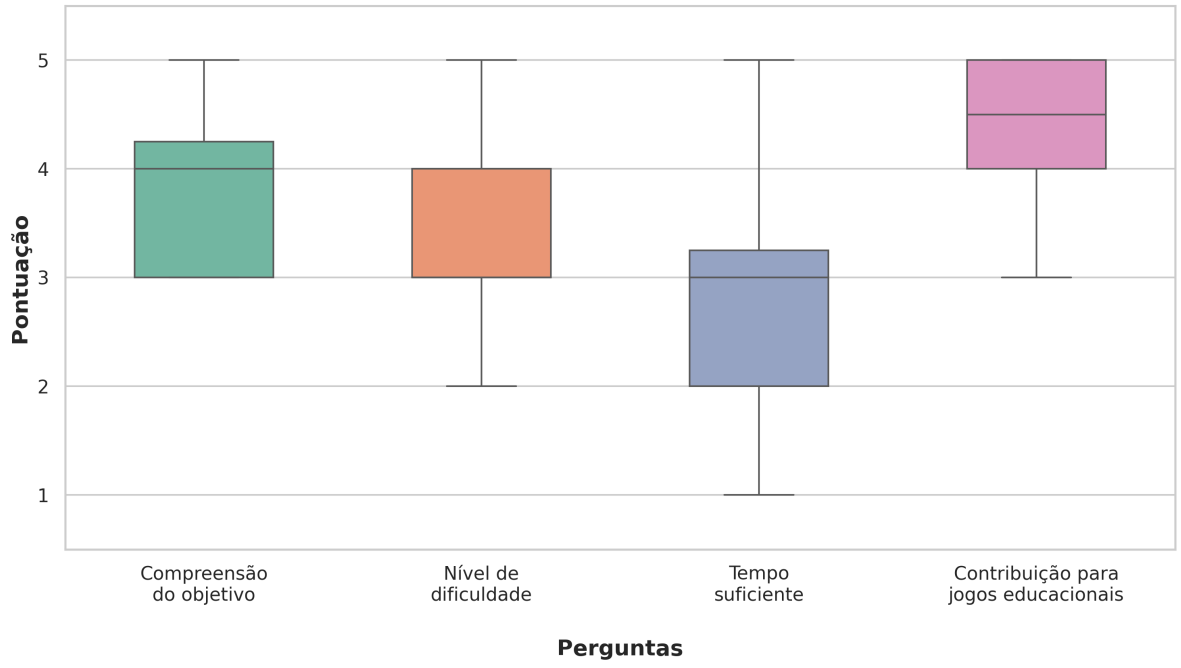


Figura 7 – *Boxplot* com a avaliação das questões quantitativas.

com informações sobre os jogos em questão (n=3, 19%). Por exemplo: P16 – "identificar como o jogador perdeu ou ganhou a fase". Relacionando as respostas com a questão anterior, observa-se que os participantes compreenderam o objetivo da modelagem de dados, mas tiveram dificuldade em apontar o que fizeram bem nesse processo.

Nível de dificuldade (questão 7): somente um (6%) participante achou "fácil" modelar dados, enquanto os demais consideraram que foi "médio" (n=8, 50%) ou "difícil" (n=7, 44%). Esse ponto reforça o entendimento de que tarefas de modelagem de dados em GLA tendem a ser complexas para estudantes que atuam como *designers* de aprendizagem, sobretudo ao considerar que: (i) mais da metade do participantes já tiveram experiência com o desenvolvimento de jogos educacionais (n=9, 56%); (ii) os dois únicos participantes com experiência em implementação de técnicas de GLA (12%) consideraram o processo "médio" e "difícil"; (iii) dez participantes (63%) informaram ter um bom nível de conhecimento em lógica de programação; (iv) a maioria (n=10) estava no 6º período ou superiores do curso; e (v) o nível de conhecimento em banco de dados dos participantes, que apontaram ser "médio" (n=9, 56%) e "bom" (n=3, 19%).

As dificuldade apontadas em relação a modelagem de dados (questão 8) referiam-

se a definir os dados/variáveis para coleta (n=11, 69%) e a abstração desse processo (n=5, 31%), dos quais dois comentários dos participantes destacam-se: "Pensar em todas as condições que o jogador enfrenta" (P16) e "Sem ter muita certeza do que se procura, é um pouco difícil apontar variáveis" (P8). Tais comentários refletem a dificuldade de definir quais dados de interação devem ser coletados para fornecer informações sobre a aprendizagem dos jogadores (Cano; Fernández-Manjón; García-Tejedor, 2017).

Tempo suficiente (questão 9): em relação ao tempo destinado à modelagem dos dados, que durou aproximadamente uma hora e meia, sete participantes (44%) o consideraram insuficiente e cinco (31%) o avaliaram como mediano. Isso também pode sugerir que a modelagem é uma etapa complexa, que demanda não apenas conhecimento, mas também tempo, variável conforme as características do jogo. Por exemplo, mecânicas robustas, diferentes fases e outros elementos do *game design* impactam diretamente o tempo e a complexidade da modelagem dos dados e, consequentemente, a implementação de GLA (Macena et al., 2024).

As respostas dos participantes quanto à questão "Houve algo que você aprendeu durante o processo de modelagem que não conhecia antes?" (questão 10), foram que: não aprenderam nada novo (n=9, 56%), informações genéricas (n=2, 12%) e "sim" (n=5, 31%), com destaque a: (i) "o uso de listas para capturar mais detalhadamente alguns dados" (P1). Essa resposta refere-se à uma estratégia importante para a modelagem, pois essas estruturas facilitam a organização e hierarquia dos dados; (ii) "analisar o que as variáveis podem dizer sobre o jogador" (P4). Esse comentário reflete com exatidão o propósito da modelagem de dados: a identificação dos dados para coletar evidências que tragam *insights* sobre o comportamento e trajetória do jogador; e (iii) "ao longo do processo foi ficando mais fácil identificar quais dados eu deveria armazenar" (P15). Esse ponto está relacionado de forma direta com o tempo para modelar dados: quanto mais o participante dedica-se a esse processo, mais consegue abstrair informações sobre o jogo e localizar variáveis necessárias para a coleta.

Contribuição para jogos educacionais (questão 11): por mais que o nível de dificuldade de modelar dados tenha sido alto, bem como o tempo sendo avaliado como insuficiente, quatorze participantes (88%) acreditam que a modelagem contribui para o

desenvolvimento de jogos educacionais mais eficazes. Em relação ao "como" (questão 12), relataram: a diminuição do tempo de processo de implementação (n=1, 6%), a definição de variáveis para estimar possíveis indícios de aprendizagem (n=2), o suporte ao processo de captura de dados (n=6, 37%), a percepção melhor dos aspectos gerais e específicos do jogo (n=2), a entender o comportamento do jogador (n=2), a identificar falhas no *game design* (n=2), entre outros (n=2). Ademais, seis participantes (37%) apontaram que a modelagem foi semelhante a algo que haviam realizado previamente (questões 13 e 14), como: diagrama de classe (n=1), levantamento de requisitos (n=1), modelagem de banco de dados (n=1), entre outros (n=3, 19%). Nesse aspecto, os participantes apontaram que o conhecimento prévio de disciplinas de computação auxiliou a modelar dados (questões 15 e 16), como: programação (introdutória, estruturas de dados e orientada a objetos, n=5), modelagem de sistemas (n=3), banco de dados (n=3) e disciplinas teóricas pedagógicas (n=1).

As contribuições que essas disciplinas trouxeram para a modelagem de dados, conforme o relato dos participantes (questão 16), foram: criação de diagramas (n=1, 6%), modelagem de sistemas/banco de dados (n=2), análise de entidades e necessidades do banco de dados (n=1), identificação de variáveis (n=2) e teorias de aprendizagem (n=1). Esses pontos também podem apoiar a complexidade de modelar dados, de modo que conhecimentos de disciplinas anteriores de computação são fundamentais nesse processo. Entretanto, nem todos os participantes haviam cursado essas disciplinas quando a investigação foi realizada: 63% (n=10) cursaram Banco de Dados (BD), 50% (n=8) Modelagem e Projeto de Sistemas (MPS) e 56% (n=8) Algoritmos e Estruturas de Dados II. Além disso, os participantes apontaram que possuíam um nível de conhecimento médio ou baixo nessas disciplinas (81%, n=13, em BD e 94%, n=15, em MPS).

3.1.5 Considerações sobre a investigação

A pesquisa demonstrou que os estudantes, que atuavam como *designers* de aprendizagem, em maioria, compreenderam o objetivo da modelagem de dados e consideraram que essa etapa contribui para o desenvolvimento de jogos educacionais mais eficazes.

Esse resultado corrobora a importância da modelagem de dados para GLA. Entretanto, o nível de complexidade para realizá-la foi avaliado como alto e o tempo destinado à atividade foi considerado insuficiente. Portanto, em resposta à QP3.1, as dificuldades identificadas de modelagem de dados de jogos educacionais foram: (i) o nível de abstração desse processo; (ii) a apropriação do jogo para definir os dados a serem coletados; (iii) o tempo disponível; e (iv) os conhecimentos prévios, oriundos de disciplinas anteriores de Computação, que auxiliam na modelagem. Esses itens sugerem que, embora a modelagem seja uma etapa fundamental em GLA, não foi considerada como trivial por estudantes que atuavam como *designers* de aprendizagem.

As limitações da pesquisa incluem: (i) o tempo, pois a aplicação ocorreu numa única aula da disciplina, podendo ter interferido na experiência dos participantes e na qualidade das modelagens. Entretanto, esse aspecto constituiu um achado, indicando que o tempo influencia a complexidade de modelar dados; e (ii) a perspectiva, por considerar somente a percepção de estudantes de computação quanto à modelagem de dados. Desse modo, não é possível generalizar os resultados para *designers* de aprendizagem que não pertencem à área ou que já possuem experiências consolidadas. Apesar dessa limitação, que se pretende mitigar em pesquisas futuras, essa perspectiva correspondeu ao escopo da investigação.

3.2 Geração de estruturas de captura de GLA: LLM vs especialista

A partir da pesquisa anterior, observou-se a dificuldade de estudantes que atuavam como *designers* de aprendizagem em modelar dados para jogos educacionais. Em paralelo, analisando a emergência da área de GAI com os LLMs, observou-se sua aplicação em domínios baseados em dados, como *Learning Analytics* (Mazzullo et al., 2023; Misiejuk et al., 2025). Dessa forma, dadas as capacidades desses modelos em auxiliar com tarefas que envolvem proposta, análise e interpretação de dados, podem ser uma potencial alternativa no campo de GLA para mitigar a complexidade de modelar dados para jogos educacionais. Essa foi a motivação dessa investigação.

Portanto, definiu-se como questão de pesquisa: "QP3.2 - Qual a qualidade das estruturas de captura para jogos educacionais geradas por um LLM, em comparação às elaboradas por um especialista em GLA?". Essas estruturas são dependentes da modelagem de dados, visto que se tratam da organização das variáveis de GLA em um *template* específico. Um Experimento Controlado foi conduzido para responder a QP3.2, elaborado conforme as diretrizes de Wohlin et al. (2012) e descrito a seguir.

3.2.1 Escopo

A definição de metas e objetivo do experimento foi realizada com base na abordagem GQM (*Goal-Question-Metric*) (Basili; Rombach, 1988), descrito na Tabela 3.

Analisar	Estruturas de captura de dados em jogos educacionais conforme *template* do GLBoard
Para o propósito de	Comparar técnicas de modelagem de dados (manual e por LLM)
No que diz respeito a	Qualidade
Do ponto de vista	De especialistas em jogos educacionais
No contexto de	*Game Learning Analytics*

Tabela 3 – Objetivo da pesquisa, conforme a abordagem GQM.

3.2.2 Planejamento

Contexto: este experimento focou-se em comparar estruturas de captura de dados no *template* GLBoard, geradas por diferentes técnicas de modelagem de dados: (i) manualmente; e por meio de interação com um LLM, o ChatGPT, especificamente suas versões (ii) 3.5; (iii) 4; e (iv) 4o. Um especialista em GLA e com experiência em LLM, sendo o mestrando desta dissertação, foi o responsável por elaborar o *prompt* para enviar ao LLM a partir de um jogo educacional, e assim gerar as estruturas.

Hipóteses: conforme a questão de pesquisa e seu objetivo, foram formuladas as hipóteses estatísticas a seguir, visando compreender se as estruturas geradas por LLM apresentariam diferenças de qualidade quando comparadas à estrutura elaborada manualmente por especialista.

- **Hipótese Nula (H0):** não há diferença estatística na qualidade das estruturas de GLA entre as técnicas de modelagem de dados utilizadas;
- **Hipótese Alternativa (H1):** há diferença estatística na qualidade das estruturas de GLA entre as técnicas de modelagem de dados utilizadas, em pelo menos um par de técnicas.

Variáveis: a Variável Independente (VI) deste experimento é a técnica de geração de estruturas de GLA, manipulada em quatro níveis (tratamentos): manual, ChatGPT 3.5, ChatGPT 4 e ChatGPT 4o. O objetivo é analisar seu efeito sobre a Variável Dependente (VD), que é a qualidade das estruturas, operacionalizada a partir das pontuações atribuídas por especialistas no instrumento "*Player Level Up!*". Esse instrumento considera os aspectos de coerência, redundância/completude e evidência de evolução, descrito no item "Instrumentos".

Seleção de participantes: a investigação contou com a participação de cinco especialistas em GLA, que avaliaram todas as estruturas de captura (dados pareados). O critério de seleção foi por conveniência, pertencendo ao mesmo grupo de pesquisa (agilizando o contato) e pela experiência com GLA. Em relação ao perfil dos especialistas: todos (100%) possuíam entre quatro a sete anos de experiência com desenvolvimento de jogos e já haviam implementado técnicas de GLA utilizando o GLBoard; dois (40%) eram acadêmicos de cursos de graduação em computação, enquanto os outros três (60%) faziam pós-graduação em informática (dois mestrandos e um doutorando); e quanto a publicações científicas relacionadas a jogos, todos possuíam quantidades distintas: uma, quatro, nove, dezenove e vinte e dois.

Design Experimental: o padrão de *design* utilizado é um fator com mais de dois tratamentos, cujo fator é a técnica de modelagem de dados e os tratamentos são "manual", "ChatGPT 3.5", "ChatGPT 4" e "ChatGPT 4o". Em relação à estratégia estatística, considerando que o experimento contaria com dados ordinais e pareados, definiu-se o teste não-paramétrico de Friedman (Friedman, 1937).

Materiais: para a condução da pesquisa, foi definido o jogo educacional e os dados correspondentes que iriam compor o *prompt* enviado ao LLM para gerar a estrutura de captura. O jogo selecionado foi o "Cadê minha pizza?" (Honda et al.,

2023), descrito no Apêndice A, cujo objetivo é auxiliar na aprendizagem de operações matemáticas e caminho mínimo (Teoria dos Grafos), em que o jogador deve entregar pizzas pela cidade considerando a gasolina dos entregadores e os custos das rotas para as residências. Sua escolha deu-se por ser de autoria do mestrando e pelas mecânicas serem simples, o que favoreceu a geração das estruturas, dado o conhecimento do especialista sobre suas mecânicas e elementos de aprendizagem. Para a seleção de quais dados seriam incorporados ao *prompt* para enviar ao LLM, considerou-se elementos da documentação do jogo (EGDD)[4] aos quais o mestrando e orientadoras consideraram ser fundamentais para o modelo compreendê-los, sendo a base de um jogo educacional:

- Público-alvo: define a faixa etária pretendida e o perfil cognitivo dos estudantes;
- Conteúdo educacional: determina os conteúdos curriculares presentes no jogo, estabelecendo os objetivos de aprendizagem;
- História: detalha a narrativa que contextualiza, de forma lúdica, as ações do jogador;
- *Gameplay*: representa o conjunto de interações e regras do sistema as quais o jogador está sujeito;
- Mecânica de aprendizagem: relaciona a forma que o *level design* foi entrelaçado aos conteúdos curriculares para promover o exercício e eventual aprendizagem.

Instrumentos: dentre os instrumentos utilizados nesta pesquisa, estão:

- (i) a ferramenta de GLA: as estruturas de captura geradas basearam-se no *template* de dados do GLBoard (Silva et al., 2022), devido ao manual estar no idioma nativo dos autores (língua portuguesa) e a experiência prévia com o modelo;
- (ii) o LLM selecionado para gerar as estruturas: o ChatGPT, devido à possibilidade de utilizar suas três versões (3.5, 4 e 4o) e por ser uma das ferramentas de IA Gerativa mais utilizadas, com mais de 28.000 publicações científicas que a mencionam na base de dados Scopus (dados de janeiro de 2026). Apesar das

[4] Informações sobre o EGDD: drive.google/Guia_para_EGDD

possibilidades de personalização oferecidas por modelos abertos (como o LLaMa e Mistral), seu uso requer conhecimento técnico e uma infraestrutura mínima, o que poderia comprometer a replicabilidade da pesquisa. Esse aspecto também foi considerado na escolha do ChatGPT, que possui uma interface *on-line* de fácil acesso para interação com o *chatbot*;

- (iii) a técnica de *prompt*: optou-se pelo uso de *basic prompting* (Shin et al., 2023; Liu et al., 2023). A escolha considerou a simplicidade e a facilidade de uso da técnica: elaborar um conjunto de instruções ao modelo para obter as respostas desejadas;

- (iv) o aprendizado por contexto (ICL) utilizado para interagir com o modelo: fez-se uso de *zero-shot learning*, visando analisar o desempenho do modelo sem exemplos no *prompt*. Essa decisão foi para poder analisar, de forma isolada, a geração das estruturas de GLA a partir da base de dados do modelo;

- (v) o instrumento de avaliação das estruturas: "*Player Level Up!*", com sete questões quantitativas (Likert-5) organizadas em três dimensões (coerência, redundância/completude e evidência de evolução), exibido na Tabela 4. Esse instrumento foi elaborado de forma autoral por quatro especialistas em GLA, distintos dos avaliadores desta investigação, em decorrência de dois fatores principais: as estruturas são baseadas no *template* de dados do GLBoard e não se identificaram instrumentos na literatura que façam avaliação da modelagem de dados. Para a construção do instrumento, realizou-se uma sessão assíncrona de discussão, mediada pelo mestrando, onde uma proposta inicial de questões avaliativas foi apresentada. Cada especialista analisou o instrumento, verificou a coerência das questões com as dimensões propostas e sugeriu ajustes, que posteriormente foram incorporados. Esse processo serviu também para assegurar a validade da VD, que foi operacionalizada a partir desse instrumento.

Dimensão	**Nº**	**Questão**
Coerência	1	Os nomes das variáveis são consistentes com os dados, ou seja, está claro identificar ao que a variável se refere
	2	As variáveis definidas na estrutura são direta e claramente relacionadas à mecânica de jogo
	3	Os tipos de dados atribuídos a cada variável aparentam ser apropriados para o que se pretende capturar
	4	As variáveis definidas pertencem às interações do jogador ($path\_player$)
Redundância e completude	5	Não existem variáveis que capturam informações sobrepostas/duplicadas
	6	Não se identificou, na estrutura, variáveis que poderiam ser importantes de capturar e não foram consideradas
Evidência de Evolução	7	A estrutura permite análise temporal do percurso, como a duração gasta em cada fase ou tarefa ou *timestamps*

Tabela 4 – Instrumento de avaliação das estruturas de captura no *template* GLBoard.

3.2.3 Operação

Neste tópico, consta o detalhamento do processo de condução do experimento, que se divide em duas partes principais: (i) geração das estruturas de captura; e (ii) coleta de dados com as avaliações das estruturas pelos especialistas.

Geração manual da estrutura: para a estrutura elaborada manualmente, o especialista em GLA (mestrando) analisou as mecânicas e objetivos de aprendizagem do jogo. Em seguida, definiu, em um documento, os dados que iria capturar e a justificativa pelo qual são necessários para expressar evolução do jogador, realizando assim a modelagem de dados. Dado o objetivo do jogo, em que o jogador deve entregar pizzas em residências considerando a gasolina disponível e os custos das rotas, as variáveis definidas foram: (i) entregador selecionado (ou nulo, caso não haja); (ii) o identificador da residência de entrega; (iii) o status da entrega, podendo ser falha (não adjacente ou sem gasolina disponível) ou sucesso; e (iv) o registro de tempo (*timestamp*) da ação. Diante disso, as variáveis foram organizadas conforme o *template* de captura do GLBoard (formato JSON), resultando na estrutura da Figura 8, com um exemplo.

Geração das estruturas via LLM: refere-se a geração das estruturas de captura por meio de um LLM, utilizando as versões 3.5, 4 e 4o do ChatGPT e a técnica de *basic prompting* com *zero-shot learning* (sem exemplos). Diante disso, o mestrando elaborou o *prompt* que seria enviado ao modelo, segmentado em quatro componentes principais: (i) introdução: explicando o contexto e objetivo da atividade; (ii) GLBoard: apresentando o *template* de captura de dados do GLBoard e explicando cada variável; (iii) requisitos:

E1

```
"path_player": [
    {
        "delivery_person": string
        "destination": string
        "delivery_status": string
        "timestamp": DateTime
    }]
```

Exemplo

```
"path_player": [
    {
        "delivery_person": "Deliverer_person 1"
        "destination": "House 5"
        "delivery_status": "Non-adjacent"
        "timestamp": 22:59:40
    }]
```

Figura 8 – Estrutura JSON de captura e exemplo, respectivamente.

indicando ao modelo o que deveria ser realizado, dando ênfase no $path\_player$; e (iv) solicitação de espera: pedindo ao modelo que aguarde o recebimento dos dados do jogo para gerar as estruturas em formato JSON seguidamente. O *prompt* completo pode ser visualizado no Apêndice C. Em relação aos dados informados sobre o jogo, optou-se por inseri-los em um segundo *prompt*, para evitar sobrecarregar o modelo em uma única entrada. Em seguida, o par de *prompts* (instruções e informações do jogo) foi aplicado nas versões do ChatGPT (iniciando pela 3.5, depois 4 e 4o), onde cada conversa foi iniciada em um novo *chat* de interação, com o objetivo de evitar que o modelo "aprendesse" com as respostas anteriores. As respostas completas de cada modelo podem ser consultadas no link[5] e as estruturas estão disponíveis na Figura 9.

Ordem de definição das estruturas: as estruturas foram elaboradas inicialmente de forma manual pelo mestrando e, em seguida, geradas de forma artificial pelas versões 3.5, 4 e 4o do ChatGPT. O mesmo *prompt* e dados do jogo foram informados ao modelo, a fim de assegurar comparabilidade e evitar interferência humana nas respostas.

Coleta de dados: utilizou-se um formulário em *Google Forms* para coletar as avaliações dos especialistas sobre as estruturas de capturas de dados, contendo: (i) *links* da

[5] https://doi.org/10.6084/m9.figshare.29433452



Figura 9 – Estruturas geradas, respectivamente, com o ChatGPT 3.5, 4 e 4o.

planilha de informações e da pasta com as telas do jogo, para facilitar na compreensão; (ii) as imagens das estruturas geradas, tanto de forma manual quanto pelo ChatGPT; (iii) sete questões quantitativas correspondentes ao instrumento "*Player Level Up!*" para avaliar a qualidade das estruturas; e (iv) uma questão aberta (opcional). As estruturas foram apresentadas aos avaliadores no formulário em uma ordem fixa (manual, ChatGPT 3.5, 4 e 4o), de forma anônima para evitar vieses[6]. Embora a sequência não tenha sido aleatorizada, as avaliações foram feitas de modo independente e assíncrono. Uma nota individual foi gerada para cada estrutura pelo especialista avaliador, calculada a partir da soma dos valores das questões em escala Likert. Desse modo, a nota poderia variar entre 7 (o especialista avaliou todas as questões com a nota 1) e 35 (nota máxima atribuída à todas as questões).

[6] Os participantes não foram informados sobre a autoria das estruturas.

3.2.4 Análise e interpretação

Na análise de dados, fez-se uso de gráficos *boxplot* para a visualização dos resultados de avaliação dos grupos (estruturas) e, para verificar a diferença estatística, utilizou-se o teste não paramétrico de Friedman (Friedman, 1937). A amostra foi composta pelos cinco especialistas em GLA, que avaliaram todas as estruturas de captura (n=5, dados pareados). O teste é compatível com dados ordinais e pareados, como neste experimento. A Figura 10 apresenta o *boxplot* com as avaliações, obtidas por meio do somatório das questões de cada estrutura por especialista.

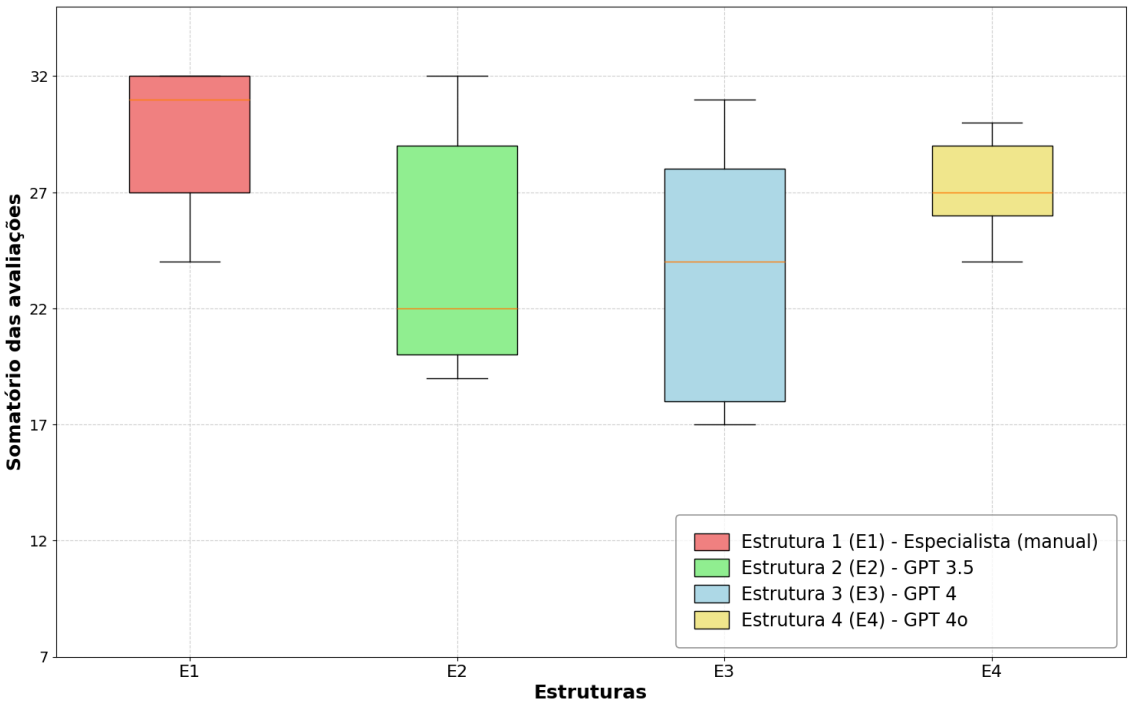


Figura 10 – *Boxplot* com as avaliações gerais de cada estrutura.

O gráfico evidencia que a E1, elaborada de forma manual pelo especialista, foi a mais bem avaliada quando comparada com as geradas pelo ChatGPT (E2, E3 e E4). As estruturas E2 e E3, geradas, respectivamente, pelas versões 3.5 e 4 do ChatGPT, obtiveram os menores valores de avaliação, com grande dispersão dos dados Já E4 foi a segunda mais bem avaliada, com a menor dispersão entre as estruturas analisadas e valores mais próximos de E1 no somatório das avaliações.

Para verificar possíveis diferenças significativas na qualidade das estruturas, operacionalizada por meio das avaliações dos especialistas, fez-se uso do teste não-

paramétrico de Friedman. Desse modo, os grupos foram organizados com base na planilha de avaliação, segmentando os dados conforme os tratamentos analisados: manual e versões 3.5, 4 e 4o do ChatGPT. Em seguida, o teste foi aplicado, calculando-se o valor de χ^2 (qui-quadrado) e o *p-value*. O resultado obtido foi $\chi^2 = 2.22$ e $p \approx 0.528$ ($p > 0.05$), indicando que não há evidência estatística suficiente para apontar diferenças na qualidade das estruturas de GLA geradas pelas diferentes técnicas de modelagem de dados. Desse modo, a Hipótese Nula (H0) não pôde ser rejeitada.

Com o objetivo de compreender com mais clareza a distribuição dos dados, a Tabela 5 inclui a estatística descritiva, com as estruturas avaliadas e suas medianas, modas e amplitudes. A Tabela apresenta uma preferência sutil de E1 pelos especialistas, com a maior mediana (31) e moda (32). Como apontado pelo *boxplot*, E2 e E3 obtiveram uma alta dispersão de dados, com os maiores valores de amplitude (13 e 14, respectivamente). A estrutura E3 (versão 4 do ChatGPT), inesperadamente, apresentou resultados inferiores à E2 (versão 3.5), com a menor moda (17) e maior dispersão. Em contrapartida, E4 demonstrou uma melhor consistência entre as estruturas geradas por LLM, com mediana de 27 e menor amplitude geral (6), indicando uma estabilidade maior de resposta dos especialistas à estrutura gerada pelo ChatGPT 4o. Ademais, as medianas e modas elevadas de E1 e E4, assim como os menores valores de amplitude, sugerem proximidade descritiva entre as avaliações atribuídas a essas estruturas. A análise, de forma geral, apresenta que as estruturas foram avaliadas positivamente: modas superiores à 16 e medianas maiores que 21 (valor intermediário da escala).

Tabela 5 – Estatística descritiva por estrutura.

Estrutura	Mediana	Moda	Amplitude
E1	31.0	32	8
E2	22.0	19	13
E3	24.0	17	14
E4	27.0	24	6

Inicialmente, esperava-se que as estruturas geradas por LLM fossem mais bem avaliadas nas dimensões de qualidade adotadas no instrumento "*Player Level Up!*" do que a estrutura gerada manualmente pelo especialista, o que poderia indicar diferença estatística entre as técnicas analisadas. No entanto, a Hipótese Nula (H0) não pôde

ser rejeitada. Complementarmente, fez-se uso de *boxplots* e estatística descritiva para compreender a distribuição das notas atribuídas pelos avaliadores. A partir da análise com esses recursos, observou-se que a E4 (gerada pelo ChatGPT 4o) obteve avaliações positivas e próximas às da E1 (gerada manualmente pelo especialista). Esse resultado, ainda que não permita afirmar equivalência ou superioridade estatística, sugere indícios de que LLMs podem apoiar a modelagem de dados para gerar estruturas de GLA.

Embora não seja possível generalizar os achados da investigação devido à amostra reduzida, os resultados sugerem que modelos de IA Gerativa podem apoiar a construção de estruturas de GLA personalizadas, avaliadas, em sua maioria, com notas superiores ao valor intermediário da escala (21). Em resposta à QP3.2, a análise descritiva sugere que as estruturas de GLA geradas pelas versões 3.5 e 4 obtiveram avaliações inferiores à estrutura elaborada pelo especialista, enquanto a estrutura gerada pelo ChatGPT 4o apresentou qualidade mais próxima à dela. Esse resultado pode indicar o potencial do apoio de LLMs à modelagem de dados, visto que: (i) não é considerada uma atividade trivial, sobretudo por estudantes que atuaram como *designers* de aprendizagem; (ii) o especialista possuía anos de experiência com jogos educacionais; e (iii) o modelo foi avaliado em condições limitadas, com poucas informações sobre o jogo educacional e utilizando *zero-shot learning* (sem exemplos no *prompt*).

Entretanto, não se pode afirmar que as estruturas sejam equivalentes, pois não foi realizado um teste de equivalência, nem que sejam diferentes, uma vez que a Hipótese Nula não foi rejeitada. Nesse sentido, o *boxplot* e a análise descritiva apenas complementam a análise sob outra perspectiva, não substituindo o teste estatístico. Uma das limitações da pesquisa é considerar somente a percepção dos especialistas, que podem não ter compreendido algum aspecto do jogo com clareza, refletindo em uma avaliação não realista. É válido também ressaltar que, por mais que o mestrando seja autor do jogo e possua experiência com modelagem manual, isso não influenciou a concepção dos *prompts* a ponto de ser considerado um viés, pois seguiram um formato fixo e apenas incluíram as informações do jogo, sem ajustes entre as versões. Outra limitação refere-se ao tamanho reduzido da amostra, o que dificulta a generalização dos resultados e pode limitar o poder estatístico do teste.

3.3 Investigação de uma técnica de *prompt* interativa de LLM para GLA

A investigação anterior revelou indícios do potencial de LLMs para auxiliar no campo de GLA, ainda que as estruturas geradas não tenham sido superiores à de um especialista. Esse aspecto pode ter relação com o contexto limitado fornecido ao modelo, a partir do qual se supõe que a incorporação de mais dados do jogo na entrada e o uso de uma técnica interativa de *prompt* podem aprimorar os resultados. Essa é a motivação desta investigação, em forma de estudo empírico, cuja questão de pesquisa é: "QP3.3 - Como uma técnica de *prompt* com questionamentos e a quantidade de dados de jogos inseridos na entrada influenciam na qualidade das estruturas de captura de dados geradas por um LLM no *template* GLBoard?". Este estudo resultou em uma publicação científica, disponível nos Anais do EduComp 2026[7].

3.3.1 Objetivo, contexto e participantes

Nesta investigação, buscou-se analisar a geração de estruturas de captura de dados no *template* do GLBoard, investigando o impacto da inclusão de mais dados do jogo no *prompt* e da utilização de uma técnica interativa de questionamentos. Desse modo, três abordagens distintas foram adotadas para a geração de estruturas de GLA: (i) manualmente, por estudantes; (ii) com um LLM e técnica de *basic prompting*, por um especialista em GLA; e (iii) com um LLM e uma técnica interativa de questionamentos inspirada em *chain-of-thought prompting* (Wei et al., 2022), pelos mesmos estudantes. Diante disso, três grupos de participantes foram envolvidos na pesquisa:

- Estudantes de computação, que geraram as estruturas manualmente e com o LLM de forma interativa. Dez universitários da UEA foram selecionados, sendo 90% (n=9) do curso de Licenciatura em Computação e 10% (n=1) de Sistemas de Informações. A maioria encontrava-se no 6º período (n=4, 40%) ou períodos superiores ao 8º (n=6, 60%), cuja seleção foi por conveniência: no âmbito da

[7] https://sol.sbc.org.br/index.php/educomp/article/view/42634

disciplina de FSE, os estudantes tinham como objetivo implementar técnicas de GLA em seus jogos, devendo realizar previamente a modelagem de dados. Dessa forma, todos atuavam como *designers* de aprendizagem de jogos educacionais;

- O especialista em GLA que gerou as estruturas com o LLM através de *basic prompting*, sendo o autor da pesquisa (mestrando), com experiência em jogos educacionais, GLA e projeção de *prompts*;

- Três especialistas em GLA, que avaliaram as estruturas geradas, selecionados pelos anos de experiência em construção de jogos educacionais e implementação de técnicas de GLA, bem como por conveniência (mesmo grupo de pesquisa). Todos possuíam formação em Licenciatura em Computação e eram estudantes de Mestrado/Doutorado de uma universidade pública.

3.3.2 Instrumentos e procedimentos

Os mesmos instrumentos da investigação anterior (Seção 3.2) foram utilizados: GLBoard como ferramenta de GLA, ChatGPT como LLM, *basic prompting* como técnica de *prompt*, *zero-shot* no aprendizado por contexto, visando analisar a resposta do modelo unicamente de sua base de dados (sem exemplos), e "*Player Level Up!*" para avaliação das estruturas. No entanto, alguns aspectos diferiram-se: (i) a versão do ChatGPT utilizada foi somente a 3.5, pois a 4 não era gratuita; e (ii) outra técnica de *prompt* foi investigada: uma estratégia interativa com questionamentos inspirada em *chain-of-thought prompting*, dados seus resultados positivos na pesquisa de Wei et al. (2022). As etapas conduzidas na pesquisa podem ser visualizadas na Figura 11.

Mapeamento de atributos de jogos educacionais: o primeiro passo consistiu na definição de quais elementos de jogos seriam enviados para o LLM. Na investigação anterior, com somente quatro elementos, o ChatGPT foi capaz de gerar estruturas de GLA que foram avaliadas, em maioria, com notas positivas por especialistas. Assim, pressupõe-se que, ao aumentar a quantidade de elementos, a qualidade das estruturas de GLA também pode aumentar. Nesse aspecto, utilizando como base o modelo de EGDD (Pires et al., 2021), 26 elementos foram selecionados para composição do

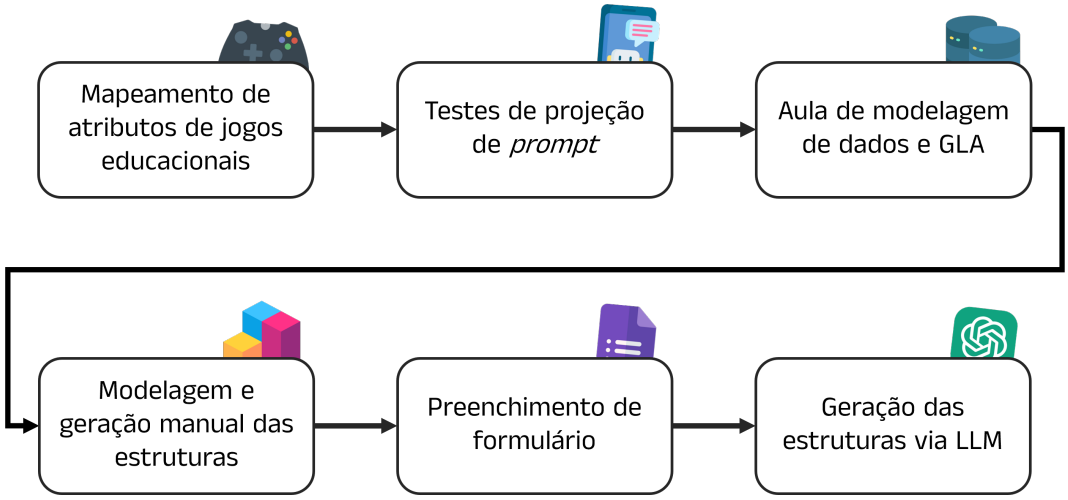


Figura 11 – Etapas da pesquisa sobre *Prompting* em LLM para GLA

prompt a ser enviado ao LLM, dos quais nove são obrigatórios (nome do jogo, tema, conteúdo, objetivo educacional, público-alvo, história, jornada do herói, mundo do jogo e *gameplay*). Esses elementos foram organizados em três categorias (nome do jogo e seus fundamentos, informações de *game design* e aspectos de aprendizagem), cujos nomes e descrições podem ser consultados no Apêndice D.

Testes de projeção de *prompt*: nessa etapa, o especialista em GLA conduziu testes para projetar um *prompt* de forma a obter os resultados desejados pelo LLM. Como resultado dos testes com o ChatGPT 3.5, a versão final do *prompt* foi projetada, disponível no Apêndice E. Ele é composto por quatro componentes principais: (i) contexto, explicando o objetivo geral ao modelo; (ii) *template* do GLBoard, exibindo o padrão da estrutura de captura do modelo e detalhando seus dados; (iii) objetivos, indicando ao modelo o que fazer e as especificações; e (iv) conclusão, informando ao modelo para aguardar o envio das informações do jogo e ressaltando que a estrutura deve ser em formato JSON. Esse último tópico foi necessário, devido a ocasiões onde o modelo não retornava a estrutura em formato JSON e indicava variáveis relacionadas ao contexto do jogo, mas não que não expressavam a evolução do jogador em uma fase.

Aula de modelagem de dados e GLA: a próxima etapa tratou-se da condução de uma aula para introduzir o conteúdo de modelagem de dados e GLA aos participantes, que incluiu os seguintes tópicos: (i) definição e importância de modelar dados; (ii) exemplos práticos; (iii) alertas e dicas para a modelagem de dados; (iv) modelo GLBoard:

definição, arquitetura e *template* de dados; e (v) conclusão, atividade e referências.

Modelagem e geração manual das estruturas: a atividade solicitada durante a aula refere-se ao processo de modelagem de dados. Desse modo, nessa etapa, os estudantes definiram quais dados iriam capturar de seus jogos, junto com a justificativa de coleta, por meio de uma planilha *on-line*. Em seguida, iniciaram a criação das estruturas de forma manual (Abordagem 1): colocaram os dados no *template* do GLBoard em formato JSON, focando-se no campo $path\_player$ para a coleta de variáveis que possam evidenciar a evolução do jogador nas fases. Quando finalizaram, preencheram um formulário em *Google Forms* para o envio das estruturas criadas e avaliação sobre a percepção de modelar dados e elaborar as estruturas (Apêndice F).

Preenchimento de formulário: no passo seguinte, um segundo formulário foi compartilhado com os estudantes para que incluíssem os atributos de seus jogos educacionais (com base nos elementos descritos no Apêndice D). Essa etapa foi necessária para armazenar os dados dos jogos que seriam enviados ao ChatGPT para geração das estruturas de GLA. Para facilitar o acesso, uma planilha foi gerada a partir do formulário e compartilhada com os estudantes.

Geração das estruturas via LLM: o último passo da pesquisa envolveu o uso do ChatGPT para gerar as estruturas GLA. Essa etapa dividiu-se em dois momentos: (i) geração pelo especialista (Abordagem 2): o especialista em GLA, utilizando o *prompt* elaborado com a técnica de *basic prompting* e os dados dos jogos fornecidos pelos estudantes, utilizou o LLM para gerar as estruturas correspondentes; e (ii) geração pelos estudantes (Abordagem 3): os estudantes criaram um *prompt* para orientar o ChatGPT a receber os dados dos jogos, fazer questionamentos sobre o jogo educacional para entendê-lo melhor e gerar a estrutura de GLA. Essa técnica interativa com questionamentos foi inspirada na técnica de *chain-of-throught prompting*, utilizando uma etapa intermediária para o LLM "raciocionar". Os estudantes foram orientados a interagir com o ChatGPT em seus computadores, devendo enviar o *prompt* junto com os dados dos jogos e responder aos questionamentos do modelo. Após a geração das estruturas, os estudantes preencheram um terceiro formulário com os *links* de interação com o ChatGPT e as percepções sobre a atividade (criação do *prompt*, respostas do modelo,

comparação com a modelagem manual), descrito na Seção 3.3.3.

3.3.3 Coleta e análise de dados

Na coleta de dados, os estudantes utilizaram um terceiro formulário para enviar as estruturas geradas e fornecer *feedback* das atividades (Apêndice G). Esse formulário foi elaborado de forma autoral pela especificidade da pesquisa, contendo questões qualitativas e quantitativas a respeito da experiência de uso e da dificuldade ao elaborar *prompts*, do apoio e entendimento fornecido pelo ChatGPT, da orientação a novos usuários e das percepções sobre adequação, limitações e possibilidades de melhoria das estruturas geradas. Além disso, inclui uma questão para autorizar a disponibilização dos dados de forma anônima, com finalidade única para pesquisas.

No total, 24 estruturas de GLA foram geradas: oito pelos estudantes de forma manual, oito pelo especialista com técnica de *basic prompting* no ChatGPT e oito pelos estudantes com técnica interativa de questionamentos no ChatGPT. As estruturas foram baseadas em oito jogos distintos, visto que quatro estudantes optaram por formar duplas. Em relação às avaliações dos especialistas, utilizou-se formulários com o instrumento "*Player Level Up!*" (Tabela 4), que contém sete questões quantitativas em escala Likert-5 e uma questão aberta opcional para cada estrutura.

Para as análises quantitativas, fez-se uso de um *boxplot* para comparar as avaliações dos especialistas para as três abordagens de geração das estruturas de captura e um gráfico de linhas para ilustrar a distribuição de notas por questão. No *boxplot*, o eixo X são as abordagens, onde cada barra representa a distribuição das notas dos especialistas para as estruturas correspondentes. Já o eixo Y refere-se aos limites das notas, que varia de 7 (todas as questões foram marcadas com a menor nota) a 35 (todas as questões atribuídas com nota máxima). No gráfico de linhas, o eixo X é a numeração das questões (1 a 7) e o eixo Y são as notas de avaliação, que variam de 24 (os três especialistas atribuíram a nota mínima para as 8 estruturas da abordagem respectiva) a 120 (nota máxima dos três especialistas). Quanto às questões qualitativas, referentes às percepções dos estudantes, utilizou-se análise temática: (i) uma leitura preliminar

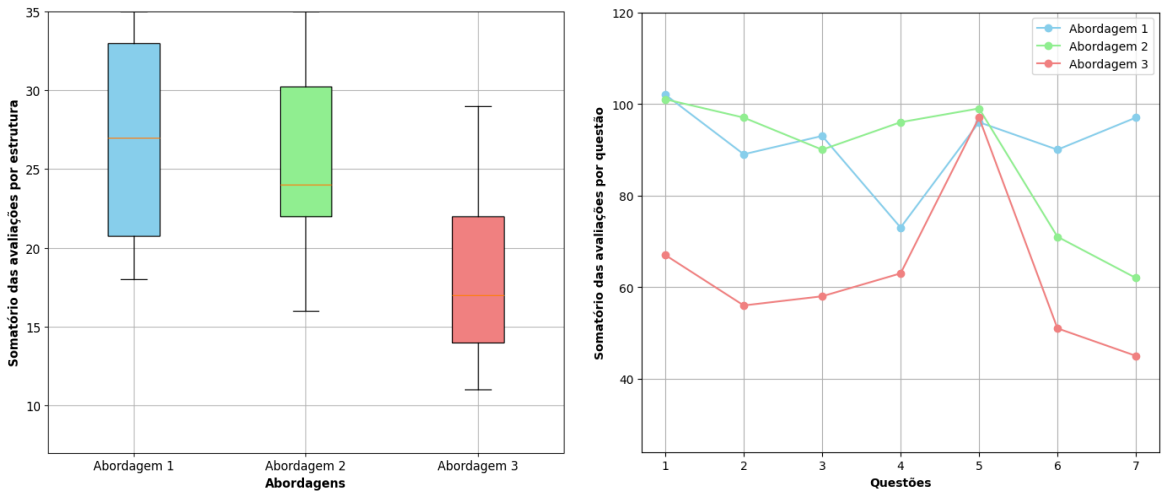
das respostas foi feita; (ii) categorias temáticas foram criadas para agrupá-las; e (iii) as respostas foram atribuídas conforme as categorias para obter *insights*.

3.3.4 Resultados

A Figura 12 apresenta gráficos que ilustram as avaliações dos especialistas para as estruturas de GLA conforme o instrumento "*Player Level Up!*". No gráfico *boxplot* da Figura 12(a), cujas avaliações são organizadas por abordagem, pode-se observar que a Abordagem 1, onde os estudantes geraram as estruturas de captura de forma manual (sem LLM), foi a melhor avaliada pelos especialistas. A Abordagem 2, em que o especialista gerou as estruturas com *basic prompting* via ChatGPT, foi ligeiramente inferior a Abordagem 1, mas apresentou a menor distribuição dos dados e uma mediana similar. Já a Abordagem 3, onde os estudantes utilizaram um *prompt* orientado a questionamentos ao ChatGPT para gerar as estruturas, foi a pior avaliada: mediana menor que o valor intermediário da escala (21) e, em maioria, notas abaixo de 25.

Em relação às questões, a Figura 12(b) apresenta o somatório das avaliações por questão em cada abordagem. Observa-se que as Abordagens 1 e 2 apresentaram avaliações semelhantes na maioria das questões, sugerindo proximidade entre as estruturas geradas de forma manual e as estruturas geradas por meio de um LLM com intervenção de um especialista. Esse resultado é similar ao observado na investigação da Seção 3.2. Por outro lado, exceto na Q5 (ausência de variáveis duplicadas ou sobrepostas), a Abordagem 3 obteve avaliações inferiores às demais, com somatórios abaixo do valor intermediário da escala (72). Além disso, a Q6 (completude da estrutura quanto às variáveis relevantes) e a Q7 (presença de variáveis relacionadas à análise temporal) obtiveram avaliações acima desse valor na Abordagem 1, mas abaixo dele nas demais abordagens. Esse resultado sugere que, na percepção dos especialistas, as estruturas geradas por LLM, sobretudo na Abordagem 3, apresentaram limitações em contemplar variáveis para rastrear o percurso do jogador e analisar sua evolução durante a *gameplay*.

Quando os estudantes foram questionados sobre o nível de dificuldade em modelar os dados para criar as estruturas, tanto de modo manual (questão 10 - Apêndice



(a) Boxplot das Abordagens. (b) Avaliações das questões em cada Caso.

Figura 12 – Resultados das avaliações realizadas pelos especialistas.

F) quanto na criação do *prompt* para o ChatGPT (questão 6 - Apêndice G), notou-se que as respostas foram semelhantes, indicando uma dificuldade intermediária, embora variável entre os participantes. Dentre alguns desafios apontados na elaboração do *prompt* para gerar as estruturas de GLA(questão 7 - Apêndice G) , destacam-se: estruturar os dados no *prompt* (n=5, 50%), limitações do ChatGPT (n=2, 20%) e responder aos questionamentos do LLM (n=1, 10%). No que diz respeito aos dados, os estudantes apontaram que o ChatGPT auxiliou a entender quais eram importantes (n=5, questão 8 - Apêndice G) e o "porquê" (n=6, 60%, questão 11 - Apêndice G).

De modo geral, nenhuma abordagem apresentou notas inferiores a 10 (Figura 12(a)), sugerindo que as estruturas, na percepção de especialistas em GLA, não foram consideradas totalmente inadequadas para apoiar o rastreamento do percurso do jogador nos jogos respectivos. A Abordagem 1, em que os estudantes geraram as estruturas de GLA de forma manual, foi a mais bem avaliada por especialistas na área. Entretanto, isso não indica que a utilização de um LLM seja menos adequada, uma vez que a Abordagem 2 obteve resultados menos dispersos e avaliações próximas às da Abordagem 1. De modo semelhante, as notas mais baixas da Abordagem 3 não implicam que a técnica interativa de questionamentos prejudique os resultados do modelo, mas sugerem que a elaboração de *prompts* pelos estudantes pode ter sido desafiadora para aplicá-la. Esse desafio pode estar relacionado ao fato de que os estudantes não tiveram aulas prévias

sobre elaboração de *prompts*, ainda que o ChatGPT fosse uma ferramenta familiar.

3.3.5 Considerações sobre a investigação

Em resposta à QP3.3, os questionamentos e a quantidade de dados de jogos no *prompt* influenciam na qualidade das estruturas de captura de GLA de formas distintas:

- **Questionamentos:** essa investigação traz indícios de que os questionamentos têm potencial para gerar estruturas de captura mais adequadas, conforme concordância de oito estudantes (80%, questão 10 - Apêndice G). Por outro lado, os estudantes apresentaram dificuldades em elaborar *prompts*: 40% (n=4) avaliaram a dificuldade como "mediana" e 20% (n=2) como "difícil" ou "muito difícil" (questão 6 - Apêndice G), principalmente para estruturar a geração das estruturas pelo modelo (questão 7 - Apêndice G). Essas dificuldades, no entanto, podem não ter sido plenamente percebidas pelos participantes, visto que 40% (n=4) informou não haver necessidade de melhoria nas estruturas geradas pelo ChatGPT (questão 15 - Apêndice G) e 50% (n=5) relatou que não havia elementos importantes faltantes para coletar (questão 13 - Apêndice G). Essa percepção contrasta com a avaliação dos especialistas, pois a Abordagem 3 apresentou os piores resultados nas questões Q6 e Q7, indicando que variáveis importantes para a identificação de indícios de aprendizagem não foram contempladas nas estruturas.

- **Inclusão de mais dados de jogos no *prompt*:** pode ter atrapalhado o modelo, que desconsiderou alguns deles durante a geração das estruturas. Esse ponto está em conformidade com a pesquisa de Liu et al. (2024), que apontou que modelos tendem a ignorar informações do *prompt*, sobretudo no meio da entrada. Além disso, a quantidade excessiva de informações poderia ter causado alucinação ao modelo, gerando resultados incoerentes e imprecisos (IBM, 2023a). Tais aspectos reforçam a necessidade da elaboração de um *prompt* adequado ou uma estratégia que não necessite de *prompts* complexos e atinja os resultados desejados.

Dentre as limitações da investigação, destacam-se: (i) o tempo reduzido das

aulas, que não permitiu a aplicação de aulas de engenharia de *prompt* aos estudantes; (ii) o uso do ChatGPT 3.5 que, embora tenha sido escolhido por ser gratuito, apresentava limitações em relação às versões 4 e 4o disponíveis à época; (iii) o tamanho reduzido da amostra, decorrente também da exclusão de participantes que não participaram de todas as etapas ou que não preencheram todos os formulários; (iv) a análise concomitante dos questionamentos e das informações adicionais dos jogos no *prompt*, cujos efeitos não foram isolados e podem ter prejudicado as inferências; e (v) a formação de duplas pelos estudantes, que pode ter influenciado a forma como os *prompts* foram elaborados e dificultado a identificação das dificuldades individuais no processo.

3.4 Síntese do Capítulo

Este capítulo apresentou investigações sobre o problema de pesquisa (desafios em modelagem de dados) e de que forma os LLMs, uma tecnologia emergente, podem mitigá-lo. Diante disso, na primeira investigação, estudantes de Computação que atuavam como *designers* de aprendizagem modelaram dados para seus jogos educacionais. As principais dificuldades identificadas incluem: (i) o nível de abstração necessário para modelar dados; (ii) a necessidade de compreender, com clareza, o jogo e seus objetivos de aprendizagem para definir os dados que serão coletados; (iii) o tempo utilizado em sala de aula não foi suficiente para modelar dados; e (iv) alguns conhecimentos prévios de computação são fundamentais para a modelagem, mas nem todos os estudantes já haviam cursado as disciplinas correspondentes. Esses pontos sugerem que a modelagem de dados não é considerada uma etapa trivial.

Já na segunda investigação, analisou-se se os LLMs podem contribuir para o campo de GLA. Uma das principais atividades de GLA é elaborar estruturas de captura (*templates* de dados), que é diretamente dependente da modelagem de dados. Portanto, dado um jogo educacional, um especialista em GLA concebeu uma estrutura de captura de forma manual, e utilizou três versões distintas do ChatGPT para gerá-la artificialmente. As quatro estruturas foram comparadas e a estrutura manual foi a mais bem avaliada por outros especialistas na área. Entretanto, a estrutura gerada pelo

ChatGPT 4o obteve avaliações semelhantes, mesmo sem utilizar exemplos no *prompt* e com conhecimentos limitados sobre o jogo educacional. Essa pesquisa revelou indícios de que os LLMs podem contribuir para GLA, tanto na definição dos dados quanto na criação e organização das variáveis de GLA em uma estrutura de captura.

Dando continuidade à pesquisa anterior, realizou-se uma investigação para verificar se as estruturas de captura de GLA geradas por LLM podem ser aprimoradas ao incluir mais informações do jogo no *prompt* e ao empregar uma técnica interativa de questionamentos. Para isso, estruturas de GLA foram comparadas: (i) oito geradas manualmente por estudantes; (ii) oito geradas artificialmente com um LLM por um especialista em GLA; e (iii) oito geradas artificialmente com um LLM, utilizando uma técnica interativa de questionamentos, por estudantes. Especialistas em GLA avaliaram as 24 estruturas e, de forma geral, as manuais foram as mais bem avaliadas, as do especialista obtiveram avaliações semelhantes e as com a técnica interativa foram as piores avaliadas. Analisando as percepções dos estudantes, observou-se que tiveram dificuldade na projeção dos *prompts* aos modelos, mas que consideraram que a técnica de questionamentos auxilia na modelagem. Os resultados também sugerem que há potencial dos LLMs para apoiar a modelagem de dados para jogos educacionais.

4

CONSTRUINDO E AVALIANDO UM AGENTE ESPECIALISTA EM GLA

As contribuições das investigações do capítulo anterior ressaltam que: (i) modelar dados para jogos educacionais é considerada como uma atividade complexa por estudantes que atuam como *designers* de aprendizagem; (ii) existe potencial no uso de LLMs para mitigar essa complexidade, sobretudo na interpretação do jogo educacional para a proposta de variáveis de GLA correspondentes; e (iii) técnicas de questionamentos podem orientar o *designer* de aprendizagem a refletir sobre seu jogo e a definir os dados de coleta, mas projetar *prompts* para LLMs que os conduzam a gerar estruturas de GLA pode não ser trivial. Diante disso, este capítulo contempla a proposta e desenvolvimento de uma solução (prova de conceito - POC) para dar suporte à modelagem de dados: um agente de conversação no ChatGPT. Além disso, detalha a avaliação do agente por *designers* de aprendizagem, sendo especialistas na área (Seção 4.1) e estudantes que atuam nessa função em disciplinas (Seção 4.2). O capítulo está relacionado ao terceiro objetivo específico da dissertação[1] e corresponde ao Ciclo de *Design* da DSR.

[1] Desenvolver uma proposta (prova de conceito - POC) que dê suporte a *designers* de aprendizagem na modelagem de dados.

4.1 Construção e Avaliação de Agente Inteligente em GLA

Diante das dificuldades em projetar *prompts* para GLA, uma estratégia que não requer *prompts* complexos pode ser promissora. Essa é a motivação desta investigação, cuja questão de pesquisa é: "QP4.1 - Como construir e avaliar um agente especialista baseado em LLM que auxilie na modelagem de dados?". Esta investigação encontra-se publicada nos anais do SBIE - CBIE (2024), disponível no link[2].

4.1.1 Objetivo, contexto e participantes

Esta investigação propõe e avalia um agente de conversação, em formato de POC, para auxiliar na modelagem de dados. Optou-se por alinhá-lo ao GLBoard (Seção 2.4), um modelo para captura e análise de dados de jogos educacionais, que disponibiliza um *template* de dados genérico. Essa escolha deu-se: (i) pelo modelo pertencer ao grupo de pesquisa do mestrando; (ii) pela experiência com sua utilização; (iii) pela facilidade de uso, sobretudo em relação à personalização das estruturas de GLA; e (iv) pela disponibilidade de materiais complementares gratuitos e em língua portuguesa.

A motivação da pesquisa surgiu após a introdução da funcionalidade "MyGPTs" pela OpenAI, que permite a criação de versões customizáveis (agentes) do ChatGPT (OpenAI, 2023). O ambiente proporciona a configuração do comportamento do agente por meio da definição de instruções e da atualização de sua auto-descrição, bem como a incorporação de uma base de conhecimento composta por arquivos (textos, documentos, imagens, entre outros). A partir desses materiais, o LLM utiliza *Retrieval-Augmented Generation* (RAG) para recuperar trechos da base de conhecimento semanticamente relacionados à consulta do usuário e gerar respostas fundamentadas nessas informações. Desse modo, o usuário é capaz de criar um agente de conversação sem a necessidade de conhecimentos avançados em *prompt* ou programação. A restrição do agente ao seu escopo (instruções e base de conhecimento) permite a geração de respostas orientadas ao domínio específico, levando-o a se comportar como um especialista.

A investigação descreve o processo de construção do agente, bem como sua

[2] https://sol.sbc.org.br/index.php/sbie/article/view/31447

avaliação. Nesse sentido, os participantes da pesquisa são os três especialistas em GLA que conceberam o agente (mestrando e orientadoras) e outros sete especialistas em GLA convidados que realizaram sua avaliação. Eles foram selecionados por conveniência, com as seguintes características: (i) formação: 43% (n=3) com graduação em computação em andamento e 57% (n=4) com pós-graduação (mestrado e doutorado) iniciando ou em andamento; (ii) experiência em desenvolver jogos: 86% (n=6) possui dois ou anos ou mais de expertise, enquanto 14% (n=1) com um ano; (iii) experiência em GLA: 57% possui menos de um ano e 14%, um ano; (iv) quantidade de jogos em que implementou GLA: 71% (n=5) para um jogo e 29% (n=2) para dois jogos ou nenhum. Todos consentiram em disponibilizar os dados avaliativos de forma anônima, com finalidades únicas para pesquisas científicas.

4.1.2 Procedimentos

As etapas conduzidas para construir o agente estão contidas na Figura 13.

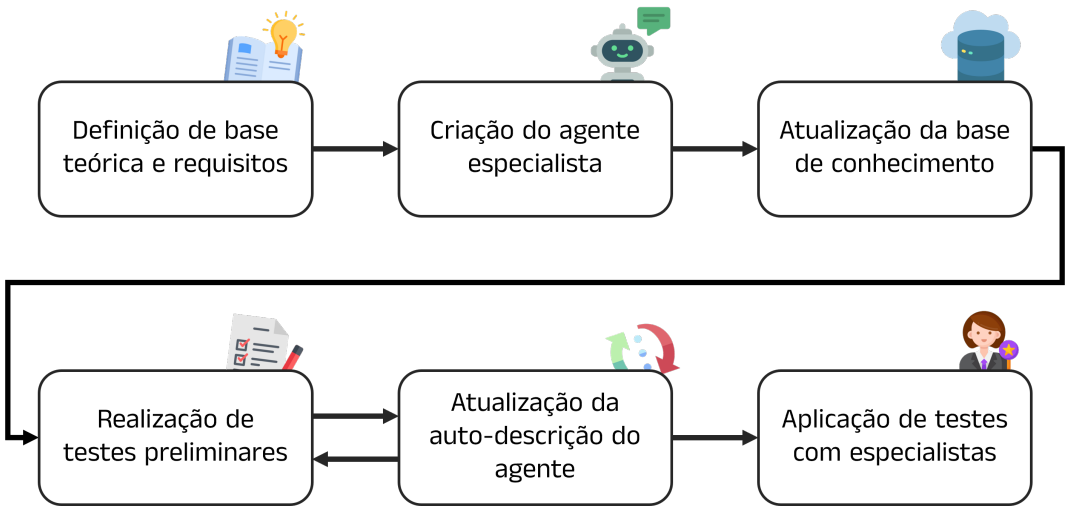


Figura 13 – Etapas da investigação sobre o agente Inteligente em GLA.

Definição de base teórica e requisitos: a primeira etapa consistiu no levantamento da base teórica para construir o agente, em que se analisou trabalhos da literatura que propõem GPTs customizados para diversos contextos (Almasre, 2024; Kıyak; Kononowicz, 2024; Sathe et al., 2024). Apesar de não haver nenhum trabalho sobre agentes em GLA, notou-se que a criação de um *chatbot* customizável traz benefícios como:

evita a generalização do ChatGPT convencional, possibilita a inclusão de uma base de conhecimento específico, não requer *prompts* complexos para interação, não necessita de codificação para personalizá-lo, entre outros. Já os requisitos para a modelagem de dados e implementação das técnicas de GLA são baseados no GLBoard, considerando seu *template* de dados representado em JSON para estruturas de captura de dados, podendo ser customizável para quaisquer jogos educacionais.

Criação do agente especialista: trata-se da construção do agente pelos especialistas em GLA (mestrando e orientadoras), que iniciou com o acesso ao "MyGPTs" do ChatGPT, condicionado à assinatura da versão Plus. Em seguida, o agente foi configurado na área "GPTBuilder", que possui duas abas para definição das propriedades do agente: "Criar", através de interação com *chatbot*, e "Configurar", com campos para preenchimento (nome, descrição, instruções, quebra-gelos[3], conhecimento, capacidades e ações). Utilizando ambas as opções, as propriedades do agente foram definidas, considerando que ele: (i) auxilie na modelagem de dados e implementação das técnicas de GLA; (ii) adote um comportamento semelhante ao de um especialista humano; e (iii) tenha capacidade orientar o *designer* de aprendizagem, estimulando-o a refletir sobre seu jogo por meio de questionamentos. O agente foi denominado "*GLA Specialist*".

Atualização da base de conhecimento: refere-se ao mapeamento de trabalhos para incluir na base de dados do agente. A ferramenta "Google Acadêmico" foi utilizada para localizar e selecionar trabalhos com relevância sobre GLA, cujo critério era estar indexado nas primeiras páginas da pesquisa. Também incluiu-se o artigo do GLBoard, sua documentação e seu *template* de dados (JSON). Esses materiais foram baixados e incorporados ao "Conhecimento" do agente, em formatos TXT, PDF e PNG.

Realização de testes preliminares: diz respeito aos testes exploratórios conduzidos pelo mestrando na área de "Pré-visualizar", onde se pode interagir com o *chatbot* antes de publicá-lo. O objetivo foi analisar as respostas do agente para identificar possíveis inconsistências, cujos testes incluíram: questionamentos sobre GLA e GLBoard, pedidos de auxílio na implementação das técnicas, solicitações para preencher o *template* de dados (estruturas) conforme os dados dos jogos informados, entre outros. Nesse momento, percebeu-se ressalvas quanto a não conformidade de estruturas geradas com

[3] Mensagens diretas clicáveis para auxiliar o início de uma conversa.

o padrão do GLBoard (formato, lógica), etapas não realizadas e outros pontos.

Atualização da auto-descrição do agente: as inconsistências identificadas foram corrigidas no agente a partir da atualização de sua auto-descrição. Dentre os ajustes realizados, destacam-se: (i) inclusão de etapa de confirmação ao usuário para questionar sobre o jogo educacional; (ii) foco na personalização do *template* de dados; (iii) inserção dos nomes dos criadores do agente e um breve resumo histórico para contextualização; e (iv) descrição com mais detalhes sobre a variável $path\_player$, que armazena dados do percurso do jogador. Diante disso, retornava-se à etapa anterior para a realização de novos testes e atualizava-se a auto-descrição quando necessário, mantendo um processo cíclico entre as etapas. A auto-descrição atual do agente está disponível no Apêndice I.

Aplicação de testes com especialistas: após iterações de melhorias do agente, determinou-se que suas respostas estavam atendiam aos critérios esperados: conformidade com o *template* do GLBoard, variáveis adequadas com o contexto do jogo informado, presença de etapa de confirmação sobre os questionamentos, entre outros. Em seguida, especialistas em GLA foram convidados a realizar testes com o agente, cujo *link* de acesso foi disponibilizado por meio da opção de "Compartilhar". Os testes ocorreram de modo *on-line* e assíncrono, visando maior flexibilidade de horários, onde os especialistas foram orientados a interagir com o agente (perguntar sobre conceitos de GLA, modelar dados, elaborar estruturas de GLA, entre outros) e preencher um formulário avaliativo (descrito na subseção seguinte). Os especialistas utilizaram um roteiro para guiar o teste (Apêndice J), mas não se limitando a ele.

4.1.3 Coleta e análise de dados

A coleta dos dados ocorreu por meio de um formulário avaliativo (*Google Forms*) com os especialistas em GLA, composto por três seções: (i) perfil do testador: questões quantitativas sobre o tempo de experiência com jogos educacionais e GLA, bem como a quantidade de jogos em que o especialista implementou GLA; (ii) questões fechadas (quantitativas): em escala Likert-5, visando avaliar: a precisão e a relevância das respostas sobre GLA e GLBoard, a clareza na explicação das variáveis pertencentes

ao $path\_player$, a adequação dos exemplos e o potencial do agente para auxiliar na modelagem de dados; e (iii) questões abertas (qualitativas): sobre pontos positivos e negativos, limitações e sugestões de melhoria. As questões foram elaboradas de forma autoral pela especificidade da pesquisa, visando verificar se o agente contribui para o campo de GLA. O formulário completo está presente no Apêndice H.

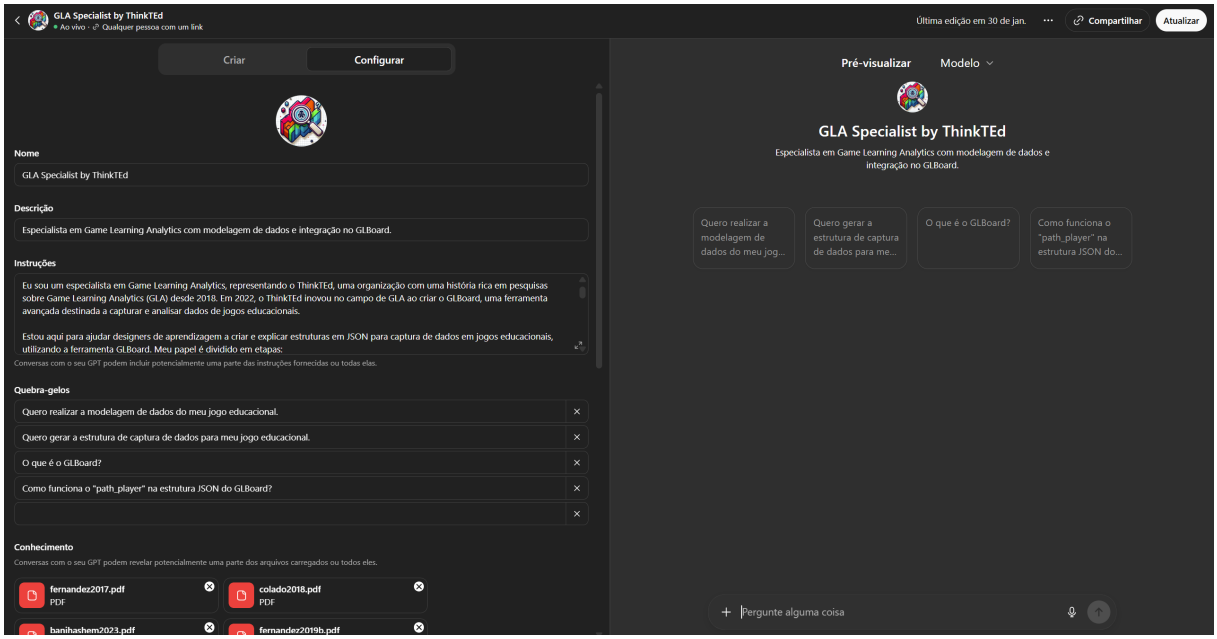
Para a análise de dados, utilizou-se: (i) tabela com a a distribuição de notas por questão (Tabela 6) e histogramas individuais das avaliações para as questões quantitativas; e (ii) análise temática para as questões qualitativas: uma leitura preliminar das respostas foi realizada, onde se criaram categorias. Em seguida, as respostas foram organizadas conforme a categoria e analisadas em conjunto, visando a identificação de padrões e interpretação dos resultados conforme a categoria.

4.1.4 Resultados

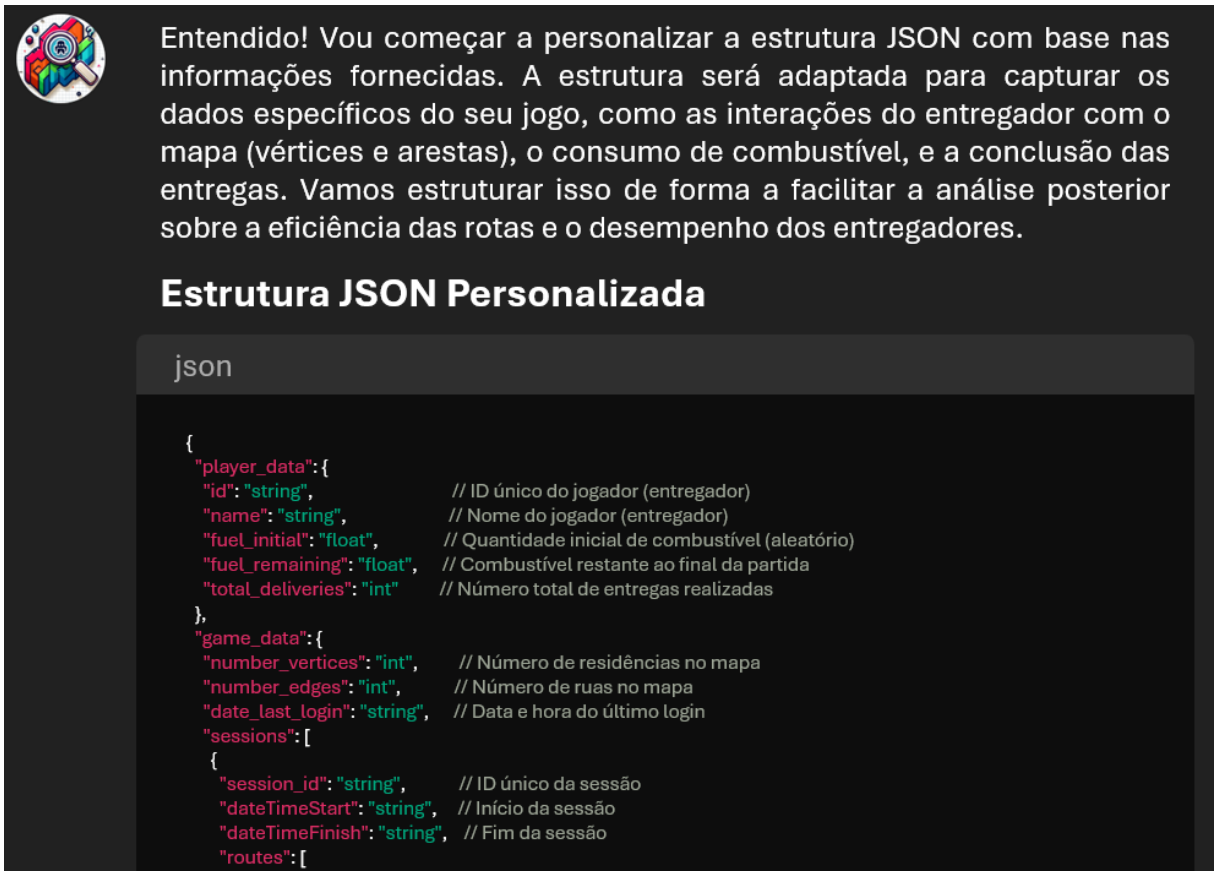
Como resultado da investigação, obteve-se a versão preliminar do artefato "*GLA Specialist*": um agente inteligente especialista criado no ChatGPT para auxiliar em GLA, com ênfase no modelo GLBoard. O agente realiza as seguintes etapas: (i) auxilia na compreensão de GLA e GLBoard; (ii) recebe e analisa dados de jogos educacionais para gerar as estruturas de captura; (iii) aprimora a estrutura a partir de questionamentos ao usuário, caso permitido; (iv) personaliza a estrutura (variáveis e tipos de dados) a partir dos dados dos jogos e das respostas do usuário aos questionamentos; (v) insere comentários nas estruturas para auxiliar na modelagem de dados; (vi) cria um exemplo da estrutura JSON com dados genéricos para ilustrar a coleta; (vii) faz a validação da estrutura com o usuário, que pode solicitar modificações; e (viii) foca-se em propor variáveis para substituir o $path\_player$ e armazenar informações do percurso do jogador.

O agente está disponível publicamente sob o link[4] e é ilustrado na Figura 14, contendo sua página de edição (Figura 14(a)) e um exemplo de geração de estrutura de GLA (Figura 14(b)). A interação completa está disponível no link[5]

[4] https://chatgpt.com/g/g-musWTetpp-gla-specialist-by-thinkted
[5] https://chatgpt.com/share/678d78cc-5670-8013-ab29-76ffa1014f0a



(a) Página de edição do agente.



(b) Interação com o agente especialista.

Figura 14 – Agente "*GLA Specialist*".

Em relação à avaliação do agente pelos especialistas em GLA, a Tabela 6 apresenta as questões quantitativas, que correspondem, respectivamente, às questões 6 a

12 do Apêndice H, incluindo a distribuição das notas, as somatórias e as medianas. Nota-se uma convergência dos especialistas, que atribuíram somente notas 4 (concordo) ou 5 (concordo completamente), apresentando somatória[6] superior a 29. Esse resultado pode expressar o potencial do agente em experiências que envolvem GLA e GLBoard.

Tabela 6 – Distribuição das Notas por Questão.

Questão	Descrição	Nota 4	Nota 5	Somatória	Mediana
Q1	O agente forneceu respostas precisas e relevantes às suas perguntas sobre GLA?	29%	71%	33	5.0
Q2	O agente forneceu respostas precisas e relevantes às suas perguntas sobre GLBoard?	43%	57%	32	5.0
Q3	O agente foi capaz de personalizar a estrutura JSON de acordo com as informações fornecidas sobre o seu jogo educacional?	57%	43%	31	4.0
Q4	O agente explicou adequadamente as variáveis propostas no $path\_player$?	29%	71%	33	5.0
Q5	Na sua percepção, os exemplos que o agente gerou eram adequados para capturar informações sobre dados e implementação técnicas de GLA?	71%	29%	30	4.0
Q6	Você acha que o agente tem potencial para auxiliar na modelagem de dados e implementação técnicas de GLA?	29%	71%	33	5.0
Q7	Como você classificaria sua experiência geral ao interagir com o "*GLA Specialist*"?	86%	14%	34	5.0

A Tabela 7 apresenta as perguntas abertas (qualitativas), correspondentes às questões 13 a 17 do Apêndice H, utilizadas para coletar o *feedback* dos avaliadores sobre o agente. A coluna "Respostas" contém uma síntese das opiniões dos especialistas.

Tabela 7 – Questões e Respostas sobre o Agente.

Questão	Título da Pergunta	Respostas
Q1	O que você mais gostou do agente?	As respostas precisas e específicas para entender o jogo (n=4, 57%), a facilidade na criação das estruturas somente com informações do jogo (n=1, 14%), a geração do JSON (n=4) e a sugestão de gráficos para interpretá-lo (n=1).
Q2	O que você não gostou do agente?	Enquanto uns indicaram que "nada" (n=2, 29%), outros pontuaram: respostas longas (n=2), falta de exemplos de aplicação e tecnologias existentes (n=1), o tamanho do JSON (n=1) e a sugestão de variáveis cuja captura é complexa (n=1).
Q3	Quais melhorias você sugeriria para aprimorar o "*GLA Specialist*"?	Alguns testadores não apontaram sugestões (n=3, 43%), outros sugeriram incluir exemplos com imagens (n=1), que o modelo solicitasse a documentação do jogo para fornecer respostas mais precisas (n=1), mais exemplos em JSON (n=1) e respostas não tão extensas (n=1).
Q4	Houve algo que o agente não conseguiu fazer ou responder conforme suas expectativas? Se sim, por favor, explique.	A maioria disse que não (n=5, 71%), com exceção de dois avaliadores que: (i) ficou curioso em implementar a estrutura no jogo, coletar dados de jogadores, enviá-los ao modelo e analisar como o agente iria melhorar a estrutura; (ii) informou que o agente não compreendeu por completo o jogo para gerar a estrutura, mas que auxiliou nos primeiros passos para modelar dados.
Q5	Você acha que o agente é uma boa opção para auxiliar estudantes a entender/implementar GLA? Por quê?	Os avaliadores concordaram de forma unânime que sim (n=7, 100%), pontuando benefícios como: precisão e especificidade sobre o assunto nas respostas (n=2), auxilio no processo de modelagem de dados (n=1), apoio introdutório à GLA (n=3), otimização de tempo e exemplos no *template* em JSON (n=1).

[6] Valor agregado das notas de todos os especialistas para uma questão, variando de 7 (todos os especialistas atribuíram a nota 1) a 35 (nota máxima atribuída pelos especialistas).

4.1.5 Considerações sobre a investigação

Respondendo à QP4.1, um agente baseado em LLM para realizar atividades de GLA, como a modelagem de dados, pode ser criado com a função "MyGPTs" do ChatGPT, utilizando materiais de referência da área como base de conhecimento, definindo uma auto-descrição detalhada e realizando testes unitários recorrentes para garantir a precisão das respostas. Quanto à avaliação, pode ser feita por especialistas de GLA, por meio de testes que envolvam: conversar com o agente sobre conceitos, pedir para que uma modelagem de dados de um jogo seja realizada, solicitar para preencher o *template* de dados com as variáveis de GLA, entre outros. De forma geral, obteve-se uma boa avaliação do agente pelos especialistas nas questões quantitativas e qualitativas, cujas melhorias e ressalvas serão incorporadas à nova versão. Esses resultados sugerem o potencial do agente em contribuir para modelagem de dados em GLA, fornecendo um ambiente simples de conversação que não requer o projeto de *prompts* complexos e que possui uma base de conhecimento específica que pode ser atualizada.

Limitações da pesquisa incluem: (i) possível viés, visto que os especialistas pertencem ao mesmo grupo de pesquisa, o que pode ter influenciado suas percepções e favorecido avaliações mais positivas. Ainda assim, apontaram sugestões para aprimorar o agente; (ii) perspectiva somente dos especialistas e não de estudantes, que foi realizada posteriormente (Seção 4.2); (iii) testes com foco nas respostas, sem comparação estatística com outras estratégias (versões do ChatGPT, outros LLMs); (iv) foco na geração de estruturas de captura no *template* do GLBoard, sem contemplar outras ferramentas como o xAPI-SG (Serrano-Laguna et al., 2017). O objetivo era demonstrar os primeiros passos na construção de um agente especialista em GLA, sendo o primeiro da literatura sobre esse aspecto, cujas limitações busca-se mitigar em pesquisas posteriores.

4.2 Avaliação do Agente Inteligente em GLA por estudantes de computação

Após condução da investigação anterior, obteve-se o "*GLA Specialist*", um agente inteligente para auxiliar na modelagem de dados em GLA. Avaliações com especialistas na área demonstraram potencial da ferramenta, que não requer a elaboração de *prompts* complexos. Entretanto, o agente ainda não foi avaliado por outra parte do público-alvo: estudantes de computação que atuam como *designers* de aprendizagem. Nesse aspecto, a seguinte questão de pesquisa foi definida: "QP4.2 - Quais as percepções desses estudantes ao utilizar o "*GLA Specialist*" e ao compará-lo com a técnica de modelagem manual?". Um estudo de caso foi conduzido para respondê-la. Esta investigação resultou em uma publicação, sendo aprovada para o SBIE - CBIE (2025), disponível neste link[7].

Para esta pesquisa, buscou-se analisar: (i) as experiências dos estudantes ao modelar dados com o agente especialista; e (ii) suas percepções ao comparar a utilização do agente com a técnica de modelagem manual. Diante disso, o único critério de seleção definido foi ter participado da investigação presente no Capítulo 3, Seção 3.1, onde os estudantes modelaram dados para seus jogos educacionais de modo manual. Portanto, nove pessoas atenderam ao critério e participaram da pesquisa, das quais seis (66%) se identificaram com o gênero masculino e três (34%) com o feminino, com faixa etária de 19 a 24 anos e 100% fazia parte do curso de Licenciatura em Computação da UEA.

4.2.1 Procedimentos

A condução desta investigação está organizada em quatro etapas principais, como ilustrada na Figura 15 e descrita a seguir.

Explicação da atividade: visto que a pesquisa foi realizada em sala de aula, sua condução deu-se por meio de uma atividade com os estudantes. Desse modo, essa etapa incluiu um resumo da aula anterior (investigação da Seção 3.1, referente à modelagem de dados manual), os requisitos da atividade e observações adicionais. Explicou-se aos estudantes que deveriam modelar dados para seus jogos utilizando o

[7] https://sol.sbc.org.br/index.php/sbie/article/view/38508

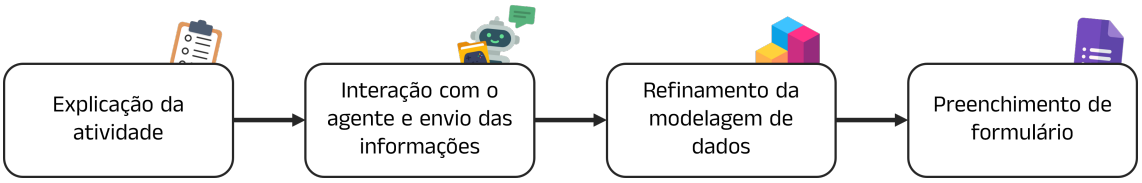


Figura 15 – Etapas da investigação sobre aplicação do Agente Inteligente.

agente inteligente "*GLA Specialist*" e preencher um formulário de avaliação ao final da atividade. As orientações incluíram: (i) modelar dados com o agente até julgarem os resultados adequados; (ii) interagir livremente com o agente sem regras específicas; (iii) utilizar um ou mais *chats*, conforme a necessidade, guardando os *links* das interações; e (iv) preencher o formulário ao final da atividade.

Interação com o agente e envio das informações: todos os estudantes acessaram o agente por meio de uma conta específica do ChatGPT e e do mesmo *link* de acesso, iniciando suas interações com o *chatbot* de modo individual. Com isso, informaram no *prompt* que gostariam de realizar a modelagem de dados dos seus jogos, cujo agente, logo em seguida, solicitou as características desses jogos, exemplificando-nas em alguns casos. A partir da solicitação do agente, os estudantes iniciaram a elaboração de *prompts* simples com informações de seus jogos, tais como: nome, público-alvo, tema e conteúdo, habilidades da BNCC relacionadas, *gameplay*, história, mecânica de aprendizagem, e enviaram-nas ao agente. Ressalta-se que nem todos os participantes enviaram a mesma quantidade e tipos de características de seus jogos. Desse modo, a partir das informações, o agente elaborou uma tabela com a modelagem de dados dos jogos, contendo: nº, nome do dado, exemplo de coleta e justificativa de captura.

Refinamento da modelagem de dados: após a elaboração da modelagem de dados dos jogos, os estudantes prosseguiram utilizando o agente para refiná-las. Dentre os refinamentos, destacam-se: explicação de mecânicas de outras fases, solicitação de explicação sobre variáveis definidas, indicação de melhorias e inconsistências em alguns dados, inclusão, exclusão ou modificação de alguns campos e adição de mais detalhes sobre aspectos específicos. Considerando a etapa anterior e a atual, as interações com o agente estenderam-se por aproximadamente 1 hora e 30 minutos.

Preenchimento de formulário: ao passo em que finalizavam suas modelagens, os

estudantes preencheram um formulário em *Google Forms*. Esse formulário foi elaborado para registrar as interações com o agente inteligente, a experiência ao utilizá-lo e a percepção ao compará-lo com a técnica de modelagem de dados manual, criado de forma autoral devido a especificidade da pesquisa. Assim como na investigação anterior, os estudantes informaram no formulário que consentem em disponibilizar seus dados de forma anônima, com o objetivo único de pesquisa.

4.2.2 Coleta e análise de dados

Para a coleta dos dados, o formulário descrito anteriormente foi utilizado. Ele está disponível no Apêndice K, contendo questões quantitativas e qualitativas sobre: experiência de uso e percepção de dificuldade, avaliação das estruturas e das variáveis de GLA geradas, percepção sobre a utilidade e potencial do agente, e comparação com a técnica manual de modelagem de dados.

Em relação à análise dos dados, utilizou-se: (i) gráficos *boxplot* e de linhas para as questões quantitativas (avaliações dos estudantes); e (ii) análise temática para as questões qualitativas: a partir de uma leitura inicial das respostas, foram criadas categorias, nas quais as respostas foram posteriormente organizadas. Em seguida, estas foram analisadas em conjunto, com o propósito de identificar padrões, semelhanças e diferenças, bem como interpretar os achados.

4.2.3 Resultados

O *boxplot* da Figura 16 reúne as avaliações dos estudantes quanto às questões quantitativas de avaliação do agente especialista, as quais são descritas a seguir e enumeradas conforme sua ordem no formulário (Apêndice K).

Interação geral: refere-se ao nível de interação com o agente (questão 5), em escala de "muito fácil" (nota 1) até "muito difícil" (nota 5). Mais da metade dos participantes (n=5, 56%) achou fácil/muito fácil, enquanto que 22% (n=2) considerou "médio". Em relação ao "porquê" (questão 6), os aspectos positivos apontados pelos estudantes

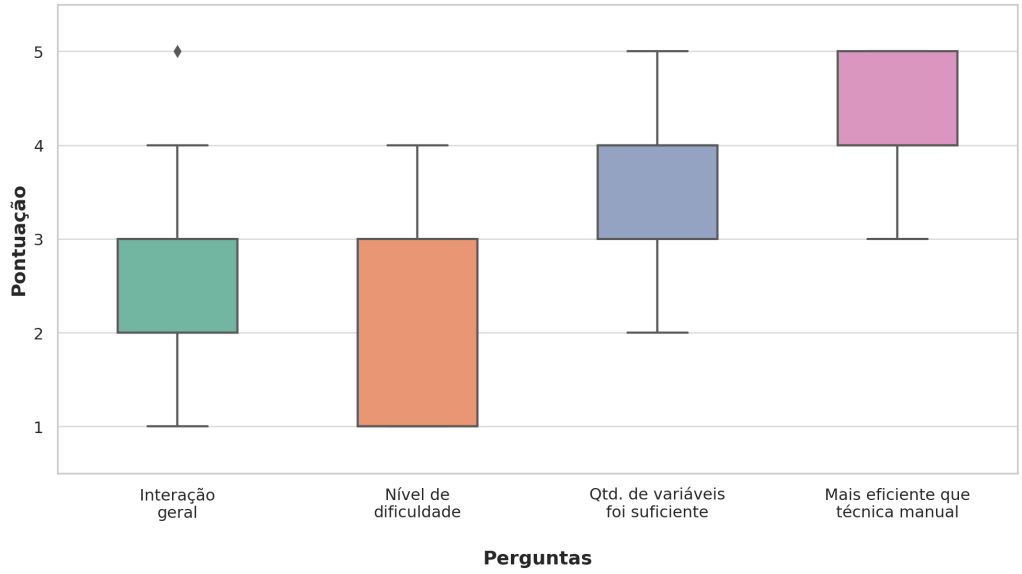


Figura 16 – *Boxplot* das questões quantitativas avaliadas.

foram: rapidez das instruções (n=1), não necessidade de um *prompt* específico para o agente (n=1), precisão (n=2), principalmente em comparação ao ChatGPT convencional, e apoio à compreensão de modelagem de dados (n=2). Os pontos negativos levantados foram: a quantidade de informações solicitadas/apresentadas (n=2), repetição de texto (n=1, 11%), informações desnecessárias (n=2) e limitação da versão do ChatGPT-Plus (n=1), onde o agente não pôde mais ser utilizado depois de atingir o limite)[8].

Nível de dificuldade: quanto a esse aspecto (questão 7), sete (78%) consideraram que a modelagem com o agente foi "fácil" e dois permaneceram neutros (22%). Quanto ao que mais acharam difícil (questão 8), a maioria (n=6, 67%) não apontou dificuldade, enquanto outros mencionaram o envio dos detalhes dos jogos (n=2) e a ausência de autonomia em comparação ao método manual (n=1). Na questão "Você acha que o agente especialista é uma boa opção para auxiliar a modelar dados de jogos educacionais?" (questão 14), houve concordância de 100% dos estudantes, assim como dos especialistas em GLA na investigação anterior (Seção 4.1).

Dentre o que acharam bom do agente (questão 15), os estudantes apontaram a precisão nas respostas (n=2, 22%), facilidade de uso (n=3, 33%), auxílio nas ideias de coleta e justificativa dos dados (n=5), e agilidade (n=2). Já sobre aspectos ruins (questão

[8] Essa situação ocorreu com dois participantes. Foi necessário a utilização de uma nova conta do ChatGPT com a versão Plus para que pudessem prosseguir com o uso do agente.

16), informaram que o agente não mostrava totalmente as variáveis (n=1), a expiração de uso (n=2), relacionada ao limite do ChatGPT que impediu prosseguir a interação com o agente, informações respectivas (n=1), respostas grandes (n=1) e falta de autonomia (n=1). Essas respostas sugerem dois pontos principais: (i) que a modelagem de dados é um processo complexo, até para um agente inteligente especialista em GLA; e (ii) LLMs fechados como o ChatGPT apresentam obstáculos, como acesso aos agentes somente na versão Plus (não gratuita) e limite de uso. Tais pontos também se revelam ao analisar as interações dos estudantes: notou-se uma disparidade na solicitação de informações dos jogos pelo agente, onde em um caso foi feito de forma genérica, e nos demais, entre três a dez informações específicas foram solicitadas. Em relação às variáveis propostas na modelagem, a quantidade proposta pelo modelo variou de oito a doze.

Quantidade de variáveis: esse aspecto refere-se à suficiência das variáveis propostas pelo agente para representar o percurso do jogador e identificar indícios de aprendizagem (questão 9). 67% (n=6) consideraram a quantidade suficiente, enquanto 22% (n=2) a consideraram insuficiente, sentindo falta de dados (questão 11) específicos, como consultas ao algoritmo, e genéricos, como tempo para realizar uma ação. Quanto às justificativas das variáveis propostas (questão 10), 78% (n=7) as consideraram boas e apontaram que o agente auxiliou na identificação de variáveis nas quais não haviam pensado (questões 12 e 13), como nível de dificuldade (n=2), tentativas por fase (n=2), tempo (n=2) e dados específicos do jogo (n=4, 44%).

Eficiência em comparação com a técnica manual: ao comparar a técnica manual de modelagem de dados (descrita na Seção 3.1) com o uso do agente, 89% (n=8) dos participantes acreditam que o agente tenha sido mais eficiente (questão 20) e preferiram-no para modelar dados (questão 18). Os aspectos apontados pelos estudantes (questão 19) incluem: praticidade (n=3, 33%), apoio na geração de mais ideias (n=3) e identificação de dados que não haviam pensado (n=5, 56%). Um participante (11%, P7) informou não ter preferência pelo método de modelagem de dados, pontuando que o uso das técnicas depende da necessidade. Ressalta-se que, com o uso do agente, mais variáveis foram propostas pelos participantes em relação à modelagem manual.

A maioria dos testadores (n=8) se sentiu mais confiante ao usar o agente (questão

23), enquanto um testador preferiu a técnica manual pela autonomia em pensar nas variáveis de coleta. A respeito da dificuldade (questão 21), 67% (n=6) acharam a técnica manual mais difícil, enquanto 22% (n=2) acharam que ambas foram desafiadoras. A Figura 17 ilustra as notas individuais de dificuldade atribuídas pelos estudantes à modelagem manual (estudo descrito na Seção 3.1) e à modelagem com o agente (nesta pesquisa), organizadas conforme a ordem de resposta no formulário. Os dados evidenciam que a modelagem manual foi tida como a mais difícil.

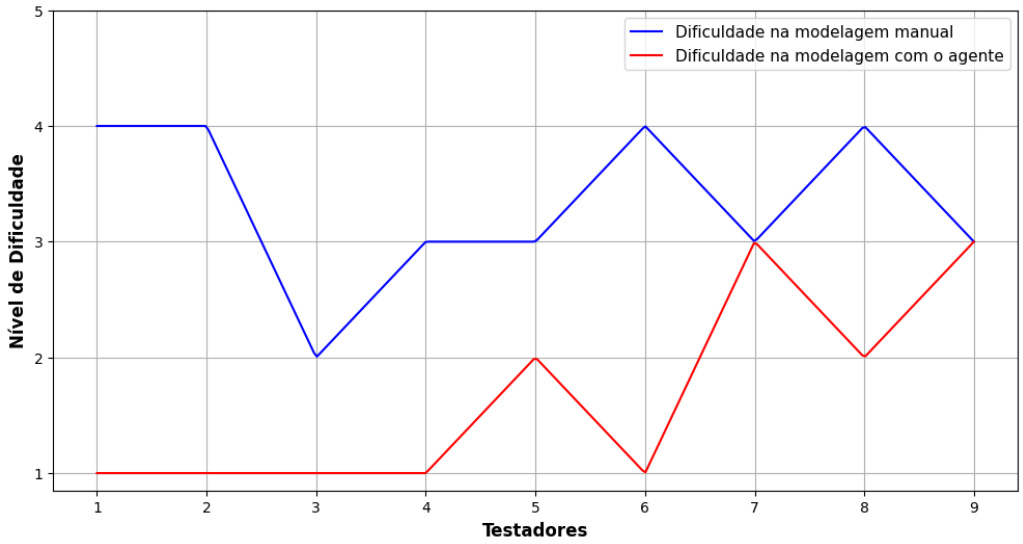


Figura 17 – Comparação do nível de dificuldade na modelagem de dados.

4.2.4 Considerações sobre a investigação

Em resposta à QP4.2, as percepções dos estudantes com o "*GLA Specialist*" foram positivas: unanimidade (100%) quanto a ser uma boa opção para modelagem de dados (questão 14), 56% (n=5) achou a interação fácil (questão 5), 78% (n=7) não o considerou difícil de usar (questão 7), 67% (n=6) julgou suficiente a quantidade de dados proposta (questão 9) e 78% considerou boas as justificativas apresentadas (questão 10). Na comparação com a modelagem manual, 89% (n=8) achou o agente mais eficiente (questão 20) e sentiu-se mais confiante para usá-lo (questão 23), 78% considerou a técnica manual mais difícil (questão 21) e 89% dos participantes prefere utilizar o agente (questão 18). Um dos estudantes (P7) se sentiu mais confiante ao utilizar a técnica manual, destacando a

autonomia que ela proporciona à modelagem de dados.

O desenho do estudo considerou a seleção de participantes que já haviam modelado dados de forma manual (Seção 3.1), de modo que fosse possível compará-la com a modelagem apoiada pelo "*GLA Specialist*". Nesse sentido, os participantes compreendiam as dificuldades na modelagem manual e puderam reconhecer com clareza os aspectos nos quais o agente forneceu de suporte adicional. Em contrapartida, como limitação, essa experiência prévia pode ter introduzido um viés nas percepções dos participantes, possivelmente favorecendo uma avaliação mais positiva da modelagem apoiada por LLM, bem como uma percepção de menor dificuldade durante o uso do agente. Outra limitação refere-se ao limite do uso do ChatGPT, que pode ter afetado a uniformidade das interações e a reprodutibilidade do estudo. No entanto, espera-se que o impacto tenha sido pontual, pois apenas dois participantes foram afetados e, posteriormente, conseguiram prosseguir com o uso do agente em outra conta.

Os *insights* dessa investigação, que buscou não somente avaliar o agente com o público-alvo, mas também comparar a técnica com a modelagem manual, incluem: (i) na percepção de estudantes que atuavam como *designers* de aprendizagem, o agente demonstrou-se como uma boa alternativa para o processo de modelagem de dados; (ii) o uso de LLMs fechados é um desafio, por conta da limitação de uso, instabilidade e impossibilidade de potencializar as respostas de outras formas, tais como *fine-tuning*; e (iii) as técnicas não se excluem e ambas possuem sua importância: enquanto o uso do agente visa auxiliar na modelagem, a técnica manual permite autonomia dos *designers* de aprendizagem, podendo ser usadas de forma combinada.

4.3 Síntese do Capítulo

Neste capítulo, buscou-se projetar, construir e avaliar uma estratégia baseada em LLM para auxiliar *designers* de aprendizagem a modelar dados em GLA, que não exigisse *prompts* complexos. O primeiro passo foi uma investigação em que se utilizou a função "MyGPTs" do ChatGPT. Por meio dela, elaborou-se um agente inteligente intitulado "*GLA Specialist*", ao qual se inseriu materiais sobre GLA e GLBoard, compondo sua base

de conhecimento. A auto-descrição do agente foi definida e refinada a partir de testes unitários, para garantir conformidade em auxiliar o *designer* de aprendizagem por meio de questionamentos e interações simples e objetivas. Sua avaliação foi realizada por especialistas em GLA, que demonstrou que o agente foi bem aceito e tido como uma ferramenta em potencial para auxiliar na área.

Em seguida, conduziu-se uma investigação para coletar percepções de estudantes de computação que atuavam como *designers* de aprendizagem. O objetivo foi analisar as experiências com o uso do agente e as percepções em comparação com a técnica manual. Os nove participantes foram os mesmos da investigação sobre modelagem de dados manual (Capítulo 3, Seção 3.1), que utilizaram o agente para modelar dados conforme seus jogos educacionais. Os resultados demonstram uma boa aceitação do agente pelos estudantes, que indicaram que sua interação foi fácil, que é uma alternativa eficiente e que, em maioria, as variáveis e justificativas propostas são adequadas. Quanto à técnica manual, os participantes acharam-na mais difícil e preferem utilizar o agente. Entretanto, assim como apontado por um estudante (P7), ela proporciona autonomia na modelagem. Dessa forma, ambas as técnicas não são exclusivas, e sim complementares: o uso do agente facilita a modelagem, enquanto que a técnica manual traz protagonismo ao *designer* de aprendizagem.

5

GLORIA: ANALISANDO JOGOS E ORIENTANDO A INSERÇÃO DE GLA

Este capítulo descreve a realização de um *Design* Participativo (Seção 5.1) para auxiliar na elicitação de requisitos da proposta, detalha o processo de construção do GLORIA e apresenta seu funcionamento (Seção 5.2). Dessa forma, este capítulo pertence aos Ciclo de Relevância e de *Design*, relacionado ao objetivo específico 4[1] desta dissertação.

5.1 *Design* Participativo

As investigações dos capítulos anteriores trouxeram achados relevantes para a pesquisa, sobretudo requisitos para a construção de uma ferramenta que auxilie *designers* de aprendizagem em GLA. Nesse sentido, optou-se também pela condução de um *Design* Participativo (DP) para a elicitação novos requisitos. O DP é uma abordagem utilizada em diversos contextos, como no *design* de sistemas computacionais, onde as pessoas-alvo da ferramenta atuam em sua concepção como especialistas, sendo os indivíduos que mais detém o conhecimento sobre o que fazem e o que precisam (Schuler; Namioka, 1993). Dessa forma, retornou-se ao Ciclo de Relevância[2]. Uma publicação científica sobre o DP encontra-se publicada nos Anais do WEI - CSBC (2026)[3].

[1] Construir uma ferramenta para auxiliar *designers* de aprendizagem a modelar dados em GLA, a partir dos resultados do corpo de conhecimento, das dificuldades/estratégias localizadas e do agente de conversação.

[2] Optou-se por não o incluir no Capítulo 3 para manter a ordem cronológica.

[3] https://sol.sbc.org.br/index.php/wei/article/view/43102

O DP foi utilizado para elencar novos requisitos para a ferramenta, a partir da emergência de necessidades de seus futuros usuários (público-alvo). Portanto, foram selecionados nove estudantes do curso de Licenciatura em Computação da UEA, que atuavam como *designers* de aprendizagem. Eles cursavam a disciplina de Fundamentos de *Software* Educacional (FSE), cujo objetivo é propor, construir e avaliar objetos de aprendizagem. Todos haviam participado das investigações sobre modelagem de dados, tanto de forma manual (Seção 3.1) quanto com o uso do "*GLA Specialist*" (Seção 4.2), possuindo experiência prévia na definição e organização de variáveis de GLA.

Nesse contexto, o DP foi conduzido em uma aula da disciplina, em três etapas: (i) requisitos e objetivo: apresentaram-se os requisitos e objetivos da atividade, que consistia em conceber, em trio, um sistema para auxiliar a modelar dados de jogos educacionais. Deveria ser construído um protótipo de baixa fidelidade (em papel), com as funcionalidades e os desenhos das telas. Os participantes podiam utilizar materiais como lápis de cor, régua, caneta, pincel, tesoura, *post-its* e cola; (ii) execução da atividade: os estudantes iniciaram a atividade com um *brainstorming* em equipe e elaboraram a proposta com os materiais disponibilizados (Figura 18); e (iii) apresentação das propostas: cada equipe exibiu e explicou sua proposta para a turma.

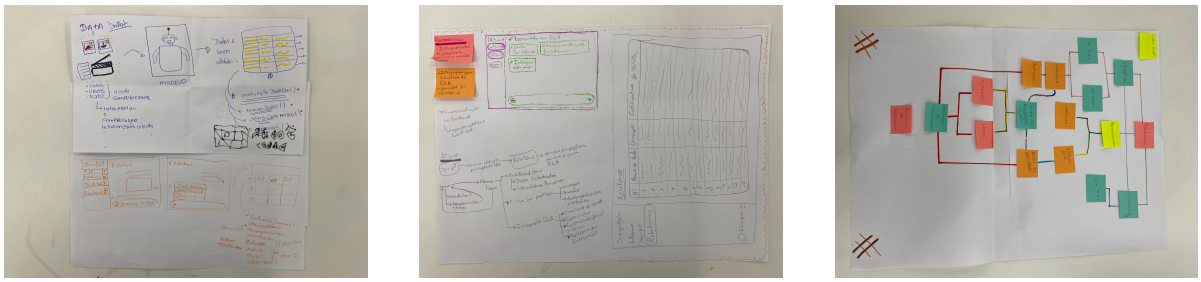


(a) Equipe 1. (b) Equipe 2. (c) Equipe 3.

Figura 18 – Equipes no *Design* Participativo.

Os resultados podem ser consultados na Figura 19, com três propostas distintas. A proposta 1, intitulada *DMBot* (Figura 19(a)), consiste em um *chatbot* para modelar dados, que recebe entradas visuais, como imagens e vídeos do jogo, além dos dados textuais. Seu propósito é compreender o jogo com mais clareza através dos recursos visuais, para reduzir a quantidade de texto informada pelos *designers* de aprendizagem ao modelo, que é um processo extenso e, dependendo do tamanho da descrição, o

modelo pode desconsiderar alguns aspectos (como observado na Subseção 3.3). Além disso, após gerar a tabela de modelagem de dados, o modelo também iria sugerir métodos genéricos para a coleta das variáveis definidas. Já na proposta 2 (Figura 19(b)), idealizou-se um sistema dependente do GLBoard, caracterizando-se como um módulo interno integrado ao LLM, que recebe como entrada os códigos do jogo implementado. A partir disso, o modelo interpreta os códigos e faz recomendações ao usuário de quais variáveis podem ser coletadas, gerando a tabela de modelagem de dados e fazendo a correspondência da parte do código em que a variável pode ser instanciada. E a proposta 3 (Figura 19(c)) trata-se de um modelo inteligente que auxilia na modelagem de dados, cuja entrada são os dados já capturados do jogo educacional do usuário. Desse modo, o modelo avalia os dados coletados e realiza uma previsão da evolução de aprendizagem do jogador, auxiliando tanto em avaliação (*insights* a partir dos dados), quanto na captura de dados (propondo melhorias na estrutura atual).



(a) Proposta 1. (b) Proposta 2. (c) Proposta 3.

Figura 19 – Propostas de sistemas para modelagem de dados.

As propostas elaboradas pelos estudantes foram fundamentais para identificar funcionalidades futuras do artefato, como: (i) inserção de recursos visuais (imagens e vídeos) para o modelo compreender o jogo com mais clareza, além das informações textuais; (ii) análise de código do jogo educacional, realizando a modelagem de dados e auxiliando o desenvolvedor a localizar trechos onde pode implementar as variáveis de GLA sugeridas; e (iii) adaptação da estrutura de captura de dados, a partir da análise dos *logs* gerados ao utilizar o jogo educacional. Além disso, pôde-se perceber que os estudantes mobilizaram conhecimentos prévios de computação para auxiliar na idealização das propostas, sugerindo funcionalidades que envolvem LLMs (presentes nas três propostas), técnicas de visão computacional e análise de estruturas computacionais

para auxiliar na modelagem de dados, como códigos e *logs* de interação. Após o *Design* Participativo, avançou-se novamente para o Ciclo de *Design*, onde se iniciou o processo de construção da ferramenta (Seção 5.2).

5.2 Proposta de ferramenta

Esta seção apresenta o processo de desenvolvimento e a ferramenta resultante, cujo objetivo é auxiliar *designers* de aprendizagem (estudantes, educadores e desenvolvedores) a modelar dados para jogos educacionais, considerando os achados das investigações anteriores e as ideias levantadas a partir do *Design* Participativo. A ferramenta é integrada a um LLM *open-source* (Qwen 2.5 14B), em que o usuário interage com um agente inteligente (*chatbot*) especialista em modelagem de dados para GLA, sem a necessidade de *prompts* complexos. O agente realiza questionamentos para compreender o jogo com mais clareza, contribuindo de forma especializada.

5.2.1 Processo de desenvolvimento

O desenvolvimento foi guiado pela metodologia de *game design* educacional de Pires et al. (2021), cujo objetivo é orientar a criação de artefatos educacionais relacionados a jogos (gamificação, *softwares*, ferramentas, jogos educacionais). As etapas dessa metodologia são ilustradas na Figura 20 e descritas a seguir.

Identificação do Problema: consistiu em localizar um problema de aprendizagem. Nesta dissertação, esse problema foi identificado em campo: observou-se que estudantes que atuam como *designers* de aprendizagem apresentam dificuldades na modelagem de dados, um passo anterior à coleta de dados em GLA. Diante disso, o mestrando realizou um *brainstorming* (tempestade de ideias) para centralizar as ideias, onde se elaborou propostas de telas e fluxo de navegação em um *sketchbook* (Figura 21(a)) e no quadro branco (Figura 21(b)). Essa etapa considerou: os achados da literatura, principalmente na identificação de trabalhos correlatos (Capítulo 2), as investigações conduzidas sobre modelagem de dados com o apoio de LLM (Capítulo 3), a proposta/avaliação do

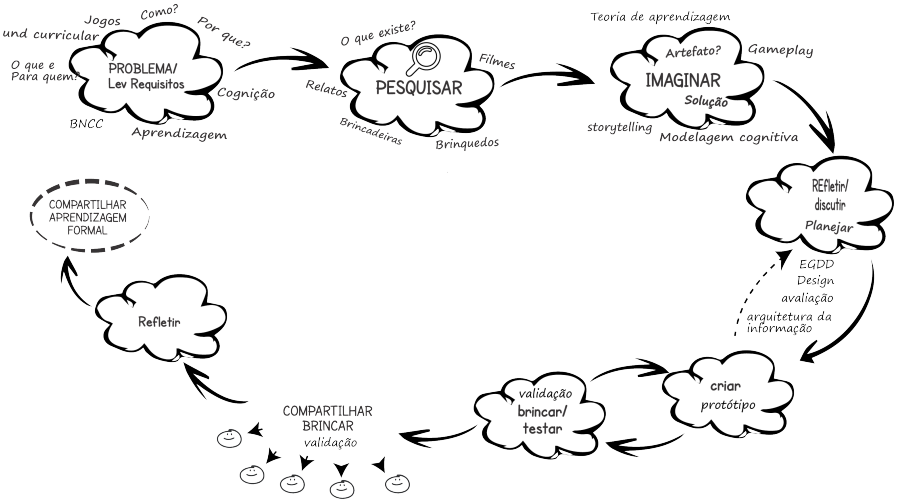
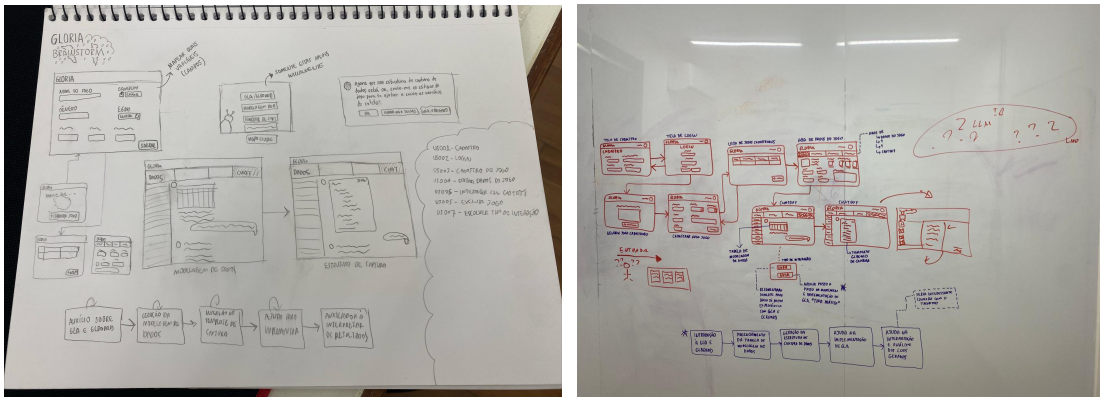


Figura 20 – Metodologia de *game design* educacional (Pires et al., 2021).

agente inteligente (Capítulo 4) e o *Design* Participativo (Seção 5.1).



(a) Brainstorming em caderno (*Sketch Book*). (b) Brainstorming em quadro-branco.

Figura 21 – *Brainstormings* do GLORIA.

Pesquisar: buscou-se conduzir uma revisão da literatura para entender o estado da arte referente ao tema. Diante disso, a pesquisa realizada no Capítulo 2 apontou que há uma lacuna na literatura referente à utilização de LLMs no campo de GLA: embora existam trabalhos que incluam o uso desses modelos para GLA, são focados em jogos específicos e não contemplam a modelagem de dados. Já os *frameworks* existentes sobre GLA auxiliam na implementação das técnicas, mas não são integrados com LLMs. A etapa também incluiu as investigações executadas nos Capítulos 3 e 4, que revelaram potencial no uso de LLMs para atividades de GLA, sobretudo na geração de estruturas de captura de dados. Resultados positivos emergiram também na criação do agente

especialista em GLA, sendo apontado como uma boa estratégia para minimizar a complexidade da modelagem de dados, na perspectiva de estudantes de computação e especialistas na área. Entretanto, também relevou uma limitação do uso do agente em modelos fechados. Esses pontos geraram *insights* que auxiliaram na etapa seguinte.

Imaginar: dadas as contribuições da etapa de pesquisa conduzida anteriormente, definiu-se que uma ferramenta seria construída. Essa escolha deu-se: (i) pela possibilidade de integração com um LLM, bem como pela sua personalização por meio de engenharia de *prompt* e *fine-tuning*, além da inclusão de técnicas de RAG; (ii) pelo uso e acesso, considerando que modelos como o ChatGPT restringem a criação e customização de agentes (*MyGPTs*) aos planos pagos, bem como limitam a quantidade de interações; e (iii) pela integração futura com o sistema GLBoard, tornando-o um módulo complementar anterior à implementação do *template* de dados nos jogos educacionais. Diante disso, definiu-se que o nome da ferramenta seria GLORIA, um acrônimo que descreve seus principais componentes:

- **G - *Game*:** os jogos são os objetos de aprendizagem aos quais a proposta destina-se, cujos dados serão necessários para os processos de GLA;
- **L - *Learning*:** a aprendizagem é o objetivo da proposta, visando auxiliar em processos que permitam identificar suas evidências em jogos educacionais;
- **O - *Observation*:** através de seus processos de GLA, possibilitará aos educadores e desenvolvedores observar/acompanhar as tomadas de decisão dos jogadores;
- **R - *Recommendation*:** diz respeito à recomendação de variáveis de GLA (proposta e refinamento), conforme os dados do jogo e as interações via *chatbot*;
- **I - *Intelligence*:** ilustra que há uma inteligência por trás da proposta, referente ao agente inteligente baseado em LLM, que é especialista em GLA;
- **A - *Analytics*:** refere-se ao alinhamento da modelagem de dados com os objetivos de aprendizagem do jogo, visando viabilizar posteriormente as análises dos *logs*.

Refletir/Discutir: realizou-se o planejamento da ferramenta, que considera a elaboração e a documentação de seus principais componentes. Nesse aspecto, inclui a

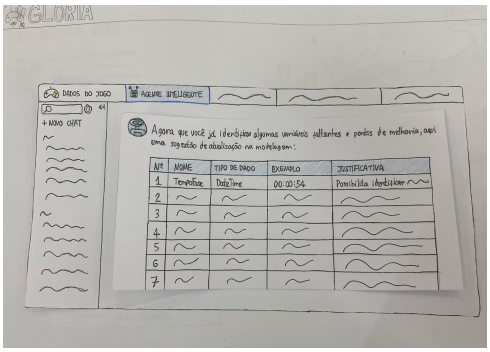
definição dos requisitos, a arquitetura da ferramenta (ambos descritos na Seção 5.2.2), as tecnologias e banco de dados, o diagrama de caso de uso e o diagrama de entidade-relacionamento (DER), disponíveis no Apêndice L.

Criar: visando refinar os requisitos e validar o escopo da ferramenta, versões preliminares (protótipos) foram concebidas. A Figura 22 ilustra os protótipos, dos quais: (i) baixa fidelidade: elaborado utilizando materiais como: papéis A4, lápis, canetas em gel, lápis de cor e borracha (Figura 22(a)). Por ser mais simples, permite que alterações mais robustas sejam aplicadas; (ii) média fidelidade (v01): concebido na ferramenta Figma, consolidando as primeiras interfaces e ilustrando o fluxo de navegação do usuário (Figura 22(b)); (iii) média fidelidade (v02): construído com o Figma Make[4], contendo o fluxo principal da ferramenta (gestão de jogos, agente inteligente e modelagem/captura de dados, Figura 22(c)). O objetivo de ambas as versões é validar se a usabilidade da ferramenta e experiência do usuário estão adequadas; e (iv) POC: Prova de Conceito (*Proof of Concept*), versão inicial implementada em ambiente de desenvolvimento. Contempla alguma das principais funcionalidades da ferramenta e antecede o protótipo de alta fidelidade (versão mais fidedigna).

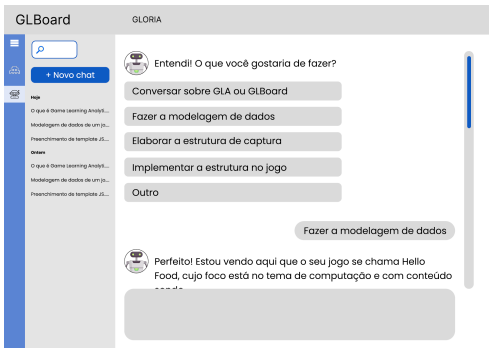
Inicialmente, a POC apresentava somente o envio do *prompt* e o retorno do LLM. Posteriormente, incluiu-se a página principal (*landing page*), o cadastro de informações sobre o jogo educacional e o agente inteligente (Figura 22(d)). Seu propósito é verificar se as funcionalidades estão adequadas e orientar o desenvolvimento dos próximos passos, ainda não incluindo os elementos estéticos de *design*. Em relação ao protótipo de alta fidelidade, suas telas e funcionalidades são apresentadas na Subseção 5.2.2.

Brincar/Testar: essa etapa opera de forma cíclica com a etapa de "Criar", onde os protótipos são validados por especialistas na área ou estudantes de computação. Desse modo, após a validação, uma lista de inconsistências e melhorias é elaborada, retornando-se para a etapa anterior para corrigi-las. Neste trabalho, a primeira versão (baixa fidelidade) foi validada por uma especialista em GLA. Ela possui oito anos de experiência com jogos educacionais, leciona na área e possui doutorado. Os principais ajustes elencados referem-se a não utilizar uma tabela de modelagem de dados (para fa-

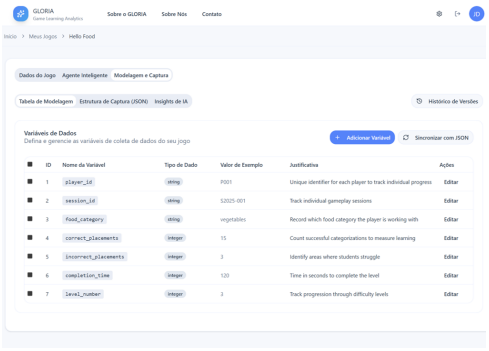
[4] Ferramenta de IA que recebe textos e imagens como entrada, capaz de criar códigos em *front-end* (como *TypeScript*) com componentes reutilizáveis, lógica e interações.



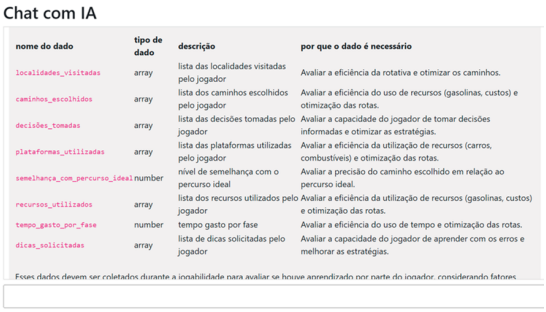
(a) Baixa fidelidade.



(b) Média fidelidade (v01).



(c) Média fidelidade (v02).



(d) Prova de conceito (POC).

Figura 22 – Protótipos: baixa, média fidelidade e POC, respectivamente.

cilitar o entendimento do processo) e a remoção da análise de *logs*, visto que aumentaria a complexidade e trataria da pós-captura de dados, enquanto que o foco da ferramenta é majoritariamente auxiliar na pré-coleta de dados. Quanto à v01 do protótipo de média fidelidade, a validação também foi realizada pela especialista em GLA, cujo principal ajuste refere-se a organização da barra lateral e a inserção de opções para visualizar a modelagem de dados e a estrutura de captura gerada.

Em relação à v02 e à POC, foram validadas mediante testes unitários e exploratórios por um especialista em GLA com experiência em LLMs (o mestrando desta dissertação). Os ajustes na v02 incluem a remoção da função "Adicionar variável" e da aba de "*Insights* de IA", pois aumentariam a complexidade da ferramenta, podendo ser incorporadas posteriormente junto ao GLBoard. Além disso, acrescentaram-se as opções de "Modo Livre" e "Modo Tutorial", para adaptar a interação com o *chatbot* conforme o nível de conhecimento do usuário em GLA. Já na POC, localizaram-se inconsistências que foram ajustadas e resultaram em versões mais atualizadas do protótipo. Os principais ajustes referem-se a: (i) correção de *bugs*, sobretudo em relação a área de

chatbot, interação com o agente e tabela de modelagem de dados; (ii) refinamentos na usabilidade, como o posicionamento e destaque visual aos botões; (iii) alterações no *design*, principalmente em relação a padronização; e (iv) implementação de melhorias, como o envio da mensagem no modo tutorial, animação de "Pensando" no *chatbot*, opção de "Parar geração" de mensagens, entre outras.

Compartilhar/Brincar: opera de forma semelhante com a de "Brincar/Testar", tratando-se dos testes conduzidos com o protótipo de alta fidelidade após implementação de ajustes na POC. A avaliação ocorre com o público-alvo, visando verificar se os objetivos da solução estão sendo efetivamente atingidos. O protótipo está descrito na Seção 5.2.2, enquanto a execução dos testes é detalhada no Capítulo 6.

Refletir: consiste nas análises realizadas a partir dos testes com o protótipo de alta fidelidade, que trazem reflexões sobre a proposta. Dessa forma, possibilita identificar inconsistências e mapear pontos de melhoria, trazendo *insights* valiosos. As discussões dessa etapa são apresentadas no Capítulo 6.

Compartilhar aprendizagem formal: nessa etapa, após a implementação dos ajustes e melhorias, obtém-se a primeira versão do produto. Também denominada de *release* (lançamento), é a versão mais fidedigna e alinhada ao público-alvo. Como a ferramenta atua na etapa introdutória do fluxo de GLA (pré-coleta de dados), sua *release* será consolidada após a integração ao GLBoard.

5.2.2 Ferramenta para modelagem de dados

O artefato em formato de ferramenta, intitulado GLORIA, tem como objetivo auxiliar *designers* de aprendizagem na modelagem de dados em GLA, sejam eles:

- **Estudantes:** atuando nessa função em disciplinas, visando a construção e avaliação de jogos educacionais. Por exemplo, acadêmicos de Licenciatura em Computação.
- **Educadores:** que utilizarão, em suas turmas, jogos educacionais com técnicas de GLA integradas, para os quais modelaram os dados. Com base nos dados coletados, poderão monitorar o progresso dos estudantes e realizar intervenções em sala de aula, quando necessário.

- **Desenvolvedores:** que atuam como *game designers* e implementarão as variáveis de GLA no código do jogo. Os dados coletados permitirão que localizem inconsistências e aprimorem o *level design*.

O GLORIA assume um papel fundamental no início do fluxo de GLA (Figura 23), contemplando: (i) o cadastro do jogo, para registrar informações sobre o jogo educacional (como os listados no Apêndice D); (ii) a modelagem de dados, apoiando a definição dos dados a coletar conforme as características e objetivos de aprendizagem do jogo, bem como sua relevância para evidenciar aprendizagem, com o apoio de um LLM; e (iii) o *template* de dados, compondo a estrutura de captura do GLBoard com as variáveis de GLA derivadas da modelagem. Esses componentes fornecem a base para as atividades posteriores de GLA, apoiadas pelo GLBoard: (iv) implementação do *template* nos jogos; (v) coleta dos *logs* dos jogadores; (vi) análise e interpretação dos dados em *dashboards* e relatórios; e (vi) refinamento do jogo educacional, com ajustes no *game design* e nos conteúdos conforme os *insights*. Portanto, o GLORIA pode representar uma importante peça do quebra-cabeça de GLA, favorecendo as etapas subsequentes e contribuindo para que, em contextos escolares e universitários, os dados coletados auxiliem na avaliação da aprendizagem e subsidiem decisões pedagógicas voltadas ao desenvolvimento de habilidades e competências alinhadas aos conteúdos curriculares.



Figura 23 – Fluxo de GLA, ilustrando a atuação do GLORIA e do GLBoard.

A ferramenta será incorporada ao modelo GLBoard para atuar como módulo complementar. Ela é integrada a um LLM, responsável por processar as informações dos jogos e as instruções fornecidas pelo usuário, e gerar respostas em linguagem natural. A partir dessa integração, é disponibilizado um agente inteligente (*chatbot*) especialista em GLA, associado a um jogo educacional. Para isso, as informações cadastradas pelo usuário são enviadas internamente ao modelo por meio de um *prompt* contextual. Dessa forma, as respostas do agente são direcionadas ao contexto do jogo, orientando o *designer* de aprendizagem de modo mais específico.

O usuário pode interagir com o agente de dois modos: (i) livremente, semelhante à modelos como o ChatGPT; ou (ii) com o tutorial, sendo guiado pelo *chatbot* sobre conceitos de GLA e GLBoard, modelagem de dados e composição da estrutura de captura (*template*). A função do agente não é atuar como protagonista, mas sim como um tutor que incentiva o usuário a compreender o processo, instigando-o a refletir e interagir com o *chatbot*. Somente quando o usuário desejar prosseguir e considerar que os dados definidos contemplam as interações do jogador relacionadas aos objetivos de aprendizagem, o agente inteligente passa a contribuir de forma mais direta na modelagem de dados, aplicando modificações e propondo novas variáveis de captura. Após a modelagem, o agente adiciona e organiza as variáveis de GLA no *template* JSON do GLBoard (descrito na Seção 2.4), de modo que o usuário possa implementar em seu projeto conforme as especificidades do jogo educacional.

As variáveis de GLA elencadas e refinadas pelo agente no *template* do GLBoard pertencem a duas variáveis gerais: (i) $path\_player$, que armazena o percurso do jogador na fase, incluindo as ações e os registros de tempo (*timestamps*); e (ii) $finalized\_challenges$, que consolida variáveis atômicas/agregadas ou conjuntos de objetos referentes aos desafios alcançados. As variáveis do $finalized\_challenges$ são obtidas após a *gameplay*, a partir de uma ou mais variáveis do $path\_player$. Assim, enquanto o $path\_player$ registra as ações realizadas (por exemplo, o momento em que o jogador pulou em uma plataforma), o $finalized\_challenges$ as quantifica (por exemplo, a quantidade de personagens com os quais interagiu). As particularidades de cada jogo influenciam a definição dos dados associados a ambas as variáveis.

Por exemplo, a Figura 24 apresenta variáveis de GLA que o GLORIA pode propor ao jogo "Cadê minha pizza?" (Apêndice A). No caso do percurso do jogador (*path_player*), as variáveis permitem rastrear suas ações ao longo da partida, como: (i) entregador selecionado, com seu identificador; (ii) residência escolhida para entrega, coletando seu identificador (nome), a verificação da adjacência e o instante da interação; e (iii) movimento realizado, com a quantidade atual de gasolina do entregador, o custo da rua percorrida e o respectivo registro temporal. Já em *finalized_challenges*, as variáveis representam indicadores agregados da partida, como a quantidade de entregas realizadas e de gasolina utilizada, o tempo total de ações, dentre outros. Desse modo, o conjunto de variáveis pertencentes ao *path_player* e ao *finalized_challenges* permite rastrear com precisão as ações do jogador e sintetizar os resultados alcançados ao longo da *gameplay*, facilitando futuramente a identificação dos indícios de aprendizagem.

path_player

```
"selected_deliverer": string
"residence_clicked": {
      "name": string
      "is_adjacent": bool
      "timestamp": DateTime }
"movement": {
      "actual_fuel": int
      "street_cost": int
      "timestamp": DateTime }
```

finalized_challenges

```
"deliverers_used": int
"total_residences_visited": int
"total_non_adjacent_clicks": int
"movements_made": int
"fuel_remaining": int
"fuel_used": int
"total_street_cost": int
"total_actions_time": int
```

Figura 24 – Variáveis do jogo "Cadê minha pizza?".

Embora as variáveis de GLA pertencentes ao *path_player* e ao *finalized_challenges* sejam específicas aos jogos educacionais, elas ainda precisam ser implementadas manualmente pelo desenvolvedor no código do jogo. Essa é uma das limitações atuais do GLBoard, sobretudo porque essas variáveis são armazenadas em formato textual (*string*), o que exige um tratamento adicional para quantificar as informações registradas nos *logs* e permitir análises com maior clareza. Como perspectiva futura, pretende-se que a estrutura de GLA personalizada, gerada com o GLORIA, seja refletida automaticamente no *template* de dados do GLBoard. Assim, o desenvolvedor precisaria apenas atribuir valores às variáveis correspondentes, sem reorganizar manualmente a estrutura

no código, facilitando também a análise dos dados coletados.

O agente presente no GLORIA também é responsável por analisar se os dados cadastrados pelo usuário são suficientes para iniciar a modelagem dos dados. Caso ele identifique campos errôneos, faltantes ou genéricos, solicitará ao usuário que inclua as informações necessárias. As variáveis de GLA elaboradas pelo agente em conjunto com o usuário têm como objetivo fornecer indícios de aprendizagem, articulando *game design* e aspectos pedagógicos. A visualização desses indícios não ocorre no GLORIA, mas sim no GLBoard, que exibe os *logs* coletados após a implementação do *template* de dados, e apresenta *dashboards* e relatórios personalizados conforme métricas, que permitem validar se os dados são suficientes para evidenciar aprendizagem.

Nesse sentido, espera-se que o fluxo de GLA ocorra de forma iterativa-incremental: os dados são modelados no GLORIA, gerando as variáveis de GLA; o desenvolvedor as implementa no jogo por meio do GLBoard e gera registros de interação a partir delas; a análise desses *logs* permite a identificação de inconsistências e melhorias; o desenvolvedor realiza modificações no *game design* e retorna ao GLORIA para aprimorar a modelagem e a estrutura de captura. A qualidade das variáveis não está sujeita apenas à capacidade do agente de propor variáveis (atividade computacional), mas também ao nível descritivo das informações do jogo inseridas pelo usuário. Mecânicas mais complexas e conteúdos mais abrangentes podem influenciar na modelagem de dados, caso não estejam bem descritos ou se o *game design* não foi projetado de forma adequada.

As funcionalidades da ferramenta foram desenvolvidas conforme as pesquisas na literatura e as investigações conduzidas, que trouxeram *insights*: (i) os modelos de GLA da literatura não possuem interação com LLMs; (ii) há uma ausência de evidências que considerem o estudante como um *stakeholder*, atuando como *designer* de aprendizagem; (iii) a modelagem de dados é uma atividade desafiadora, ainda abordada de forma limitada em pesquisas de GLA; (iv) LLMs, em condições limitadas, são capazes de modelar dados e gerar estruturas de GLA; (v) estudantes apresentam dificuldades em elaborar *prompts* para orientar um LLM a gerar uma estrutura de GLA com uma técnica de questionamentos; (vi) um agente inteligente pode contribuir para a área de GLA, auxiliando *designers* de aprendizagem a modelar dados, não necessitando de

prompts complexos; e (vii) LLMs fechados possuem limitações, como restrição de uso.

Esses *insights* orientaram os requisitos descritos na Tabela 8. O GLORIA será integrado ao GLBoard como módulo complementar, contudo, para avaliá-lo isoladamente, esta versão opera de forma independente. As informações técnicas da ferramenta, incluindo tecnologias, banco de dados e diagramas de caso de uso e Entidade-Relacionamento (DER), estão disponíveis no Apêndice L.

RF	Nome	Descrição
RF01	Exibir tela inicial	A ferramenta deve exibir uma *landing page* interativa com informações sobre seu objetivo e suas principais funcionalidades, além de acesso ao manual de uso e a uma área dedicada aos criadores
RF02	Cadastrar usuários	A ferramenta deve permitir o cadastro de usuários
RF03	Realizar *login*	A ferramenta deve incluir uma autenticação via *login* para que o usuário possa acessá-la
RF04	Cadastrar jogos	A ferramenta deve permitir o cadastro de jogos educacionais, com dados de *gameplay* e de aprendizagem
RF05	Gerenciar jogos	A ferramenta deve incluir uma opção para gerenciar (buscar, acessar, editar e excluir) os jogos cadastrados
RF06	Definir modo de interação	A ferramenta deve possibilitar que o usuário determine o modo de interação com o *chatbot*
RF07	Interagir com *chatbot*	A ferramenta deve incluir uma área para o usuário interagir com o *chatbot*, podendo interromper a geração de texto e criar novas conversas
RF08	Fornecer *feedback* de geração	A ferramenta deve incluir uma animação para indicar quando o LLM estiver "pensando" (em processamento), visando situar o usuário
RF09	Armazenar histórico	A ferramenta deve armazenar todo o histórico de interação do usuário com o *chatbot*
RF10	Gerenciar conversa	A ferramenta deve permitir buscar, renomear ou excluir uma conversa
RF11	Modelar dados	A ferramenta deve, em ambos os modos de interação, auxiliar o usuário na modelagem de dados. Inicialmente, deve orientar e fazer questionamentos, para depois incluir modificações e propor novas variáveis
RF12	Gerar estrutura de captura	A ferramenta deve, a partir da modelagem de dados definida, preencher o *template* JSON do GLBoard e retornar ao usuário uma estrutura de captura conforme seu jogo educacional
RF13	Visualizar modelagem e estrutura	A ferramenta deve permitir que o usuário visualize, em uma área própria, a modelagem de dados (tabela) e a estrutura de captura (JSON)
RF14	Exportar JSON	A ferramenta deve possibilitar a exportação da estrutura de captura em formato JSON, para uso no jogo educacional

Tabela 8 – Tabela de Requisitos Funcionais.

A representação da arquitetura da ferramenta é baseada no *C4 Model*, uma abordagem que organiza a visualização arquitetural em diferentes níveis de abstração, favorecendo a elaboração de diagramas menos complexos e apoiando uma comunicação mais clara entre as equipes envolvidas no projeto (Brown, 2013). O *C4 Model* é composto por uma hierarquia de quatro níveis: Contexto, Contêiner, Componente e Código (ou Classe). Cada nível apresenta um grau de detalhamento mais específico, possibilitando uma visão macro (geral) e micro (específica) da ferramenta, conforme a necessidade.

A Figura 25 ilustra a visão de Contêiner do GLORIA conforme o *C4 Model*, cujo "Usuário" é o *designer* de aprendizagem que irá utilizá-lo, seja um estudante,

educador ou desenvolvedor. Ele fará uso da "Aplicação Web" para acessar o *front-end* da ferramenta, desenvolvida com a tecnologia *React*. A comunicação com a "API do *back-end*" será via requisições *HTTPS*, que é responsável pela orquestração da ferramenta: envia os dados jogos para a API de IA e recebe as mensagens geradas pelo LLM. Além disso, consulta e faz gravações em um banco de dados relacional (*PostgreSQL*), que armazena os dados dos usuários, dos jogos e as modelagen de dados.

A API de IA foi construída com a linguagem *Python* (versão 3.11.13) e tem conexão direta com o *Ollama*: um "Serviço de Inferência" que executa localmente os LLMs e os disponibiliza para uso por meio de sua própria API (Ollama, 2023). Conforme os dados do jogo, um *prompt* é composto na API de IA e enviado ao *Ollama* (Apêndice M), que atualmente se encontra hospedado em um servidor local[5]. O modelo utilizado é o Qwen 2.5 (14B), que, além de não exigir elevado poder computacional (classificado como "médio" em relação à quantidade de parâmetros, conforme descrito na Seção 2.6), apresentava notas positivas no *Open LLM Leaderboard* para os *benchmarks* considerados em sua seleção, conforme detalhado no Apêndice L.

Dessa forma, o *Ollama* executa internamente o LLM com o *prompt* recebido e retorna a resposta para a API de IA. Em relação à modelagem de dados, o *Ollama* requisita da API o uso da *tool*, uma função responsável por gerá-la. Diante disso, por meio da *tool*, um novo *prompt* é enviado para o *Ollama* com os requisitos da geração (Apêndice N). Por sua vez, o *Ollama* novamente executa o LLM com o novo *prompt* e retorna a modelagem para armazená-la no banco de dados. A partir dessa modelagem, a estrutura JSON correspondente ao *template* do GLBoard é construída.

[5] Configurações: placa de vídeo NVIDIA GeForce RTX 4080, processador Intel Core i7-12700KF (12ª geração) e memória RAM DDR5 de 64 GB.

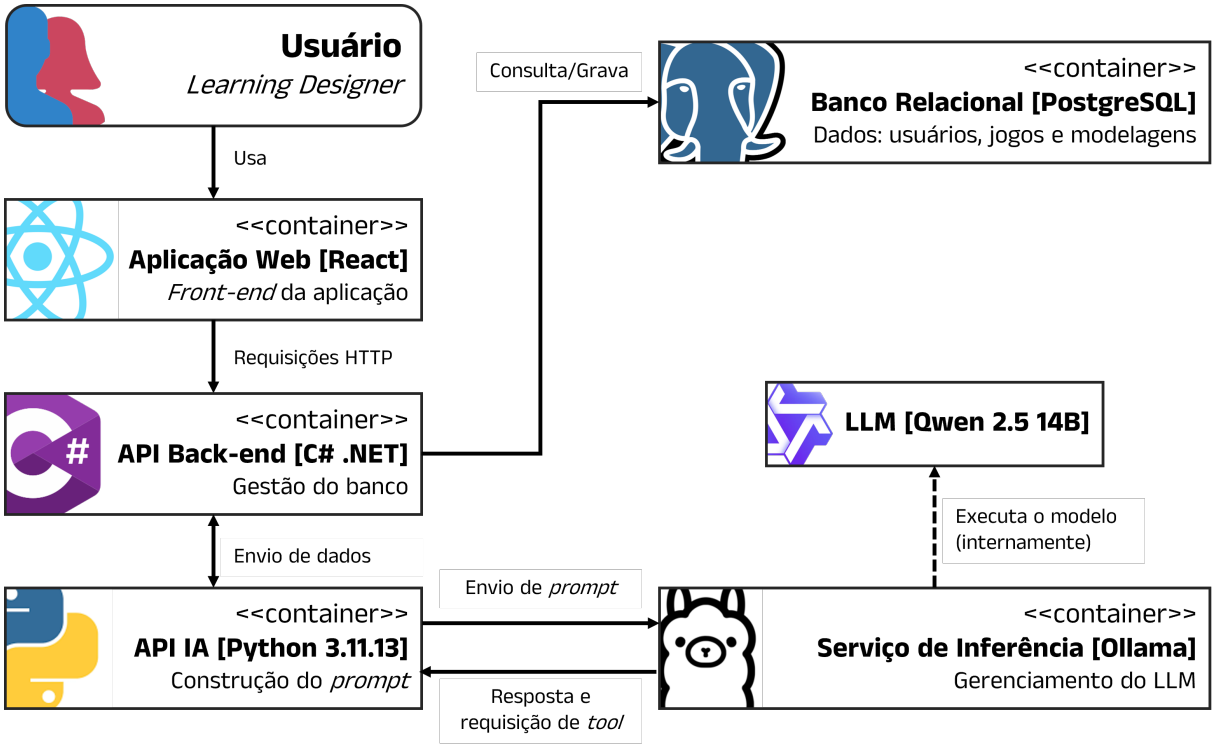


Figura 25 – Arquitetura do GLORIA. Modelo C4 - Nível de Contêiner.

O fluxo de telas da versão atual do GLORIA é apresentado na Figura 26. As cores amarelas simbolizam funcionalidades genéricas da ferramenta, enquanto que as cores azuis representam os componentes de jogos. Na cor verde, o agente inteligente, com uma área para visualização da modelagem de dados e estrutura de captura. Em seguida, cada tela é descrita, junto com as referências às figuras respectivas.

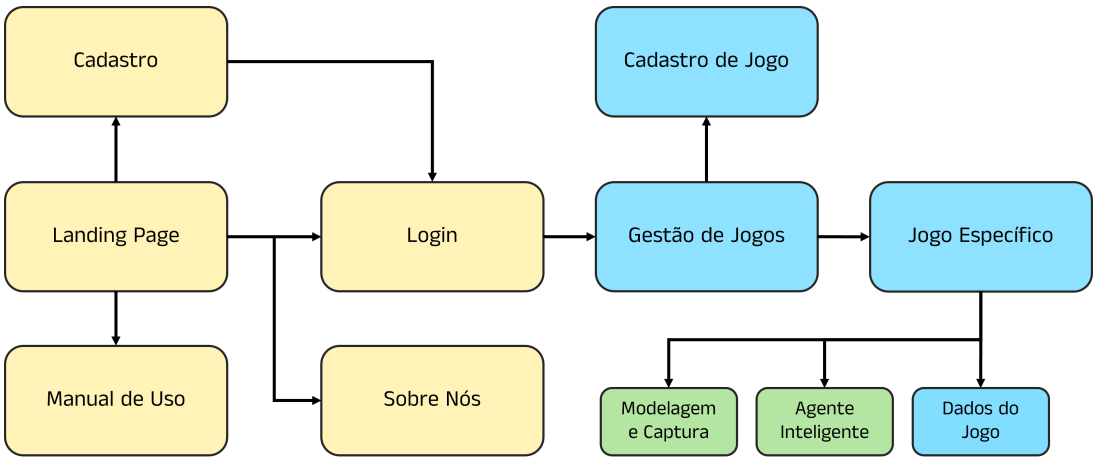


Figura 26 – Fluxo de telas do GLORIA.

- ***Landing Page* (Figura 27):** consiste na página de boas-vindas, incluindo infor-

mações sobre o que é o GLORIA, como funciona, por que usá-lo, exemplo de aplicação, público-alvo e integração com o GLBoard;

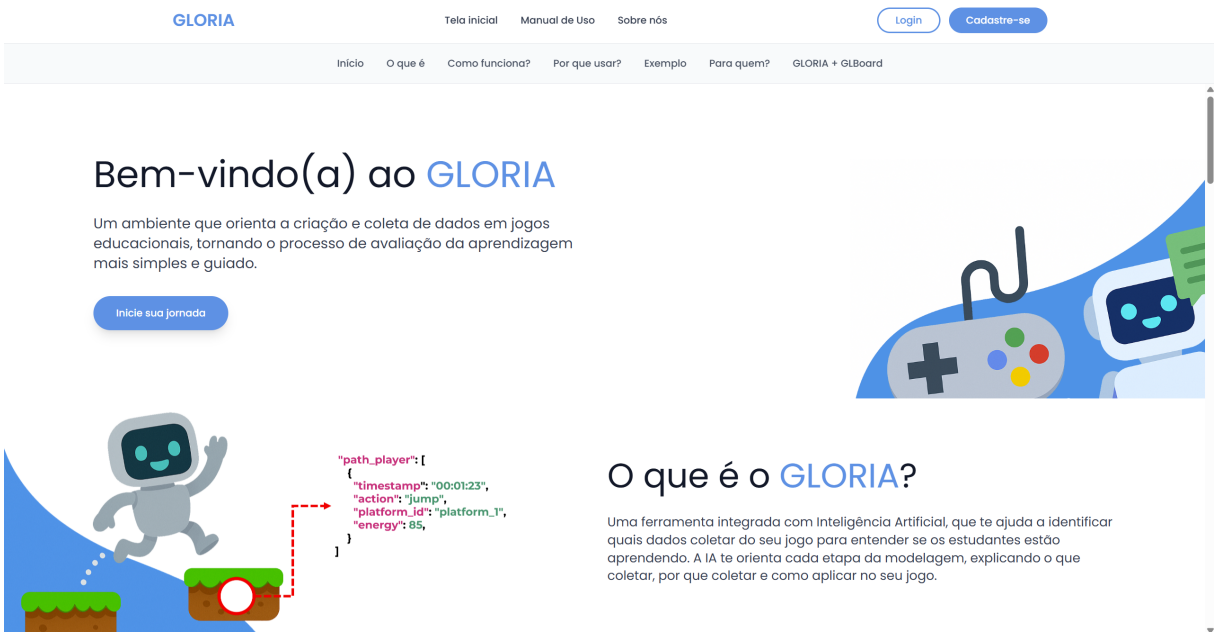


Figura 27 – *Landing page* da ferramenta.

- **Manual de uso (Figura 28):** apresenta orientações de uso da ferramenta em três seções: (i) cadastro de jogo; (ii) modelagem de dados; e (iii) *template* de dados. Cada seção contém um *card*, além de informações em tópicos. Há também um *card* sobre a integração com o GLBoard e o botão “Começar agora”;

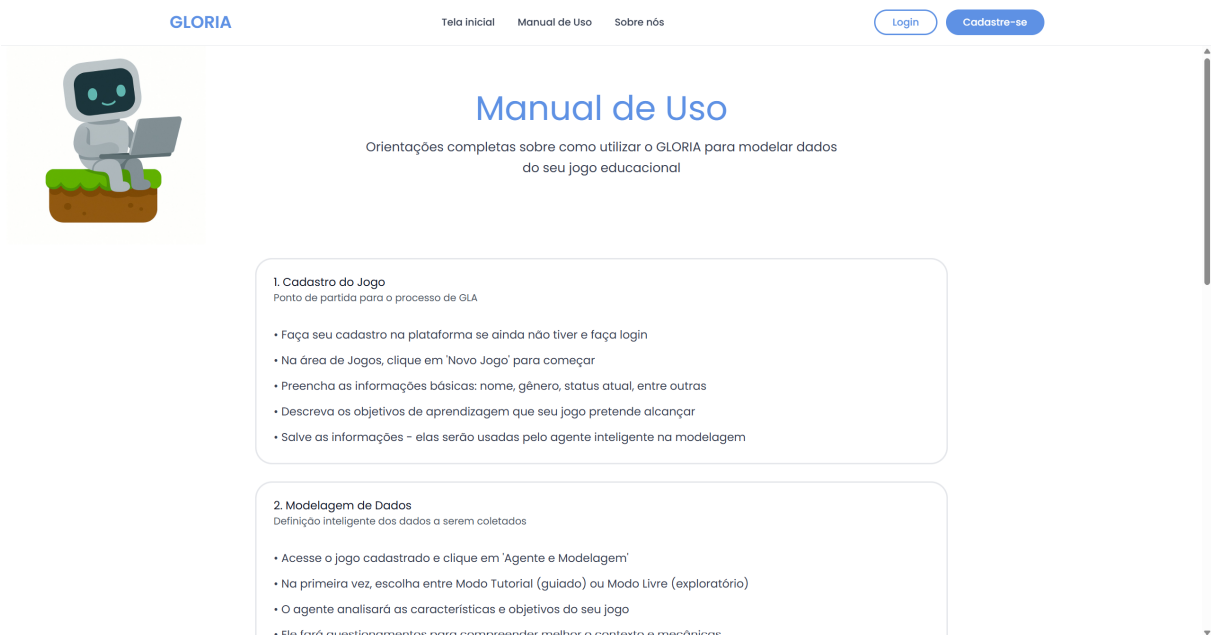


Figura 28 – Manual de uso.

- **Sobre nós (Figura 29):** apresenta a equipe responsável pela criação e desenvolvimento da ferramenta. Está organizada em (i) coordenação, sob orientação das professoras doutoras Elaine Harada, Fernanda Pires e Marcela Pessoa; e (ii) desenvolvedores, com Fabrizio Honda (mestrando desta dissertação, que atuou como *Product Owner*), Defala Filho e Victor Braga (desenvolvedores de IA), Davi Aguiar, Cleidiana Alves e José Brito (desenvolvedores *front-end*), Valdenei Lopes (desenvolvedor *back-end*) e Henrique Furtado (*DevOps*);

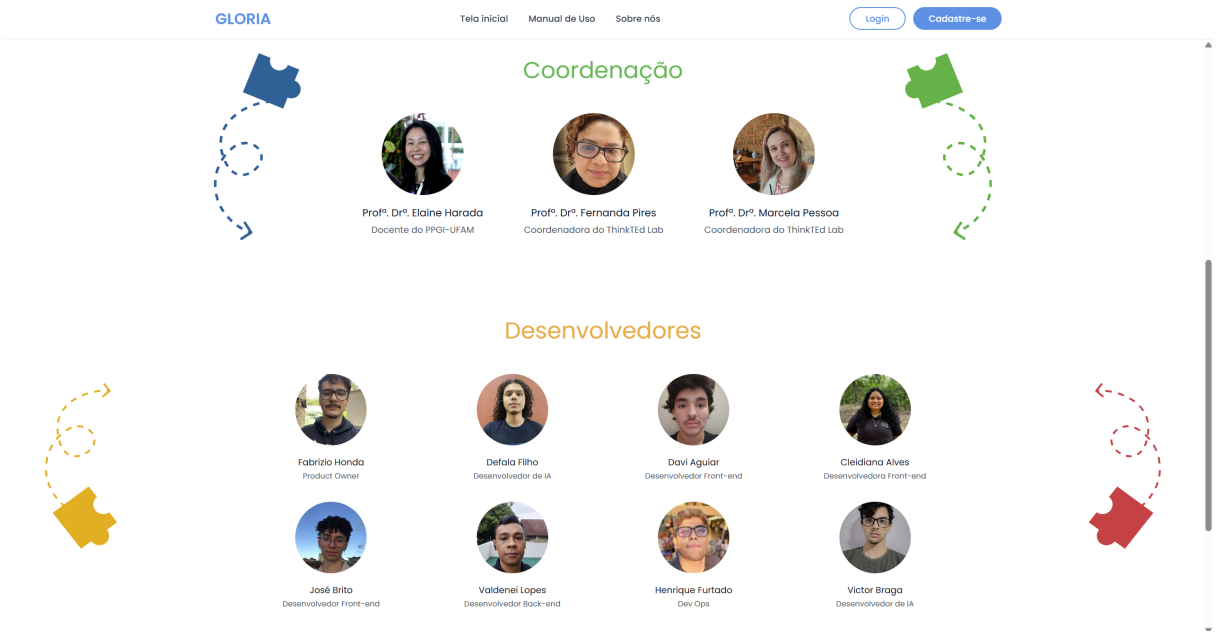


Figura 29 – Página de "Sobre nós".

- **Cadastro (Figura 33):** apresenta os campos necessários para o cadastro do usuário: nome completo, nome de usuário, e-mail, data de aniversário, escolaridade, *checkbox* para indicar se a formação foi concluída, gênero, senha e confirmação de senha. A tela também inclui os botões "Criar Conta" e "Faça *login* aqui";

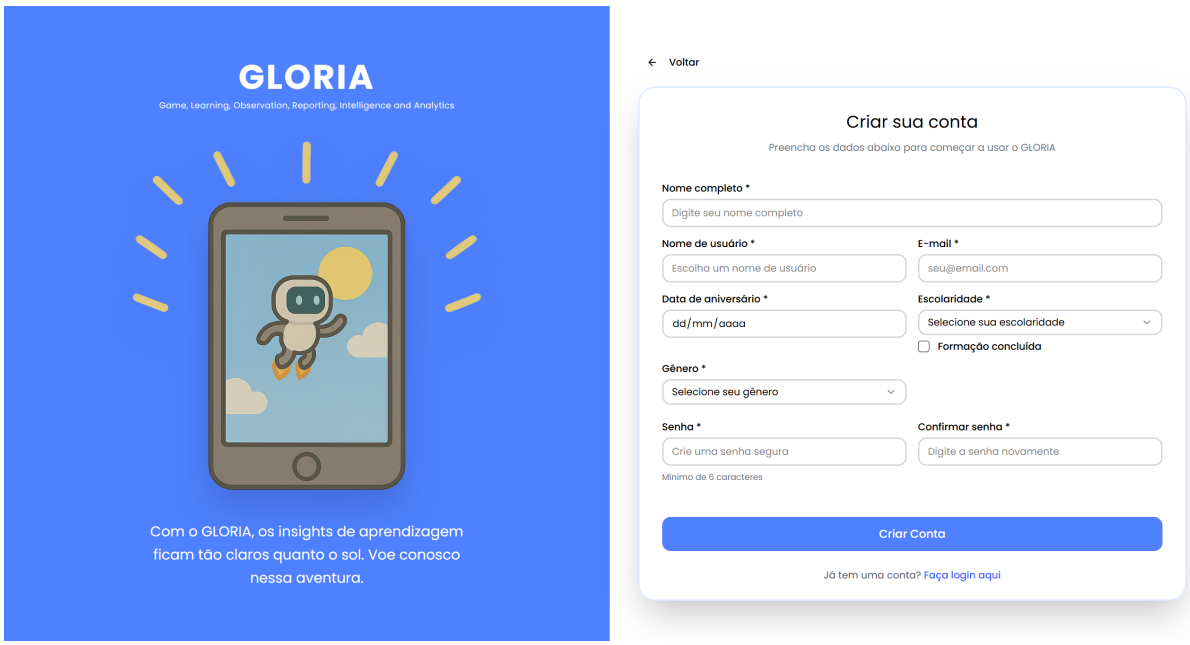


Figura 30 – Cadastro de usuário.

- ***Login* (Figura 31):** possui campos para o nome de usuário e senha (com opção para ocultá-la ou exibí-la), além dos botões "Entrar" e "Cadastre-se aqui";

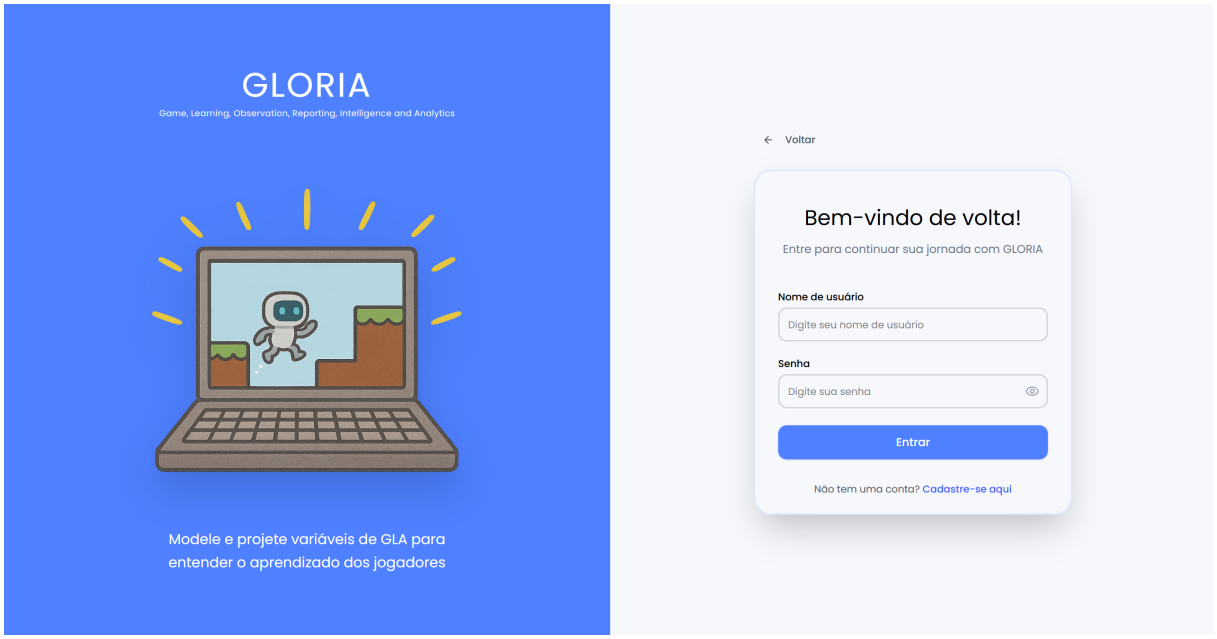


Figura 31 – *Login* da ferramenta.

- **Gestão de Jogos (Figura 32):** apresenta os jogos cadastrados em *containers*, exibindo o número, *status*, capa (tela do jogo), nome, data de criação e gênero, além dos botões para editar e excluir cada jogo. A tela também disponibiliza o botão "Novo Jogo" e uma barra de pesquisa para localizar os jogos cadastrados;

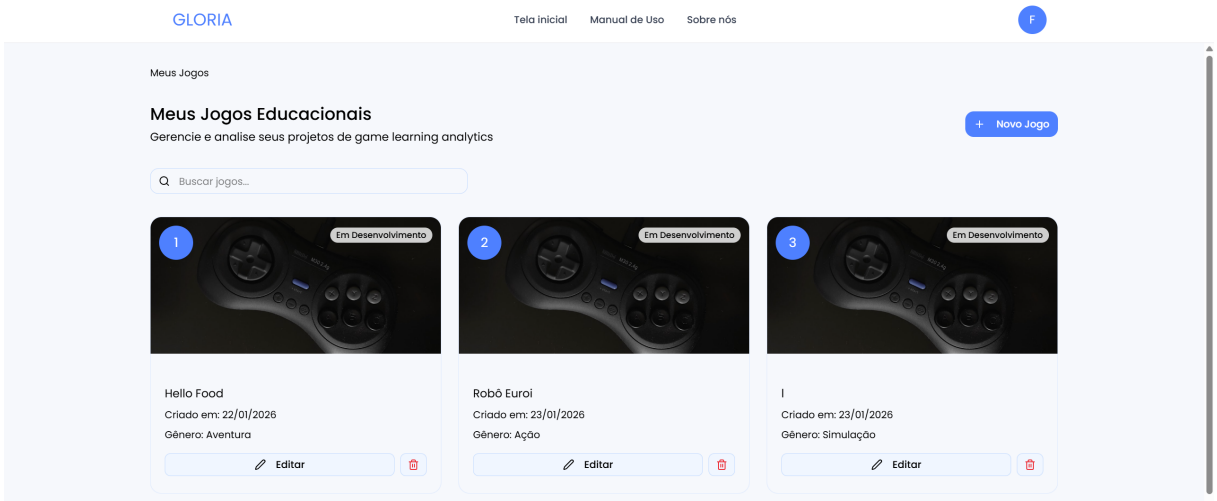


Figura 32 – Tela de "Gestão de Jogos".

- **Cadastro de Jogo (Figura 30):** permite cadastrar um jogo educacional por meio dos campos: nome, público-alvo, plataforma, gênero, história, *gameplay*, fluxo, mecânica, objetivos de aprendizagem, componente curricular (PCN, BNCC, entre outros) e *status* do projeto. Também inclui os botões "Cancelar" e "Salvar Jogo";

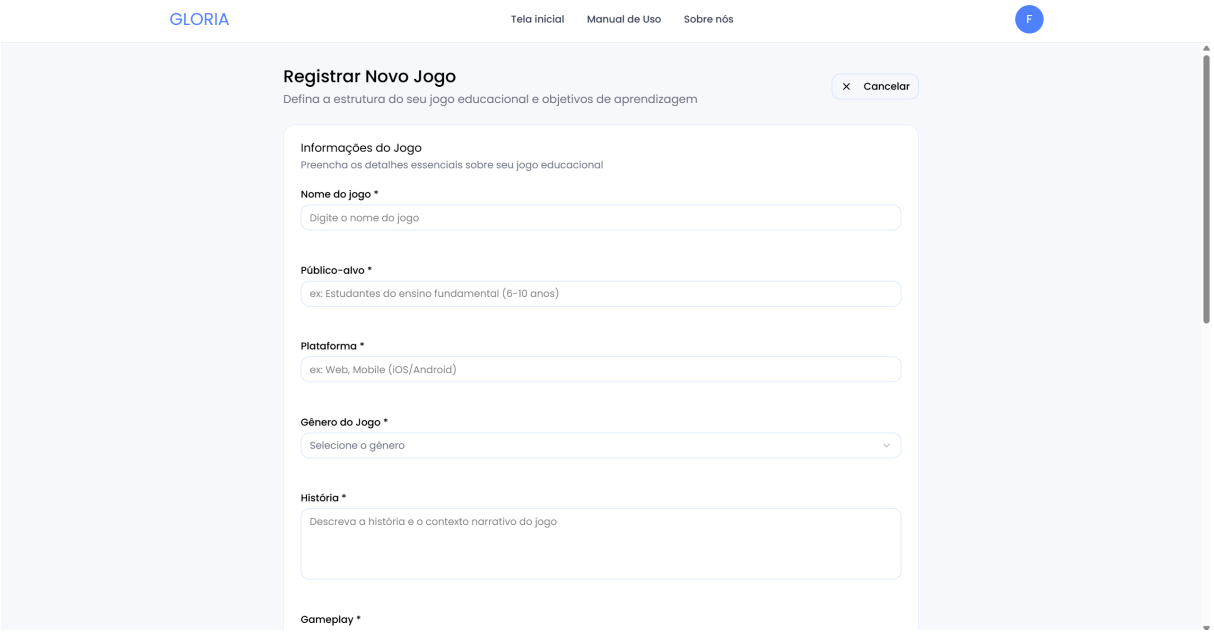


Figura 33 – Cadastro de jogo no GLORIA.

- **Jogo Específico - Dados do Jogo (Figura 34):** exibe as informações cadastradas do jogo. Inclui os botões "Editar informações", que torna os campos editáveis, e "Excluir Jogo", que abre um *pop-up* de confirmação sobre a irreversibilidade da ação e a exclusão dos dados associados;

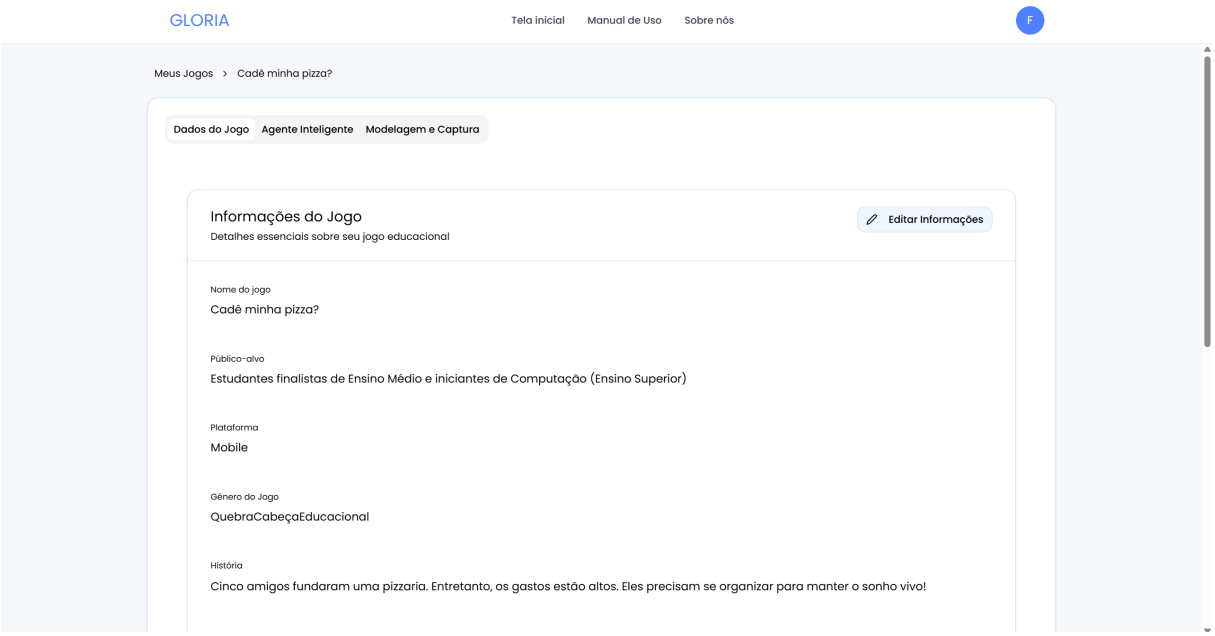


Figura 34 – Página específica de um jogo, aba de "Dados do Jogo".

- **Jogo Específico - Agente Inteligente: Primeira interação (Figura 35):** apresenta a área do *chatbot*, com duas opções de escolha ao usuário quando for sua primeira vez utilizando o GLORIA para um determinado jogo. Ele pode escolher o "Modo Tutorial", que enviará a mensagem "Olá! Sou um novo usuário aqui no GLORIA. Pode me auxiliar com os primeiros passos?" ao agente. O usuário será guiado sequencialmente por todas as etapas de GLA, até a geração do *template* de dados no modelo GLBoard; ou no "Modo Livre", caso já tenha experiência com a ferramenta e deseje usá-la de modo mais independente;

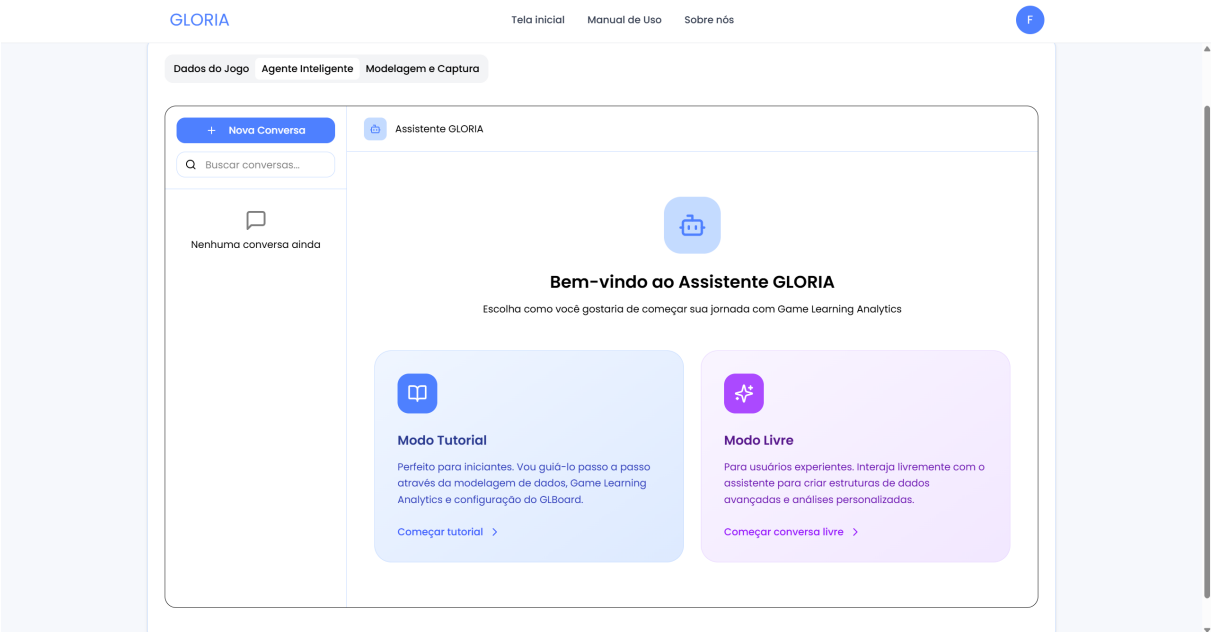


Figura 35 – Aba de "Agente Inteligente", no primeiro acesso.

- **Jogo Específico - Agente Inteligente (Figura 36):** exibe a área de interação com o agente inteligente (*chatbot*), que é integrado a um LLM. Por padrão, a tela abre na conversa mais recente. Na parte esquerda, contém: um botão de "Nova conversa", uma barra de busca e as conversas anteriores (histórico). O usuário pode criar, acessar, buscar, renomear ou excluir uma conversa. Na parte direita, é exibida a área de interação com o *chatbot*, contendo: o nome "Assistente GLORIA", as mensagens anteriores (do usuário e do agente) e um campo de envio de mensagem. Ao enviar uma mensagem, o usuário pode clicar no botão "Parar geração" para interromper a mensagem do agente;

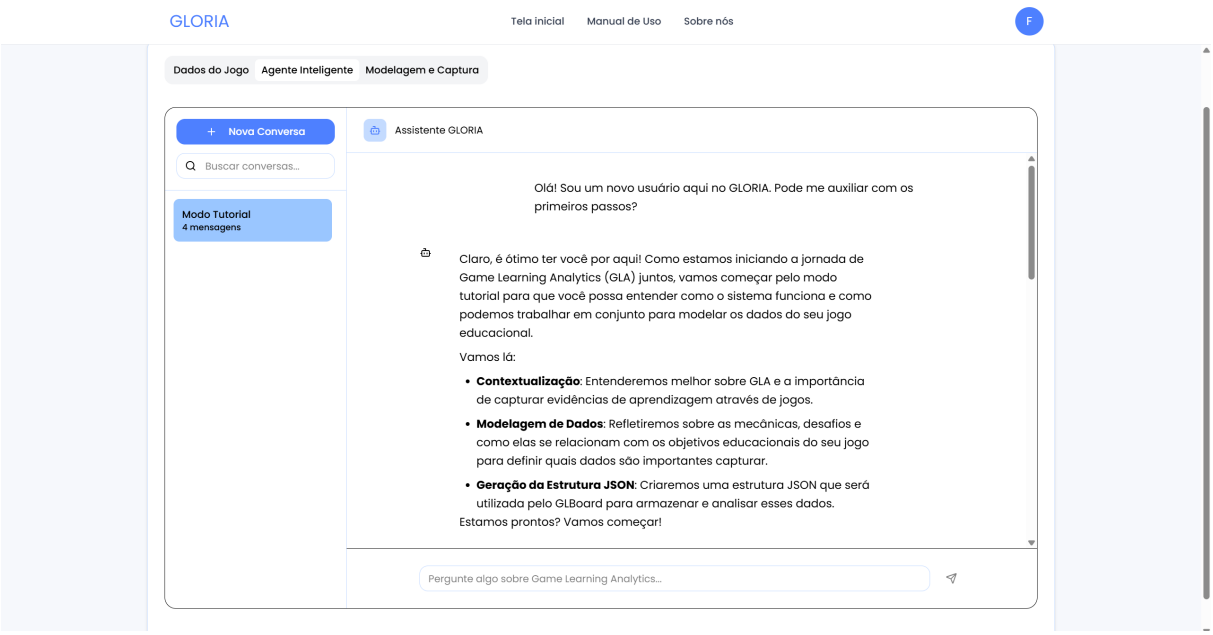


Figura 36 – Tela de interação com o agente inteligente (*chatbot*).

- **Jogo Específico - Modelagem de Dados** o primeiro subitem da aba de "Modelagem e Captura", que apresenta a "Tabela de Modelagem" (Figura 37). Ela contém colunas de id, nome da variável, tipo de dado, valor de exemplo e justificativa;

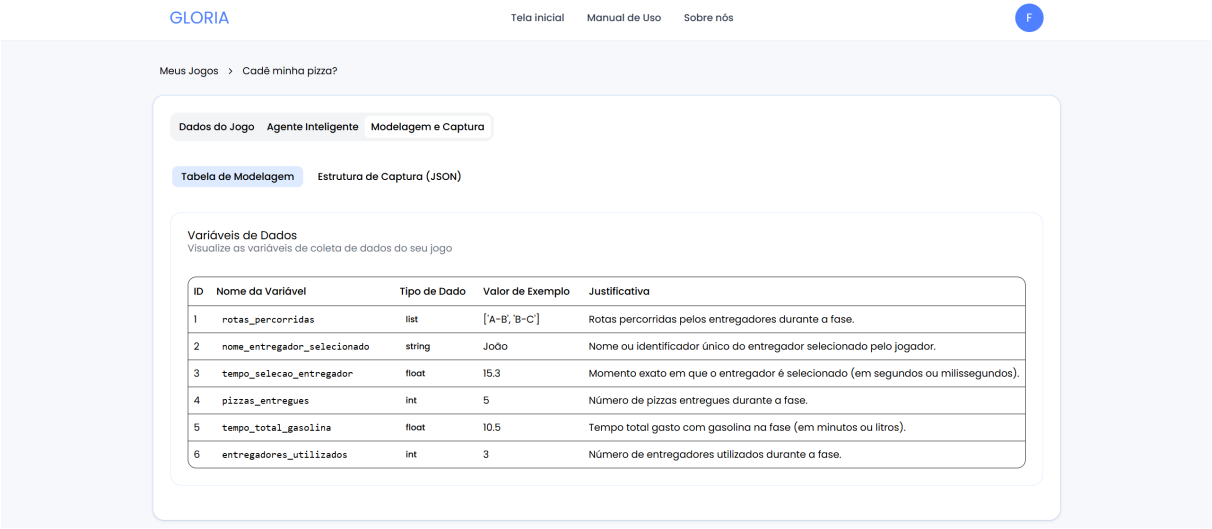


Figura 37 – Tabela de modelagem de dados do jogo específico.

- **Jogo Específico - Estrutura de Captura:** referente ao segundo subitem da aba de "Modelagem e Captura", com a "Estrutura de Captura (JSON)" (Figura 38). A

página exibe o *template* de dados do GLBoard, organizado e preenchido conforme a modelagem de dados. Inclui também um botão de "Exportar JSON".

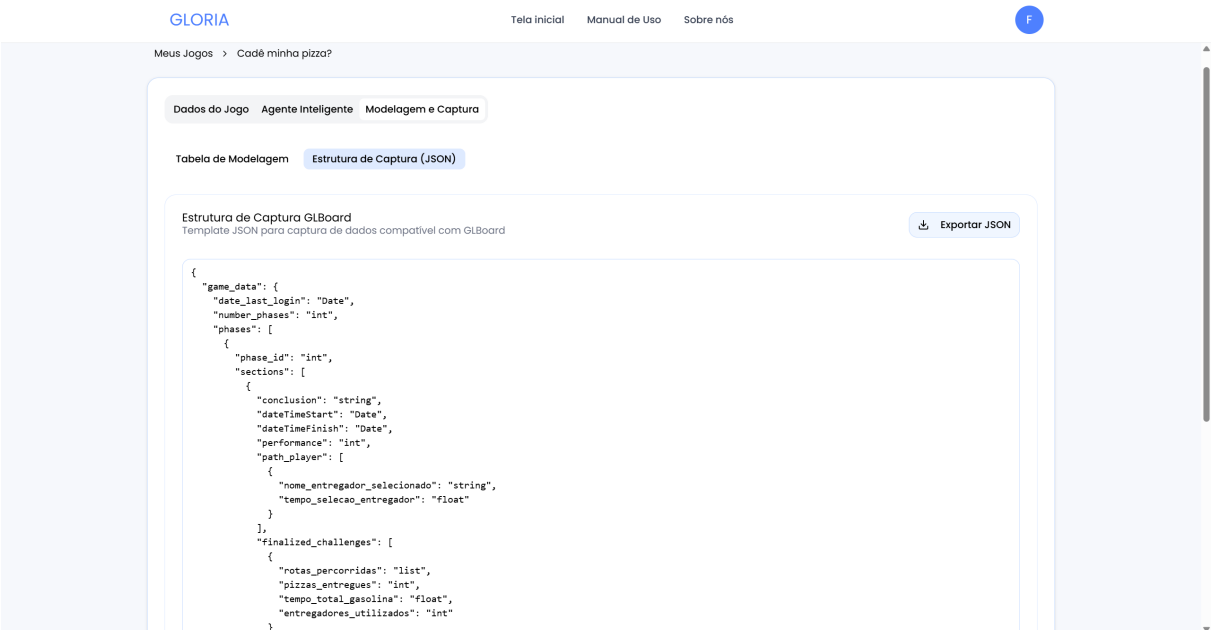


Figura 38 – Estrutura JSON gerada, conforme a modelagem.

5.3 Síntese do Capítulo

Inicialmente, este capítulo descreve a realização de um *Design* Participativo para elencar mais requisitos da ferramenta. Essa etapa foi conduzida a partir das necessidades do público-alvo: estudantes que atuavam como *designers* de aprendizagem, que propuseram e apresentaram ideias de sistemas para modelar dados de jogos educacionais. Três propostas foram elaboradas, que trouxeram contribuições futuras à ferramenta, como: inserção de recursos visuais para auxiliar na modelagem de dados, análise do código do jogo para orientar a implementação das variáveis de GLA e adaptação da estrutura de captura conforme os *logs* gerados. Além disso, por meio do DP, percebeu-se que os estudantes mobilizaram conceitos de computação para a proposta das funcionalidades, envolvendo LLMs, visão computacional e análises de códigos e *logs*.

Após o *Design* Participativo, iniciou-se o desenvolvimento da ferramenta. A metodologia de *game design* educacional (Pires et al., 2021) foi utilizada, cujas etapas

incluíram: (i) identificação do problema: localização do problema de aprendizagem, a respeito das dificuldades em modelagem de dados de jogos educacionais; (ii) pesquisa: compreensão do estado da arte, envolvendo a condução de revisões na literatura sobre GLA e LLMs, e a realização das investigações anteriores; (iii) imaginar: definição da construção de uma ferramenta integrada a um LLM, intitulada GLORIA; (iv) refletir/discutir: planejamento da ferramenta, contemplando a elaboração e documentação de seus componentes; (v) criar: construção de versões preliminares da ferramenta (protótipos); (vi) brincar/testar: validação dos protótipos por especialistas em GLA, que pontuaram melhorias e sugestões; (vi) compartilhar/brincar: condução de testes com o protótipo de alta fidelidade; (vii) refletir: análise realizadas conforme os testes, que trouxeram *insights*; e (viii) compartilhar aprendizagem formal: implementação futura dos ajustes, que consolidará a primeira versão da ferramenta no ambiente GLBoard.

Nesse aspecto, o GLORIA foi proposto: uma ferramenta integrada a um LLM para auxiliar *designers* de aprendizagem (estudantes que atuam nessa função, educadores e desenvolvedores) a modelar dados para jogos educacionais. A ferramenta foi proposta para atuar no início do fluxo de GLA, antes da coleta dos dados, contemplando as etapas de: cadastro de jogo, modelagem de dados e preenchimento do *template* de dados. O usuário pode interagir com um agente inteligente, em *chatbot*, por meio dos modos tutorial ou livre, recebendo direcionamento para definir as variáveis de GLA conforme as especificidades do seu jogo. Mediante questionamentos, o agente instiga o usuário a refletir sobre seu jogo ao invés de fornecer a resposta completa. As variáveis de GLA propostas correspondem ao percurso do jogador ($path\_player$) e aos contadores e conjuntos de objetos com os resultados alcançados, calculados pós-*gameplay* ($finalized\_challenges$). Conforme a modelagem definida com o usuário, a estrutura JSON no padrão GLBoard é construída automaticamente. Dessa forma, o usuário pode fazer o *download* da estrutura e incorporá-la ao seu jogo.

6

AVALIANDO O GLORIA PARA MODELAGEM DE DADOS EM GLA

Conforme descrito no capítulo anterior, o GLORIA foi construído: uma ferramenta para auxiliar *designers* de aprendizagem na modelagem de dados em GLA. Nesse aspecto, este capítulo apresenta a avaliação do GLORIA, incluindo: (i) uma avaliação heurística para verificar aspectos de usabilidade da ferramenta e, assim, analisar sua conformidade com os requisitos definidos; e (ii) um estudo empírico, investigando se o problema foi mitigado de forma satisfatória e se as conjecturas teóricas são válidas. O capítulo também pertence ao Ciclo de *Design* e está relacionado ao objetivo específico 4[1].

6.1 Avaliação heurística com especialistas em UX/UI

O GLORIA consiste em uma ferramenta com interface web *on-line* e integração com LLM, visando orientar *designers* de aprendizagem na modelagem de dados em GLA. Além dos objetivos de aprendizagem, é fundamental investigar se a usabilidade da ferramenta está adequada, principalmente quanto à aspectos de navegabilidade, consistência visual e qualidade do *feedback*. Nesse sentido, esta seção descreve uma avaliação heurística, um método informal de análise de usabilidade, no qual um conjunto de avaliadores examina uma interface e comenta aspectos positivos ou negativos, podendo apoiar-se

[1] Construir uma ferramenta para auxiliar *designers* de aprendizagem a modelar dados em GLA, a partir dos resultados do corpo de conhecimento, das dificuldades/estratégias localizadas e do agente de conversação.

em princípios ou heurísticas de usabilidade (Nielsen; Molich, 1990). Recomenda-se que essa avaliação seja conduzida por três a cinco avaliadores, de forma independente, buscando ampliar a identificação de problemas de usabilidade (Moran; Gordon, 2023). Desse modo, a avaliação contou com a participação de três estudantes com experiência em *design* de interface e experiência do usuário (UX/UI).

O critério de seleção foi por conveniência, priorizando pessoas com experiência em projetos em parceria com a indústria. Os participantes tinham entre 22 e 23 anos, e todos declararam alto nível de experiência em *design* de interface e experiência do usuário. Quanto ao perfil de cada participante:

- **P1:** acadêmico finalista do curso de Sistemas de Informação, possui experiência na construção de jogos educacionais, sobretudo na elaboração de ilustrações e interfaces. Atuou como *designer* de interfaces em um projeto de PD&I e integra um laboratório de tecnologias educacionais, no qual trabalha com *web design* em um sistema de gestão de jogos educacionais.

- **P2:** formação em *Design*, possui vínculo com a universidade através de projetos de PD&I, no qual atua na criação de interfaces para sistemas digitais. Além disso, participa de um laboratório de tecnologias educacionais, desempenhando atividades como ilustradora e *designer* de interfaces e experiência do usuário em sistemas relacionados a jogos educacionais.

- **P3:** graduanda em Licenciatura em Computação por uma universidade pública (6º período). Durante o Ensino Médio-Técnico, participou da construção de interfaces em seu projeto de conclusão de curso. Na universidade, há três anos participa de um laboratório de tecnologias educacionais, onde atua como ilustradora de jogos educacionais e sistemas relacionados. Além disso, exerce a função de *tech leader* (líder técnica) de *Design* em um projeto de PD&I.

Embora P1 e P3 não possuíssem formação específica em *Design*, são reconhecidos, por suas atividades remuneradas e de pesquisa, como possuidores de habilidades e competências em *design* de interface e experiência do usuário (UX/UI). Além disso,

todos os participantes possuem experiência prévia com jogos educacionais, o que contribuiu para enriquecer a avaliação.

6.1.1 Procedimentos

Nesta seção, descrevem-se os procedimentos da avaliação heurística, em forma de etapas. A condução da avaliação com os participantes selecionados deu-se de forma individual e presencial em um laboratório de pesquisa da universidade, facilitando o acompanhamento do uso da ferramenta e possibilitando observar diretamente as interações e comportamento dos testadores durante a execução das tarefas. As sessões de avaliação foram realizadas em um único dia, com duração média de 31 minutos.

Preparação do ambiente: um roteiro de avaliação foi elaborado pelo mestrando para orientar as sessões, definindo que os participantes realizariam, de forma sequencial, tarefas relacionadas às funcionalidades da ferramenta. Para a avaliação, disponibilizou-se um *notebook* e conectou-se um microfone externo para gravar os áudios pelo "Gravador de som" (ferramenta nativa do *Windows*). Em seguida, a página *web* da ferramenta foi acessada, e os participantes foram convidados individualmente a iniciar as sessões.

Introdução à ferramenta: após o consentimento para a gravação, a ferramenta foi apresentada, contemplando seu contexto de jogos educacionais, GLA e modelagem de dados. Explicou-se seu objetivo de forma geral, sem detalhes adicionais, para evitar vieses. O participante foi informado de que realizaria tarefas sequenciais para avaliar a usabilidade e aspectos de *design*. Além disso, solicitou-se que verbalizasse seus pensamentos durante a avaliação ("pensar em voz alta"), método *Think Aloud* (Someren et al., 1994), principalmente para apontar itens que lhe agradaram ou não e sugerir melhorias.

Utilização do GLORIA: o participante explorou a ferramenta, realizando as tarefas informadas pelo mestrando: (i) navegar pelas páginas iniciais; (ii) realizar o cadastro; (iii) fazer o *login*; (iv) cadastrar um jogo educacional, podendo ser real ou fictício. Nesse ponto, informou-se ao participante que poderia utilizar de artigos ou documentos de jogos para preenchimento dos campos; (v) acessar o jogo recém-cadastrado; (vi) acessar e utilizar o *chatbot* para modelagem de dados de seu jogo; (vii) visualizar a tabela de

modelagem e a estrutura de captura (JSON) respectiva; e (viii) exportar a estrutura. Durante a avaliação, cada participante pontuou inconsistências, aspectos a serem melhorados, pontos positivos e sugestões para a ferramenta. Ao fim, perguntou-se se gostaria de acrescentar algum comentário à avaliação.

Aplicação de formulário: após os comentários finais do participante, a avaliação encerrou-se. Em seguida, solicitou-se que preenchesse um formulário para avaliar a usabilidade e aspectos de *design* da ferramenta (descrito na Subseção 6.1.2).

6.1.2 Coleta e análise de dados

Nesta avaliação heurística, utilizou-se três fontes de dados: (i) observações do mestrando durante as sessões de avaliação; (ii) gravações de áudio com os comentários do *Think Aloud*, que foram transcritas e organizadas por participante; e (iii) respostas ao formulário avaliativo. Buscou-se analisar os dados em conjunto para obter *feedbacks* sobre os aspectos de usabilidade e *design* de interface do GLORIA.

O formulário avaliativo está disponível no Apêndice O, que contém seções sobre: (i) informações gerais e questão sobre consentimento de uso de dados anônimos para pesquisas; (ii) informações dos participantes (pessoais e acadêmicas); (iii) conhecimento dos participantes sobre *design* de interface, experiência do usuário e jogos educacionais; e (iv) avaliação de usabilidade, com questões em escala Likert-5 (quantitativas) e perguntas abertas sobre avaliação geral e sugestões (qualitativas). As questões quantitativas referem-se às heurísticas de Nielsen, adaptadas ao contexto do GLORIA. Essas heurísticas consistem em regras gerais de usabilidade (e não em diretrizes específicas), permitindo sua aplicação a qualquer tipo de interface (Kaplan, 2021; Nielsen, 2024).

As questões quantitativas (heurísticas) foram analisadas por meio de um gráfico *heatmap* (mapa de calor), que ilustra as avaliações dos participantes para cada questão. Quanto aos dados qualitativos, utilizou-se análise temática, cujas observações durante as sessões de avaliação, os comentários dos participantes (*Think Aloud*) e as avaliações qualitativas obtidas pelo formulário foram catalogadas em uma planilha. Esses itens foram organizados em colunas de: aspecto (descrição do item), tela (interface referida),

categoria (melhoria, inconsistência, elogio, dúvida, sugestão ou avaliação complementar), comentário, relator (identificação do participante), origem (gravação, comentário ou observação) e minuto (registro temporal, quando aplicável).

Inicialmente, cada item foi agrupado por temas recorrentes, correspondentes aos aspectos da ferramenta mencionados pelos participantes, como *chatbot*, cadastro de jogo, hierarquia visual, estrutura JSON, responsividade, entre outros. Em seguida, esses aspectos foram classificados nas categorias: (i) melhoria: funcionalidade presente na ferramenta mas que pode ser aprimorada; (ii) inconsistência: erro ou dificuldade apresentada em alguma funcionalidade já existente; (iii) elogio: pontos positivos identificados pelos participantes; (iv) sugestão: inclusão de novas funcionalidades na ferramenta; (v) dúvida: questionamentos sobre a ferramenta que envolvem aspectos futuros já mapeados; e (vi) avaliação complementar: percepção que não está diretamente relacionada às heurísticas de usabilidade, podendo envolver aspectos de modelagem de dados, geração de artefatos ou apoio à ideação.

6.1.3 Resultados

Os resultados da avaliação heurística incluem as notas dos participantes para as heurísticas de Nielsen, bem como uma análise conjunta das observações, gravações de áudio e respostas qualitativas do formulário. A Figura 39 apresenta um *heatmap* (mapa de calor) com essas notas. O Eixo Y contém os identificadores dos participantes, sendo P1 (possui experiência com a confecção de ilustrações e interfaces para jogos educacionais e sistemas), P2 (formada em *Design* e trabalha com a elaboração de interfaces) e P3 (atua como *tech leader* de *design* em um projeto em parceria com a indústria). O Eixo X representa as questões (1 a 10), referentes às heurísticas de Nielsen e correspondentes às questões 12 a 21 do formulário (Apêndice O), sendo:

- Q1 - O sistema deixa claro o que está acontecendo durante o uso (ex.: carregamentos, mudanças de etapa, respostas do agente)?
- Q2 - As terminologias, textos e instruções usados no sistema são compreensíveis e fazem sentido no contexto de jogos educacionais e GLA?

- Q3 - O sistema permite que o usuário volte atrás, cancele ações ou mude de etapa com facilidade, sem ficar travado?
- Q4 - Os elementos da interface seguem um padrão visual/comportamental consistente (botões, ícones, textos, posições)?
- Q5 - O sistema ajuda a evitar erros, oferecendo instruções claras ou evitando que o usuário siga caminhos problemáticos?
- Q6 - O sistema apresenta informações, opções e comandos de forma clara, sem exigir que o usuário memorize detalhes para continuar?
- Q7 - O sistema permite que usuários avançados realizem tarefas de forma mais rápida ou eficiente (atalhos, navegação direta, etc.)?
- Q8 - A interface é limpa, organizada e contém somente as informações necessárias, sem excesso visual ou textual?
- Q9 - Quando ocorre um erro ou problema, o sistema apresenta mensagens claras que ajudam a entender o que aconteceu e como resolver?
- Q10 - As telas, instruções e explicações fornecidas pelo sistema (incluindo o *chatbot*) ajudam o usuário a entender como realizar as atividades?

O gráfico ilustra que: Q1 (clareza do estado do sistema) e Q2 (terminologias e instruções compreensíveis) foram bem avaliadas, sugerindo que a ferramenta torna claro o que está acontecendo durante o uso e que as terminologias utilizadas são adequadas ao contexto de jogos educacionais e de GLA, respectivamente; Q3 diz respeito à possibilidade do usuário de voltar atrás, cancelar ações ou mudar de etapa sem ficar travado. Essa questão não alcançou avaliação máxima por P1 e P2, enquanto P3 a avaliou como “Baixa”, indicando que o fluxo da ferramenta ainda pode necessitar de ajustes; Q4 refere-se à padronização visual e comportamental dos elementos da interface, para a qual P1 e P2 atribuíram a nota máxima. Por outro lado, P3 a avaliou como “Médio”, sugerindo que o padrão está adequado mas que ainda há margem para melhorias; Q5 trata sobre o suporte oferecido ao usuário para evitar caminhos problemáticos, sendo

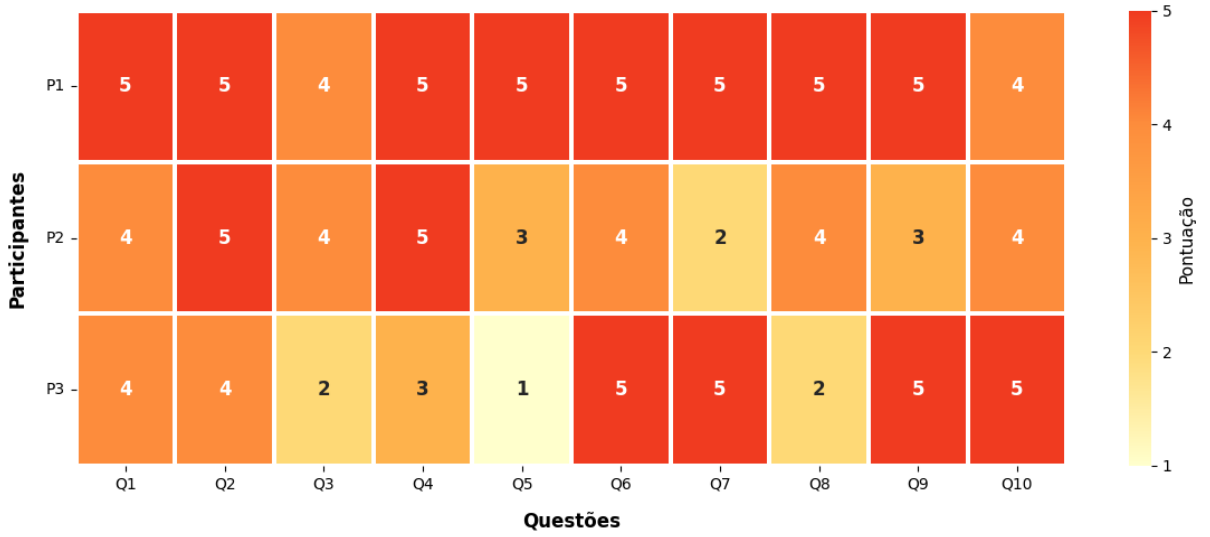


Figura 39 – *Heatmap* com as avaliações referentes às heurísticas de Nilsen.

avaliada de forma distinta por todos os participantes. Nesse sentido, a divergência nas avaliações sugere que as instruções podem ter sido percebidas como adequadas, mas insuficientes em determinadas etapas da ferramenta.

Q6 refere-se à apresentação de informações ao usuário de forma clara, evitando a necessidade de memorização. Todos os participantes concordaram quanto a esse aspecto, indicando que a ferramenta é intuitiva; Q7 trata sobre a possibilidade de usuários avançados realizarem tarefas de forma mais rápida e eficiente. P1 e P3 atribuíram notas máximas, enquanto P2 a avaliou como "Baixa", o que sugere limitações quanto a presença de atalhos ou a navegação direta entre as interfaces; Q8 está relacionada à organização e limpeza da interface, que foi bem avaliada por P1 e P2. Por outro lado, P3 a avaliou como "Baixa", indicando que algumas telas podem apresentar informações visuais ou textuais em excesso; Q9 avalia a presença de mensagens claras em situações de erro, sendo bem avaliada por P1 e P3. P2 manteve-se neutro, podendo indicar inconsistências ou a ausência dessas mensagens em alguns cenários; Q10 foi bem avaliada por todos os participantes, em relação às telas, instruções e explicações fornecidas pela ferramenta para auxiliar a compreensão de como realizar as atividades. Esse resultado também pode sugerir que a ferramenta apresenta um bom nível de intuitividade e que as orientações fornecidas são adequadas às atividades propostas.

A análise conjunta dos dados qualitativos trouxe *insights* para aprimorar a usabi-

lidade e o *design* de interface da ferramenta. Conforme mencionado na Subseção 6.1.2, esses dados estão disponíveis em uma planilha complementar[2] e foram mapeados e organizados de acordo com sua respectiva origem: observações do pesquisador, comentários de *Think Aloud* registrados nas gravações e respostas abertas do formulário. A partir desse mapeamento, foram identificadas: (i) inconsistências, como falhas no funcionamento do *chatbot* (bloqueio de rolagem e interrupção após indicação de geração do JSON), responsividade inadequada em dispositivos móveis e ausência de funcionalidades para recuperação de conta; (ii) melhorias, principalmente relacionadas à hierarquia visual e organização do conteúdo (excesso de texto, contraste elevado, ausência de elementos visuais), legibilidade e formatação do conteúdo no *chatbot* (tabela mal formatada e textos quebrados) e clareza nos campos de cadastro do jogo; e (iii) sugestões, envolvendo material de apoio para auxiliar com a implementação do GLBoard, exemplos práticos de uso da ferramenta (como vídeos), opção para seleção de data (calendário) no cadastro e fluxo de autenticação pós-cadastro.

De forma geral, a maioria dos itens pontuados pelos avaliadores refere-se à melhoria de funcionalidades já existentes, como na hierarquia visual. Poucas inconsistências foram localizadas, sugerindo que a ferramenta cumpre seu papel em usabilidade e de *design* de interface, mas que ainda pode ser aprimorada. O comentário de P1 corrobora essa afirmação: "Embora sejam muitas sugestões, a ferramenta apresenta uma excelente usabilidade, sendo muito intuitivo, além do potencial excepcional como uma ferramenta de modelagem de dados.". Outros aspectos positivos pontuados pelos participantes ("elogio" na planilha de análise) sobre a ferramenta incluem: clareza e compreensão da interface inicial, organização das informações e facilidade de uso.

Dúvidas também foram levantadas durante a avaliação, e seus esclarecimentos remetem a funcionalidades futuras, como: (i) a ferramenta será integrada ao GLBoard, não tendo como foco primário auxiliar na captura de dados utilizando outras ferramentas de GLA; (ii) a tabela de modelagem é estática, mas pretende-se possibilitar que o *designer* de aprendizagem adicione variáveis de GLA manualmente e sincronize-a de forma automática com a estrutura JSON e com o conhecimento do LLM; e (iii) ainda não há uma inteligência que automatize a implementação das variáveis de GLA na

[2] https://docs.google.com/spreadsheets/GLORIA-AvaliacaoHeuristica

game engine. Esse aspecto foi sugerido durante o *Design* Participativo (Capítulo 5), que será investigado em pesquisas posteriores.

Além disso, pela experiência prévia com jogos educacionais, os participantes também puderam contribuir com a avaliação do *chatbot* para modelagem de dados e GLA ("avaliação complementar" na planilha de análise). Por exemplo, pontuaram que há informações repetitivas nas conversas, ausência de explicação detalhada sobre o GLBoard e orientação pouco intuitiva pós-geração do JSON. Outro aspecto identificado foi a falta de suporte personalizado para usuários que não desejam interagir com o *chatbot*. Soluções futuras podem envolver a inclusão de uma mensagem para reforçar o caráter opcional das respostas e uma opção para gerar a modelagem sem interação.

Em relação à estrutura JSON, P1 destacou que "não precisa dar *ctrl + c* e *ctrl + v*" e que "implicitamente vai criando um *prompt*". Esse *insight* já havia sido identificado durante a criação do agente inteligente (Capítulo 4) e foi corroborado pelo participante: facilitar a interação com o *chatbot*, sem exigir que os usuários elaborem *prompts* complexos. P2 também relatou que o *chatbot* apresentou "sugestões interessantes sobre o que adicionar no jogo [...] fazendo ter novas ideias", indicando que o GLORIA também pode contribuir com o aprimoramento do *game design*.

6.1.4 Considerações sobre a avaliação

A avaliação heurística evidenciou boa aceitação da ferramenta pelos especialistas em UX/UI, com avaliações majoritariamente positivas para as heurísticas de Nielsen. A análise conjunta dos dados qualitativos relevou que, embora o GLORIA seja fácil de usar e intuitivo, ainda necessita de ajustes: inconsistências foram identificadas, assim como melhorias para funcionalidades já existentes. Os principais pontos destacados nas avaliações estão relacionados à estética (hierarquia visual e organização do conteúdo) e ao *chatbot* (falhas no funcionamento e legibilidade). Os participantes contribuíram com sugestões e, dada a experiência prévia com jogos educacionais, também auxiliaram na avaliação do conteúdo relacionado à modelagem de dados e GLA.

De modo geral, os participantes pontuaram um conjunto de melhorias e locali-

zaram poucas inconsistências na ferramenta. Nesse aspecto, de acordo com a DSR, os resultados sugerem que o GLORIA está em conformidade com seus requisitos, apresentando usabilidade e *design* de interface adequados. Entretanto, os achados evidenciam a necessidade de refinamentos, condizentes com o caráter iterativo do Ciclo de *Design*.

As principais limitações da avaliação incluem: (i) conveniência: os participantes pertenciam ao mesmo grupo de pesquisa do mestrando, o que pode ter introduzido vieses e favorecido avaliações mais positivas; (ii) influência do pesquisador: embora a condução presencial e individual das sessões de avaliação tenha facilitado o acompanhamento das tarefas, pode ter influenciado os participantes a agir ou avaliar a ferramenta de forma mais favorável; e (iii) amostra homogênea: a avaliação contou com três participantes que atuam na mesma área (*design* de interface e UX/UI), com faixa etária de 22 e 23 anos, o que reduz a diversidade de perfis cognitivos e limita a generalização dos resultados. Apesar das limitações, as contribuições fornecidas pelos participantes foram extensas e valiosas, auxiliando a identificar inconsistências, melhorias e sugestões.

6.2 Avaliação de suporte à modelagem de dados

Após validar o GLORIA em aspectos de usabilidade, faz-se necessário investigar se a ferramenta está atingindo seu objetivo principal: auxiliar *designers* de aprendizagem a modelar dados para jogos educacionais. Diante disso, conduziu-se um estudo empírico para avaliar se o problema foi resolvido de modo satisfatório, isto é, se a dificuldade em modelagem de dados foi mitigada. Para isso, analisaram-se as percepções dos participantes do estudo (*designers* de aprendizagem) quanto ao uso da ferramenta, sobretudo quanto à intenção de uso, apoio para modelar os dados, organização das estruturas JSON e contextualização das variáveis propostas. Além disso, buscou-se verificar se as conjecturas teóricas são válidas, isto é, se um LLM é capaz de auxiliar na modelagem de dados em GLA. Esse aspecto foi investigado a partir das avaliações das variáveis de GLA propostas pelo *chatbot*, sobretudo quanto à coerência com a mecânica e o percurso do jogador, e sua organização no *template* de dados do GLBoard.

6.2.1 Contexto e participantes

O público-alvo de GLA, cujo GLORIA é destinado, são *designers* de aprendizagem: (i) estudantes, que assumem essa função em disciplinas para construção e avaliação de jogos educacionais; (ii) desenvolvedores, os *game designers* que idealizaram e elencaram os objetivos de aprendizagem do jogo, sendo os responsáveis pela implementação das variáveis de GLA; e (iii) educadores, que irão utilizar, em sala de aula, jogos educacionais com técnicas de GLA integradas, para os quais modelaram os dados. Nesse aspecto, a seleção dos participantes baseou-se em dois critérios: (i) experiência, de pelo menos um ano, na criação de jogos educacionais; e (ii) conveniência, para facilitar a condução dos testes. A especialização em GLA não foi critério de exclusão, pois o GLORIA também visa democratizar o acesso à área e apresentá-la a usuários sem conhecimento prévio.

Portanto, quatorze pessoas participaram do estudo, com faixa etária de 21 a 31 anos, dos quais 64% (n=9) se identificavam com o gênero masculino e 36% (n=5) com o feminino. Em relação a dados acadêmicos: doze (86%) possuíam Ensino Superior completo, dos quais cinco já haviam finalizado o Mestrado. 100% dos participantes possuía formação (completa ou em andamento) em Computação, em cinco instituições públicas do país. Os participantes se consideravam professores (n=5), desenvolvedores de jogos educacionais (n=8, 57%) e estudantes (n=9).

Quanto à experiência em modelagem de dados e implementação de GLA, os participantes classificaram seu nível de conhecimento conforme a Figura 40, em uma escala Likert de cinco pontos, variando de "muito baixa" a "muito alta". Observa-se uma amostra heterogênea, composta por *designers* de aprendizagem com e sem experiência em GLA, ambos pertencentes ao público-alvo do GLORIA. Para fins de análise, os participantes foram organizados em três grupos distintos, atribuindo-se peso 3 à dimensão de modelagem de dados, considerando o foco do GLORIA, e peso 1 à experiência relacionada à implementação de GLA. Dessa forma, as notas ponderadas poderiam variar de 4 a 20. Os grupos definidos conforme o nível de conhecimento dos participantes foram: (i) Grupo A (experientes na área), com nota ponderada igual ou superior a 13 (P2, P5, P6, P12 e P14); (ii) Grupo B (experiência intermediária), com nota ponderada entre 10 e 12 (P1, P7, P9 e P13); e (iii) Grupo C (pouca experiência), com nota

ponderada igual ou inferior a 9 (P3, P4, P8, P10 e P11).

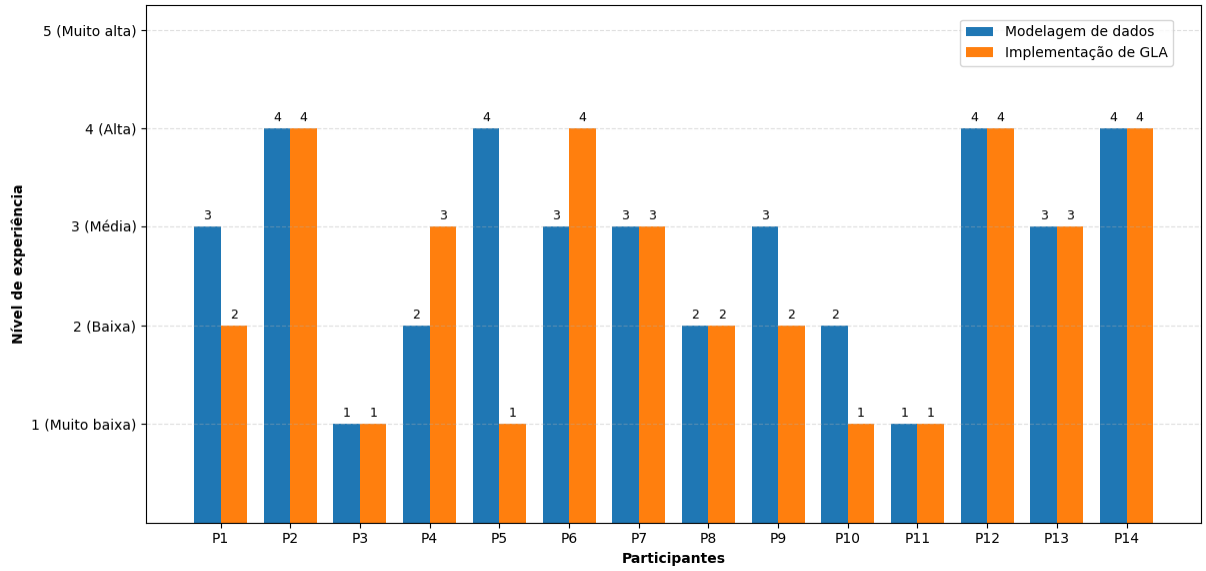


Figura 40 – Nível de experiência em modelagem de dados e implementação de GLA, por participante.

6.2.2 Procedimentos

Esta subseção descreve, em etapas, os procedimentos realizados nos testes, conduzidos em sessões: (i) individuais, facilitando o suporte aos testadores; (ii) síncronas, permitindo acompanhar o uso da ferramenta em tempo real, bem como realizar entrevistas pós-teste; e (iii) *on-line*, o que viabilizou a gravação dos testes e a participação de testadores residentes em diferentes localidades do país. Os testes tiveram duração média de 49 minutos, com tempo mínimo de 23 minutos e máximo de 1 hora e 14 minutos.

Apresentação e introdução: a etapa iniciou-se com a solicitação de autorização para gravar a reunião, visando análises posteriores. Em seguida, o mediador do teste (mestrando) apresentou-se e explicou o contexto da pesquisa: GLA, GLBoard, modelagem de dados, variáveis $path\_player$ e $finalized\_challenges$ e o GLORIA. Ressalta-se que o funcionamento da ferramenta não foi detalhado, apenas seus objetivos gerais. O mediador, então, comunicou o início do teste, no qual o participante realizaria atividades pontuais com a ferramenta. Além disso, solicitou-se que, durante a navegação,

compartilhasse a tela para acompanhamento e verbalizasse seus pensamentos ("pensar em voz alta"), método *Think Aloud* (Someren et al., 1994).

Utilização do GLORIA: em seguida, iniciou-se o teste com a ferramenta pelo participante, que realizou as tarefas sequencialmente, conforme a Tabela 9.

Tabela 9 – Atividades realizadas durante o teste com a ferramenta.

Nº	**Atividade**
1	Navegar livremente pela página inicial até julgar suficiente
2	Criar uma conta
3	Realizar o *login*
4	Cadastrar um jogo educacional, preferencialmente do qual o participante é autor ou possui experiência. Devido à extensão de alguns campos (como história, *gameplay* e mecânica de aprendizagem), foi informado que documentos complementares (como EGDD, artigo científico, entre outros) poderiam ser consultados para obter as informações necessárias
5	Acessar o jogo cadastrado e visualizar suas informações
6	Acessar a aba de "Agente Inteligente"
7	Interagir com o agente para modelar dados do jogo até optar por encerrar a interação
8	Visualizar a tabela de modelagem de dados e a estrutura de captura (JSON) na aba de "Modelagem e Captura"
9	Clicar no botão "Exportar JSON"

Condução de entrevistas: com o encerramento do teste, informou-se ao participante que alguns questionamentos sobre a ferramenta seriam realizados. Ressaltou-se que as respostas poderiam ser longas ou curtas, de preferência do testador. Diante disso, as seguintes questões foram discutidas: (i) "Quais eram suas expectativas antes de usar o GLORIA? Elas foram atendidas?"; (ii) "Você conseguiu utilizar o GLORIA? Ele te ajudou a entender modelagem de dados ou GLA?"; (iii) "Em algum momento, você acha que o agente alucinou (gerou respostas sem sentido)?"; (iv) "O que você achou do ritmo de perguntas e do tempo de resposta do agente?"; (v) "O que você achou bom do GLORIA?"; (vi) "O que você achou ruim do GLORIA?"; (vii) "O que você melhoraria no GLORIA? Tem alguma sugestão para que possamos aprimorá-lo?"; e (viii) "Você usaria o GLORIA se fosse implementar GLA em seu jogo? Por quê?".

Aplicação de formulário: após a entrevista, solicitou-se ao participante que preenchesse um formulário avaliativo sobre o GLORIA (descrito na Subseção 6.2.3) quando considerasse mais conveniente. Em seguida, o teste encerrou-se.

6.2.3 Coleta e análise dos dados

Neste estudo, utilizou-se múltiplas formas de coleta de dados: (i) observações do mestrando, por meio da análise das gravações dos testes, relacionadas a inconsistências ou melhorias; (ii) comentários obtidos via método *Think Aloud*; e (iii) respostas às entrevistas, ambos obtidos a partir da transcrição das gravações dos testes e organização dos dados por participante; e (iv) avaliação do formulário, contendo questões qualitativas em escala Likert-5 (instrumento "*Player Level Up!*", Tabela 10) e qualitativas (perguntas abertas). O formulário aplicado está disponível no Apêndice P, segmentado em: (i) descrição da pesquisa, com questão de consentimento sobre o uso de dados anônimos para fins de pesquisa; (ii) informações pessoais e educacionais; (iii) experiência a respeito de jogos educacionais e GLA; (iv) avaliação da estrutura de captura, por meio do instrumento "*Player Level Up!*"; e (v) questões abertas sobre proposta e organização das variáveis de GLA, assim como um comentário livre (opcional).

Em relação ao instrumento "*Player Level Up!*" (proposto e descrito no Capítulo 3 - Tabela 4), fez-se necessário realizar dois ajustes para alinhá-lo ao estudo: incluiu-se uma questão sobre a variável $finalized\_challenges$, visto que o GLORIA tem como objetivo propor variáveis correspondentes à ela e não somente ao $path\_player$; e refinou-se ligeiramente as demais questões, para maior esclarecimento e evitar ambiguidades. A Tabela 10 apresenta o instrumento avaliativo com os ajustes realizados, correspondente às questões 18 a 25 do formulário (Apêndice P).

Dimensão	**Nº**	**Questão**
Coerência	1	Os nomes das variáveis permitem identificar a que elas se referem?
	2	As variáveis definidas na estrutura estão claramente relacionadas à mecânica de jogo?
	3	Os tipos de dados atribuídos a cada variável são apropriados para o que se pretende capturar?
	4	As variáveis do $path\_player$ são relacionadas ao percurso do jogador?
	5	As variáveis do $finalized\_challenges$ são relacionadas a dados calculados no fim da fase?
Redundância e completude	6	Existem variáveis que capturam informações sobrepostas/duplicadas?
	7	Há variáveis importantes na estrutura que precisavam ser capturadas e não foram consideradas?
Evidência de Evolução	8	A estrutura permite análise temporal do percurso (registros de tempo)?

Tabela 10 – Instrumento de avaliação das estruturas de captura no *template* GLBoard.

Para analisar os dados quantitativos, as respostas do formulário (notas conforme o instrumento "*Player Level Up!*") foram tabuladas em uma planilha, assim como as

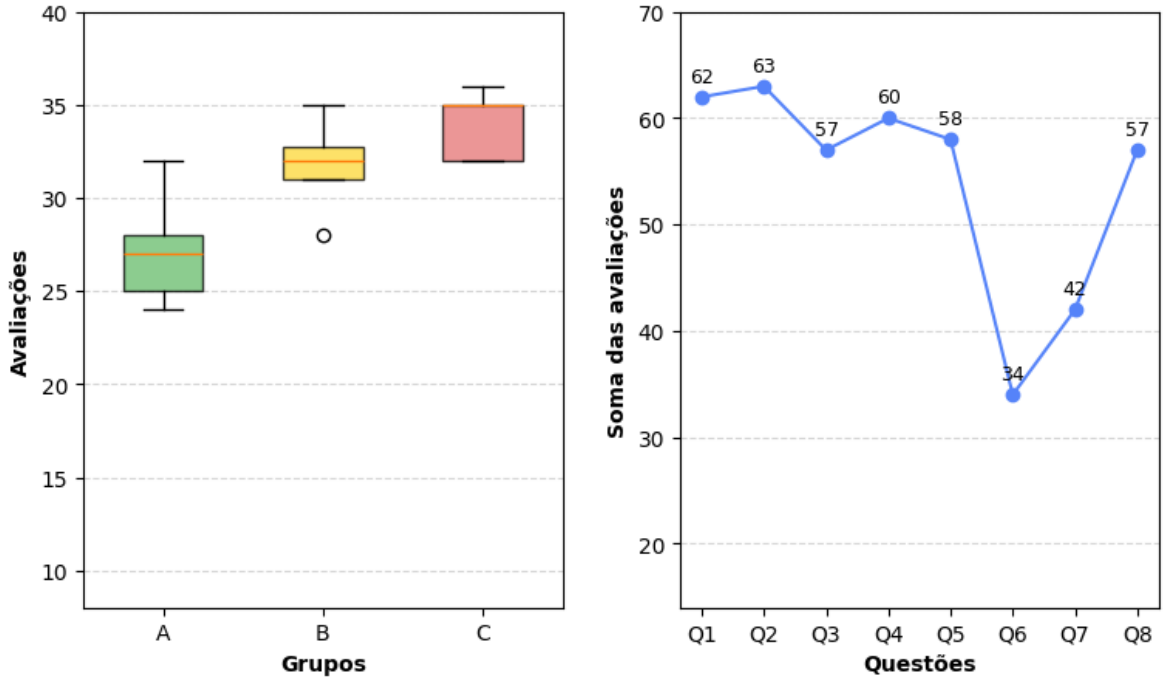
variáveis de GLA propostas durante interação dos participantes com o *chabot*, que foram mapeadas mediante análise das gravações e das estruturas JSON enviadas pelos participantes. A partir das notas dos avaliadores para as questões do instrumento, construiu-se: (i) um *boxplot* com as avaliações dos participantes para as estruturas de GLA, organizados por grupos conforme a experiência na área e (ii) um gráfico de linhas, com as somatórias das notas para as questões. Por meio das variáveis de GLA mapeadas, elaborou-se (iii) um gráfico de barras empilhadas, ilustrando as variáveis que foram incluídas nas estruturas ($path\_player$ e $finalized\_challenges$), sendo elas adequadas (usuário aprovou) ou incoerentes (não aprovadas), assim como as descartadas (excluídas ou não reforçadas pelo usuário) e desconsideradas (propostas no *chat* e aprovadas pelo usuário mas não exibidas no JSON); e (iv) uma matriz de confusão, exibindo a classificação feita pelo LLM quanto às classes de variáveis e sua acurácia.

Os dados qualitativos também foram incluídos na planilha de análise: respostas ao formulário, respostas para as entrevistas e transcrições das gravações, incluindo observações do mestrando e comentários obtidos via *Think Aloud* pelos participantes. A partir desses dados, fez-se uso de: (i) análise temática, cujas respostas dos participantes (para as entrevistas e formulários) foram analisadas, criando-se categorias conforme leitura prévia. Posteriormente, as respostas foram agrupadas conforme as categorias correspondentes, visando auxiliar na identificação de padrões e obter *insights*; e (ii) uma tabela com a síntese geral dos comentários durante os testes (*Think Aloud*) e observações das gravações, organizadas por participante.

6.2.4 Resultados

Em relação aos dados quantitativos (instrumento "*Player Level Up!*"), foram elaborados os gráficos da Figura 41. No *boxplot* da Figura 41(a), o Eixo X corresponde aos grupos e o Eixo Y às somatórias das avaliações. Os valores variam de 8 (todas as questões da estrutura JSON avaliadas com a nota mínima pelo participante) até 40 (nota máxima do participante para todas as questões). Quanto ao gráfico da Figura 41(b), o Eixo X representa as questões e o Eixo Y ilustra a somatória das avaliações por questão. As ava-

liações estão no intervalo de 14 (nota mínima atribuída por todos os participantes para a questão respectiva) a 70 (nota máxima obtida pela questão por todos os participantes).



(a) *Boxplot* dos grupos. (b) Avaliações para cada questão.

Figura 41 – Avaliações do instrumento "*Player Level Up!*".

O *boxplot* evidencia que as avaliações dos participantes dos três grupos foram superiores ao valor intermediário da escala (16), sugerindo uma avaliação positiva das estruturas de GLA geradas com o GLORIA quanto à adequação das variáveis e à organização do JSON. Os participantes do grupo C apresentaram as avaliações mais positivas, podendo ser interpretado sob duas perspectivas: (i) as estruturas foram consideradas capazes de representar o percurso do jogador no jogo, mesmo pelos participantes com menor experiência na área; ou (ii) a inexperiência pode ter limitado a identificação de inconsistências nas estruturas. O grupo B apresentou avaliações mais consistentes (menor dispersão de dados) e próximas às do grupo C, indicando boa receptividade das estruturas por participantes com conhecimento prévio em GLA. Quanto aos participantes mais experientes em GLA (grupo A), embora suas avaliações tenham sido positivas, este foi o grupo que atribuiu as menores pontuações. Esse ponto pode indicar que os participantes já possuíam expectativas mais claras de quais dados coletariam, levando-os a identificar fragilidades nas estruturas com maior facilidade.

No que diz respeito às questões Q1 a Q8, correspondentes aos itens 18 a 25 do formulário apresentado no Apêndice P: Q1 obteve uma boa pontuação (62), com 93% (n=13) dos participantes concordando que os nomes das variáveis de GLA eram intuitivos; Q2 foi a mais bem avaliada (63), indicando relação entre as variáveis e as mecânicas dos jogos (n=12, 86%); Q3 recebeu uma avaliação positiva (57), quanto à adequação dos tipos de dados. Dois participantes (14%) discordaram, sugerindo que o LLM não atribuiu tipos de dados adequados a todas as variáveis; Q4 avalia se as variáveis do $path\_player$ pertenciam ao percurso do jogador, com a qual 86% dos participantes concordaram, obtendo uma boa avaliação (60); Q5, de modo semelhante, avalia se as variáveis do $finalized\_challenges$ correspondem a dados calculados no fim da fase. Sua avaliação também foi positiva (58), dos quais doze participantes concordaram.

Q6 refere-se à presença de variáveis duplicadas. Sua pontuação baixa (34) representa um aspecto positivo: sete participantes (50%) não identificaram variáveis duplicadas. Por outro lado, quatro participantes (29%) discordaram, indicando que ainda é necessário um refinamento do agente; Q7 é similar, tratando-se de variáveis importantes que não foram sugeridas pelo *chatbot*. Seis participantes (43%) demonstraram ter sentido falta de variáveis, cuja pontuação de 42 revela um ponto de alerta, também sugerindo a necessidade de ajustes no agente; por fim, Q8 trata sobre a coleta de variáveis temporais, com uma boa avaliação (57). Nesse sentido, 83% dos participantes identificaram variáveis propostas que armazenavam aspectos de tempo do jogo.

Portanto, observa-se que, de modo geral, os participantes avaliaram positivamente as estruturas de GLA geradas com o GLORIA. Esse resultado é relevante, pois as estruturas foram construídas de formas distintas, conforme as particularidades de cada jogo educacional. Nesse sentido, as avaliações sugerem que a ferramenta foi capaz de modelar dados e gerar estruturas de GLA alinhadas aos contextos dos jogos. No entanto, ainda há espaço para melhorias, sobretudo em relação à ausência de variáveis consideradas importantes pelos participantes para a coleta de dados. Esse aspecto pode indicar: (i) a complexidade da modelagem de dados, mesmo com o apoio de um modelo de IA Gerativa; e (ii) limitações dos LLMs ao lidar com *prompts* extensos. Embora o usuário não elabore o *prompt*, as informações cadastradas sobre os jogos são encami-

nhadas ao LLM. Assim, é possível que o GLORIA tenha desconsiderado características importantes do jogo na geração das estruturas.

A Tabela 11 apresenta os participantes e seus grupos, sendo: A (experiente em GLA), B (experiência intermediária) e C (pouca experiência em GLA). Também informa o jogo ao qual modelaram dados com apoio do GLORIA, o tempo total de teste e o tempo que interagiram com o *chatbot*. Observa-se que o grupo C apresentou a maior média de interação com o *chabot* (18 minutos), seguido pelo grupo A (16 minutos) e pelo grupo B (10 minutos). Entretanto, não se pode afirmar que o nível de experiência influencia no tempo de interação, visto que foi distinto para cada participante, sugerindo que depende das preferências e estratégias individuais dos usuários. É possível observar que a maioria dos participantes (n=11, 79%) optou por usar jogos autorais, o que favorece a modelagem de dados e avaliação do impacto da ferramenta, pois pressupõe que já compreendem suas mecânicas e objetivos de aprendizagem. As informações sobre os jogos educacionais utilizados no estudo são apresentadas no Apêndice Q.

Tabela 11 – Grupos de participantes e duração dos testes.

ID	**Grupo**	**Jogo**	**Tempo total**	**Tempo (*chatbot*)**
P1	B	WAlgor (autoral e publicado)	34 min	4 min
P2	A	OptimaCorp (autoral e publicado)	1h14 min	28 min
P3	C	Animal Park (autoral e publicado)	1h08 min	21 min
P4	C	Matematicamente (fictício)	54 min	19 min
P5	A	Two Cats (fictício)	47 min	16 min
P6	A	ODC (não autoral e publicado)	34 min	8 min
P7	B	RoboHouse (autoral e publicado)	31 min	9 min
P8	C	MyName (autoral e publicado)	1h02 min	33 min
P9	B	Pensar e Lavar (autoral e publicado)	55 min	21 min
P10	C	Código Turing (autoral e publicado)	1h	12 min
P11	C	EscalaQuest (autoral e publiado)	57 min	9 min
P12	A	Robô Euroi (autoral e publicado)	41 min	15 min
P13	B	StarDust (autoral e publicado)	23 min	6 min
P14	A	Hello Food (autoral e publicado)	52 min	16 min

No que diz respeito à interação dos participantes com o *chatbot*, o gráfico de barras empilhadas da Figura 42 exibe as variáveis de GLA propostas pelo GLORIA junto aos participantes. Essas variáveis são organizadas em: (i) adequadas: inseridas no JSON que o participante explicitamente aprovou, considerou adequadas de forma implícita ou aprovou parcialmente com ressalvas[3]; (ii) incoerentes: inseridas no JSON,

[3] Identificou algum ponto de melhoria mas decidiu prosseguir com a variável da forma que estava ou fez pequenas alterações, como o nome, o tipo de dado ou a organização no JSON.

ainda que o participante as tenha considerado inadequadas ou não tenha apresentado uma avaliação a respeito; (iii) desconsideradas: não foram inseridas no JSON final, embora houvesse indícios explícitos ou implícitos de aprovação pelo participante; e (iv) descartadas: propostas pelo *chatbot*, mas não inseridas no JSON, pois o participante não as julgou importantes de forma explícita nem reforçou sua inclusão.

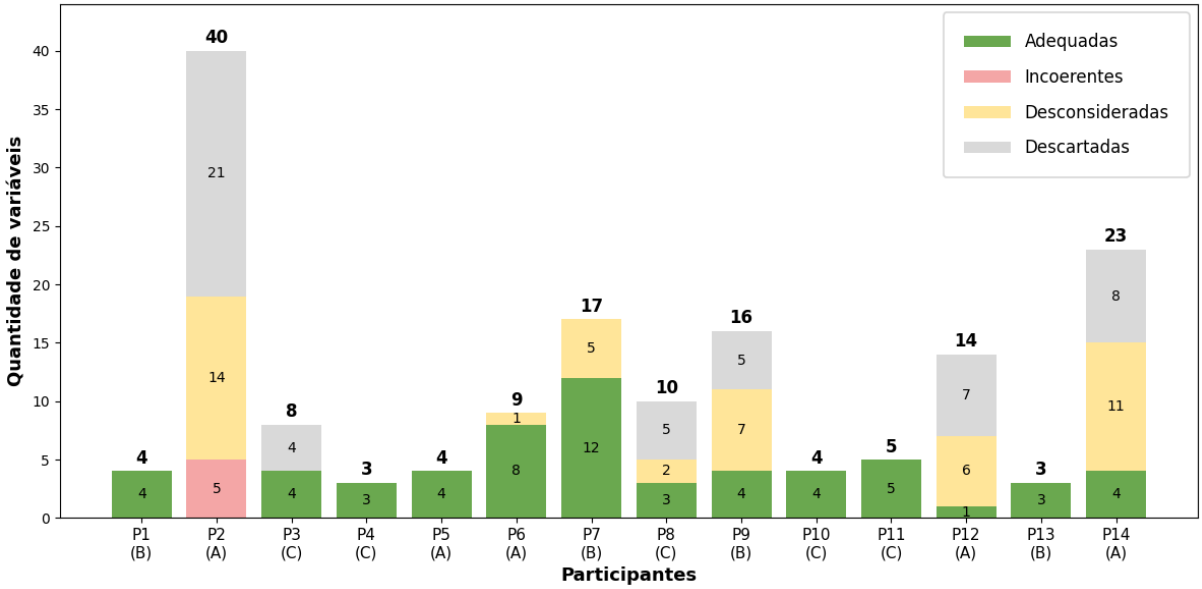


Figura 42 – Variáveis de GLA propostas com o GLORIA.

No total, 160 variáveis foram propostas durante as interações dos participantes com o GLORIA, sugerindo uma boa capacidade da ferramenta em modelar dados conforme as especificidades dos jogos. Ao analisar essas variáveis por grupo, observa-se que 90 (56%) pertencem às interações dos participantes do grupo A. Além disso, embora o grupo B seja composto por apenas quatro participantes, propôs 40 (25%) variáveis, superando o grupo C que foi responsável por somente 30 (19%) das variáveis geradas. Esse resultado sugere que a experiência prévia em GLA impactou numa geração maior de variáveis com o *chatbot*. Ressalta-se que não foi observada uma relação direta entre o tempo e a geração das variáveis: em média, o grupo C apresentou maior interação com o *chatbot* mas foi o que menos propôs variáveis. P8 e P2 foram os participantes com mais tempo de interação, respectivamente 33 e 28 minutos. No entanto, P8 propôs somente 10 variáveis, enquanto P2 propôs 40, num tempo menor de 5 minutos.

De todas as variáveis geradas junto ao *chatbot*, 64 (40%) foram incorporadas à

estrutura JSON resultante, das quais 59 (92%) foram classificadas como adequadas. Esse aspecto indica que as variáveis no JSON estavam alinhadas com as preferências dos participantes. Essa consistência também é observada nas 50 variáveis descartadas, que não foram inseridas no JSON, pois os participantes não as consideraram explicitamente necessárias para a modelagem nem reforçaram sua inclusão junto ao *chatbot*. Em contrapartida, a existência dessas variáveis também pode representar uma limitação do agente na proposição de variáveis adequadas às necessidades de modelagem de cada jogo. Além disso, observa-se que 46 variáveis foram desconsideradas, isto é, consideradas adequadas pelos participantes, mas não incorporadas ao JSON, sobretudo por P2, P7, P9, P12 e P14, que possuíam experiência prévia com GLA.

Esse resultado também corrobora o ponto de atenção para a Q7 do instrumento "*Player Level Up*", referente à variáveis importantes não consideradas. Nesse sentido, participantes com mais conhecimento na área apresentaram mais dificuldades com o GLORIA. Essa afirmação também se confirma ao analisar as gravações: observou-se uma inconsistência no *chatbot*, que tendia a "esquecer" as variáveis definidas na conversa, especialmente quando o usuário não as confirmava ou discutia sobre outros pontos. Por exemplo, em P2, nenhuma das variáveis cuja aprovação foi identificada explícita ou implicitamente foi incluída no JSON. Por outro lado, participantes com menos experiência na área (como P3, P4, P10 e P11) obtiveram resultados mais consistentes com o GLORIA, uma vez que todas as variáveis propostas durante a interação com o *chatbot* foram consideradas como adequadas e incorporadas à estrutura JSON.

Outro ponto de atenção identificado refere-se às variáveis duplicadas: embora apresentassem nomes distintos, já haviam sido mapeadas em outras variáveis, sendo elas já existentes no *template* GLBoard ou propostas pelo próprio *chatbot*. 40 variáveis duplicadas foram identificadas, o que representa 25% do total de propostas. Dentre elas, 11 foram incorporadas às estruturas JSON. Esse aspecto não ressalta somente o "esquecimento" do *chatbot* mas também identifica outra inconsistência: o *template* do GLBoard admite apenas variáveis para as classes $path\_player$ e $finalized\_challenges$[4]. Entretanto, embora o LLM tivesse conhecimento teórico sobre o GLBoard e fosse orientado para propor variáveis somente a essas classes, não tinha acesso direto ao *template* de dados

[4] Nesta seção, ambas são denominadas classes para distingui-las das variáveis que agrupam.

de GLBoard. Essa limitação ocasionou a proposta de 23 variáveis que já existiam no *template* do GLBoard, das quais 10 foram incluídas no JSON. 29% dos participantes (n=4) não avaliou positivamente a questão sobre variáveis duplicadas no instrumento "*Player Level Up!*", o que pode indicar que identificaram essa inconsistência.

Visando analisar a classificação das variáveis de GLA propostas, elaborou-se a matriz de confusão da Tabela 12. Ela contempla as classes reais das variáveis ($path\_player$ ou $finalized\_challenges$), a classe prevista pelo LLM na estrutura JSON e a respectiva acurácia por classe, calculada pela proporção de variáveis corretamente classificadas em relação ao total de variáveis pertencentes a cada classe. Ressalta-se que foram analisadas somente as variáveis efetivamente incluídas na estrutura, incluindo as duplicadas, visto que nem todas as variáveis elencadas durante a interação com o *chatbot* continham a classe de destino definida.

Tabela 12 – Matriz de confusão da classificação das variáveis

Classe Real	$path\_player$	$finalized\_challenges$	**Acurácia**
$path\_player$	36	11	77%
$finalized\_challenges$	3	4	57%

Observa-se que o GLORIA classificou corretamente 77% das variáveis pertencentes ao $path_player$, indicando maior precisão na atribuição de dados do percurso do jogador a essa classe. Entretanto, sua acurácia de 57% para a classe $finalized_challenges$ ilustra uma dificuldade do GLORIA em atribuir corretamente a essa classe variáveis agregadas, como contadores, ou relacionadas aos resultados alcançados na fase. Esse ponto passou despercebido pelos participantes durante a avaliação das estruturas com o instrumento "*Player Level Up!*": na Q5, 86% (n=12) concordaram que as variáveis dessa classe estavam relacionadas a dados calculados após a fase.

As opiniões dos participantes sobre as variáveis de GLA (questão 26 - Apêndice P) foram, principalmente: (i) positivas (n=10, 71%), destacando que elas estavam alinhadas ao contexto dos jogos e que forneceram uma base para a modelagem. Além disso, P5 ressaltou que elas contemplam informações necessárias para "fazer uma análise robusta sobre o progresso do jogador", o que está alinhado ao objetivo do GLORIA: modelar dados que permitam, em etapa futura de análise de *logs*, identificar evidências

de aprendizagem; (ii) positivas, mas com ressalvas (n=2, 14%), especialmente ao *chatbot*, que "acertou na geração de variáveis, mas não definiu variáveis importantes" (P2). P14 descreveu uma situação semelhante, ao informar que os "nomes são fáceis de entender, mas algumas precisavam ser adicionadas". Esses comentários evidenciam a limitação do LLM descrita anteriormente, que ocasionalmente "esquecia" variáveis propostas na conversa; e (iii) ressalvas (n=2), especialmente no caso de P12, em que apenas uma variável foi incorporada à estrutura, também associada à limitação de "esquecimento". Quanto ao JSON (questão 27 - Apêndice P)), os participantes avaliaram positivamente a organização das variáveis (n=11, 79%), descrevendo-a como compreensível e bem estruturada. Embora, no caso de P13, o GLORIA tenha trocado variáveis entre $path\_player$ e $finalized\_challenges$, a distinção entre ambas foi apontada como aspecto positivo (P6).

Quanto às entrevistas, os itens a seguir sintetizam as respostas dos participantes. Os enunciados completos das questões podem ser consultados na Subseção 6.2.2.

- **Q1 - Expectativas:** de forma geral, a maioria dos participantes (n=11, 79%) concordou que as expectativas foram atendidas: auxiliar na modelagem de dados (n=6, 43%) e gerar uma estrutura JSON (n=4, 29%). Em contrapartida, P2 discordou, informando que a ferramenta não priorizou o que era mais importante para a coleta. Já P8 não descreveu explicitamente suas expectativas, cujo foco com uso do GLORIA era mais técnico. P10 confirmou parcialmente, indicando não ter conhecimento o suficiente do GLBoard para afirmar com exatidão.

- **Q2 - Uso do GLORIA e compreensão sobre modelagem e GLA:** conforme a análise das respostas e das gravações, nota-se que todos os participantes conseguiram utilizar o GLORIA. Em relação à modelagem de dados e GLA, nove (64%) explicitamente informaram que a ferramenta auxiliou a entender esses conceitos. Por outro lado, cinco participantes (36%) utilizaram a ferramenta com um caráter mais técnico, voltado à definição das variáveis de GLA. Pelo conhecimento prévio na área, não utilizaram necessariamente o GLORIA para aprofundamento teórico.

- **Q3 - Ocorrência de alucinação:** embora a ferramenta tenha apresentação inconsistências (como repetição de informações, questionamento de dados já fornecidos e

esquecimento de variáveis definidas), 64% (n=9) dos participantes não considerou que houve alucinação. Em contrapartida, três participantes (21%) apontaram que o LLM alucinou, mas associaram à falta de detalhamento ou especifidade das informações do jogo. P2 e P12, por sua vez, atribuíram a alucinação ao comportamento do LLM, especialmente por propor dados na conversa, mas não os adicionar à tabela de modelagem ou à estrutura JSON.

- **Q4 - Ritmo de perguntas e tempo de resposta:** todos os participantes (100%) acharam que o tempo de resposta do LLM foi bom, embora com algumas ressalvas: geração da tabela de modelagem e do JSON mais demorada (P7), ocorrência de mensagem nula (P8) e ausência de *feedback* quando o *chatbot* "parava" (P11). Em relação ao ritmo de perguntas, dez participantes (71%) consideram-no como adequado, com destaque a resposta de P7: "não chega definindo as variáveis ou dando as respostas. Faz as perguntas e quebra a modelagem em etapas". Esse comentário reflete um dos objetivos do GLORIA: manter o protagonismo do *designer* de aprendizagem e o incentivar a refletir sobre seu jogo. Por outro lado, também identificou-se pontos de melhoria: falta de sequenciamento nas perguntas (P3), alta quantidade de questionamentos (P5), textos longos (P2, P9), repetição de informações (P10) e ausência de orientação no modo livre (P12, P14).

- **Q5 - Pontos positivos:** os principais aspectos elogiados sobre o GLORIA referem-se a: (i) facilidade de uso (n=5, 36%), sendo apontado como simples e intuitivo; (ii) apoio à modelagem de dados (n=4, 29%), caracterizado como bem explicativo, com sugestões pertinentes que podem auxiliar pessoas iniciantes na área; (iii) praticidade (n=2, 14%), sobretudo ao compará-lo com outros LLMs gerais; e (iv) velocidade das respostas (n=2). Além disso, os participantes destacaram a flexibilidade em adaptar a estrutura ao padrão fornecido (P8), a exportação do JSON (P7) e a inovação da proposta em ser uma IA (P10).

- **Q6 - Pontos negativos:** a principal consideração negativa diz respeito à clareza do fluxo e navegação (n=5, 36%), visto que não estava evidente que a aba de "Modelagem e Captura" deveria ser consultada para visualizar a estrutura JSON. Esse ponto também está associado a uma inconsistência do *chatbot*, que não a

estava exibindo na conversa. Inconsistências técnicas também foram identificadas (n=4, 29%), como: opção de pular linha no *chat* não presente (P5), ausência de botão de "editar conversa" (P10), falta de *feedback* no uso do botão de "parar geração" (P11), ocorrência de mensagem vazia enviada pelo *chatbot* (P8) e bloqueio de rolagem no *chatbot* (P5), também localizado na avaliação heurística (Seção 6.1).

- **Q7 - Sugestões de melhoria:** foram elencadas diversas contribuições, tais como: LLM multimodal (P6, P14), campos mais claros no cadastro do jogo (P1), modo escuro (P7), refinamento do *chat* para que não se assemelhe tanto a um "*bot*" (P10), entre outros. Sugestões de usabilidade e *design* também foram pontuadas (n=4, 29%), como: botões do *chat* mais interativos (P11), *dropdown* para seleção da plataforma no cadastro do jogo (P13), atalho no *chat* para acessar a aba de "Modelagem e Captura" (P10, P13), aumento do tamanho da interface do *chat* (P7) e redução de destaque ao botão "excluir jogo" (P10). Quanto às sugestões de GLA (n=4), destacam-se: adição de opção para definir o foco da modelagem, à serem sugeridas pelo *chatbot* ou criadas do zero em conjunto (P2), inclusão de mais informações sobre as variáveis do *template* do GLBoard (P14) e segmentação da tela para acompanhar o *chat* e a estrutura JSON em paralelo (P12).

- **Q8 - Uso do GLORIA para GLA:** as respostas dos participantes foram unânimes (n=14, 100%) quanto à utilização do GLORIA para auxiliar na implementação de GLA. As justificativas incluem, principalmente: (i) apoio à modelagem dos dados (n=7, 50%), com sugestões de variáveis, ideias e estímulo à reflexão; (ii) especialidade na área (n=3, 21%), principalmente comparando-o com LLMs genéricos; (iii) acessibilidade para pessoas iniciantes (n=2, 14%); e (iv) agilidade (n=2), especialmente pela economia de tempo no processo de modelagem.

Além das contribuições mapeadas com as questões do formulário e respostas às entrevistas, outros pontos de melhoria foram identificados a partir de observações das gravações e dos comentários dos participantes durante o teste (*Think Aloud*). Além de aprimoramentos no *design* e usabilidade, os itens contemplam refinamentos na área do *chatbot* e no LLM, conforme apresentado na Tabela 13.

Tabela 13 – Evidências observadas na interação com o GLORIA (*Think Aloud*).

ID	**Evidências observadas na interação**
P1	*Chatbot* repetiu mensagens idênticas. Não está claro que "Modo Tutorial" e "Modo Livre" são botões. Deve-se remover o botão "Nova mensagem" quando for a primeira interação. Usuário mencionou uma avaliação de performance e sugeriu uma estratégia para auxiliar na economia de *tokens* do LLM
P2	Sugeriu incluir informações do GLBoard na tela inicial. Limitação identificada: estrutura JSON não exibida no *chat* mesmo com insistência do participante
P3	A marcação de campos incorretos no cadastro está na cor preta, devendo ser vermelha. O usuário sugeriu incluir opção para selecionar mais de um gênero de jogo
P4	Embora o *chatbot* tenha orientado a como implementar a estrutura de captura na Unity, os códigos propostos não consideraram o GLBoard. Esse ponto está relacionado à inconsistência identificada anteriormente sobre acesso ao *template* de dados
P5	Sutil alucinação do LLM: trocou o termo "gravetos" por "gravo". A tabela de modelagem não foi gerada no *chat*: o *chatbot* utilizou um padrão com tópicos
P6	Usuário não havia notado a barra superior na tela inicial. O participante optou por usar IA Gerativa pra auxiliar a cadastrar o jogo, indicando que incorporar o LLM à essa etapa pode ser uma alternativa promissora
P7	Mesmo no "Modo Tutorial", o *chatbot* perguntou ao usuário qual modo preferia
P8	Mensagem de erro não formatada ao inserir um ano de nascimento superior ao atual (cadastro). Alucinação do LLM, que gerou uma palavra em outro idioma (supostamente em tailandês) e apresentou os argumentos da chamada à *tool* responsável para geração da tabela de modelagem. Identificou-se também uma inconsistência: mesmo que listas de variáveis sejam propostas no *chat*, seus elementos internos são inseridos na raiz de $path\_player$ ou $finalized\_challenges$ no JSON
P9	Sugeriu que, ao clicar no nome "GLORIA", direcionasse-o para a tela inicial. O participante comentou que o *chatbot* realizou um questionamento que já havia sido respondido. Além disso, achou interessante que a ferramenta sugeriu uma ideia que não havia pensado, o que ressalta seu apoio ao *game design*
P10	Notou-se que não há um botão de "ver senha" no cadastro. Erro do servidor local que hospeda o LLM, causando instabilidade na ferramenta. Por conta disso, deu-se continuidade ao teste em outra ocasião. Percebeu-se que pode ser interessante a opção de "responder" a uma mensagem no *chat*
P11	O *chatbot* solicitou informações do jogo, mesmo esses já tendo sido cadastrados. Quando uma informação foi disponibilizada, ele "relembrou" parcialmente
P12	Recomendou incluir mais gêneros de jogos e o ajuste da cor do campo para simbolizar que está selecionado (cadastro de jogo). Usuário realizou um questionamento, mas o LLM interpretou como uma solicitação de ajuste e atualizou a modelagem
P13	Achou uma quantidade elevada de campos no cadastro de jogo. Futuramente, pretende-se incluir uma opção para anexar arquivo e utilizar técnicas de RAG para recuperar as informações automaticamente
P14	Sugestões: exibir o nome completo de GLA ao lado da sigla e explicar sobre o GLBoard (tela inicial); incluir *containers* no manual de uso; marcar campos incorretos no *login*; permitir cadastrar opções de público-alvo e incluir seleção múltipla (cadastro de jogo); e, no *chatbot*, destacar as fases em negrito. Uma mensagem de erro foi exibida sem formatação ao ultrapassar o limite de caracteres da história do jogo. O usuário sugeriu manter esse limite, adicionando um contador e um *scroll* interno

De modo geral, o GLORIA foi avaliado positivamente pelos participantes, como

evidenciado pela intenção de uso da ferramenta para apoiar a implementação de GLA (Q8 - entrevista), indicada por todos os participantes (n=14, 100%), sobretudo pelo apoio à modelagem de dados (n=7, 50%). Esse resultado também é reforçado pelas notas atribuídas às estruturas de GLA, ilustradas no *boxplot* da Figura 41(a), que foram, em sua maioria, superiores ao valor intermediário da escala (24). Por outro lado, ressalvas foram identificadas, sobretudo em relação à interação com o *chatbot*. A versão posterior da ferramenta irá incorporar a correção das inconsistências, que incluem: (i) o "esquecimento" das variáveis propostas no *chat* e a ausência de orientação no "Modo Livre", aspectos que contribuíram para que usuários mais experientes em GLA apresentassem maior dificuldade de uso; (ii) o comportamento do *chatbot*, que eventualmente exibia informações internas ao usuário (chamada da *tool* e classe das variáveis), bem como gerava a modelagem sem confirmação prévia ou em um formato que destoava do padrão (tabela). Além disso, ocasionalmente não apresentava o JSON na conversa; e (iii) o acesso limitado do *chatbot* ao GLBoard, apesar de compreender seu funcionamento geral e as variáveis $path\_player$ e $finalized\_challenges$. Essa limitação resultou na proposta de variáveis que já estavam presentes no *template* do GLBoard e na sugestão de tipos de dados que não correspondem à linguagem C# (utilizada pela *game engine Unity*), como *dict* ou *list* sem os parâmetros genéricos obrigatórios. Também se observou que a justificativa para a coleta de determinados dados se aproximava mais de uma descrição das variáveis de GLA do que de uma explicação sobre sua relevância para a identificação de evidências de aprendizagem.

6.2.5 Considerações sobre o estudo

Um estudo empírico foi conduzido com *designers* de aprendizagem (iniciantes e experientes na área), com o objetivo de verificar se o problema foi resolvido de modo satisfatório e se as conjecturas teóricas são válidas (ambos conforme a DSR). Os participantes utilizaram o GLORIA, consentiram com a gravação do teste, pontuaram comentários via *Think Aloud*, responderam à entrevistas e preencheram um formulário para avaliar as estruturas de GLA geradas junto à ferramenta.

O problema investigado era a dificuldade de *designers* de aprendizagem em modelar dados para GLA. Evidências de sua mitigação incluem: (i) 100% usaria o GLORIA para implementar GLA (Q8 - entrevista); (ii) 79% apontou que as expectativas foram atendidas (Q1 - entrevista); (iii) 64% afirmou que a ferramenta ajudou a entender modelagem de dados e GLA (Q2 - entrevista); (iv) as estruturas JSON foram avaliadas positivamente, conforme o *boxplot* da Figura 41(a) ("*Player Level Up!*", questões 18 a 25 do formulário - Apêndice P); (v) 71% dos participantes apresentaram comentários favoráveis sobre as variáveis de GLA propostas, destacando sua relação com o contexto dos jogos e o apoio à definição dos dados a coletar (Q26 - formulário); e (vi) 79% dos participantes avaliaram positivamente a organização do JSON, destacando a diversidade das variáveis e sua disposição nos campos correspondentes (Q27 - formulário).

A conjectura teórica de que os LLMs podem apoiar a modelagem de dados para GLA é sustentada pelas seguintes evidências: (i) 86% (n=12) concorda que as variáveis de GLA na estrutura JSON são coerentes com a mecânica dos jogos (Q19 - formulário); (ii) 86% afirmou que as variáveis pertencem ao percurso do jogador (Q21 - formulário), o que também se confirma pela acurácia de 77% do GLORIA em organizar os dados no $path\_player$; (iii) os nomes das variáveis (Q18 - formulário) e seus tipos (Q20 - formulário) foram avaliados positivamente por 93% (n=13) e 86% dos participantes, respectivamente; (iv) 64% (n=9) apontou que a ferramenta auxiliou a compreender modelagem de dados e GLA (Q2 - entrevista); (v) das 64 variáveis incorporadas às estruturas JSON, 59 (92%) foram classificadas como adequadas (aprovadas explícita ou implicitamente, ainda que com ressalvas); (vi) das 54 variáveis da matriz de confusão, 74% foram classificadas corretamente; e (vii) os participantes destacaram como pontos positivos (Q5 - entrevista) a facilidade de uso (n=5, 36%), o apoio à modelagem de dados (n=4, 29%), a praticidade (n=2, 14%) e a velocidade das respostas (n=2).

Entretanto, foram identificadas ressalvas tanto em relação à mitigação do problema quanto à sustentação da conjectura teórica: (i) 43% dos participantes sentiram falta de variáveis (Q24 - formulário); (ii) 46 variáveis propostas no GLORIA foram classificadas como desconsideradas, isto é, não foram incorporadas ao JSON, embora tenham sido identificados indícios explícitos ou implícitos de aprovação pelos participantes.

Essa limitação pode ter relação com uma inconsistência identificada no *chatbot*, que "esquecia" as variáveis não confirmadas na conversa, principalmente com os participantes com experiência prévia em GLA; (iii) 50 variáveis propostas foram descartadas por não terem sido consideradas necessárias para a modelagem nem reforçadas pelos participantes. Esse resultado pode estar relacionado ao "esquecimento" do *chatbot*, pela proposição de variáveis duplicadas, ou sugerir dificuldades dos LLMs na modelagem de dados; (iv) uma baixa acurácia (57%) de variáveis pertencentes à classe de $finalized\_challenges$; e (v) o acesso limitado ao *template* de dados do GLBoard, ocasionando a proposta de variáveis já existentes e sugestões de tipos incompatíveis com C#.

De modo geral, considerando os pontos positivos e ressalvas localizadas, os resultados sugerem que o problema foi parcialmente mitigado, mas não totalmente resolvido, e que as conjecturas teóricas são válidas, mas exitem limitações técnicas associadas à versão atual da ferramenta. A maioria das inconsistências e melhorias estão relacionadas ao comportamento do LLM: como opera com engenharia de *prompt*, sua atuação depende da interpretação das instruções e informações fornecidas. Apesar dessas limitações, os resultados sugerem que o GLORIA pode apoiar a modelagem de dados para GLA e estima-se que o aumento na quantidade de parâmetros do LLM, o treinamento especializado (*fine tuning*) e a inclusão de técnicas de RAG possam contribuir para aumentar a qualidade e a consistência das respostas e estruturas geradas.

Dentre as limitações do estudo, destacam-se: (i) a amostragem por conveniência, que pode ter introduzido viés na avaliação; (ii) o perfil homogêneo (todos da área de Computação), o que dificulta a generalização para *designers* de aprendizagem de outras áreas; (iii) a presença do pesquisador durante os testes, que pode ter influenciado o comportamento dos participantes; (iv) a ausência de mensuração da aprendizagem sobre modelagem de dados e GLA, uma vez que a análise se baseou apenas na percepção dos participantes; (v) o uso de jogos distintos pelos *designers* de aprendizagem, cujas características e eventuais limitações de *game design* podem ter influenciado as interações e as variáveis geradas, aspecto que não foi investigado; e (v) a classificação das variáveis como "adequadas", que foi realizada mediante análise das gravações. Nem todos os participantes explicitamente mencionaram concordância com as variáveis, de forma

que a classificação também tenha se baseado na interpretação subjetiva do mestrando. Essa estratégia pode ter introduzido vieses.

Apesar das limitações, todos os participantes pertenciam ao público-alvo do GLORIA: *designers* de aprendizagem com e sem experiência em GLA, que contribuíram com avaliações a partir de diferentes níveis de conhecimento. Além disso, os testes síncronos permitiram identificar inconsistências e sugestões de melhoria que possivelmente não seriam mapeadas em uso não supervisionado. Quanto à avaliação de aprendizagem dos conceitos, embora baseada somente na percepção dos participantes, trouxe evidências iniciais sobre o potencial da ferramenta para o apoio à modelagem de dados. A escolha de jogos distintos, embora as diferenças entre seus *game designs* não tenham sido controladas, permitiu o uso de jogos autorais e familiares aos participantes, o que pode ter favorecido o foco nas funcionalidades da ferramenta. A respeito da classificação das variáveis, para garantir maior precisão e minimizar o viés, foram classificadas internamente em: "explícitas" (usuário concordou com a inclusão ou exclusão), "implícitas" (o usuário expressou concordância ou discordância de forma implícita) e "desconhecidas" (não houve evidências suficientes para a classificação).

6.3 Síntese do Capítulo

Este capítulo contempla a avaliação do artefato GLORIA, ao qual se conduziu: (i) uma avaliação heurística, para verificar aspectos de usabilidade e analisar a conformidade do artefato com seus requisitos; e (ii) um estudo empírico, para avaliar se o problema (dificuldade na modelagem de dados para GLA) foi resolvido de modo satisfatório e validar as conjecturas teóricas (LLMs podem apoiar a modelagem de dados).

Na avaliação heurística, participaram três especialistas em *design* de interface e experiência do usuário (UX/UI). Os participantes foram introduzidos à ferramenta, utilizaram o GLORIA mediante a realização de tarefas, pontuaram comentários (*Think Aloud*) e responderam a um formulário com as heurísticas de Nielsen e perguntas abertas sobre avaliação geral. As gravações de áudio da avaliação foram registradas com consentimento e analisadas em conjunto com os demais dados coletados. Os resultados

indicaram boa receptividade à ferramenta, destacada como intuitiva e de fácil uso, além de avaliações majoritariamente positivas quanto às heurísticas de Nielsen.

A experiência prévia com jogos educacionais também possibilitou que os participantes contribuíssem com a avaliação do *chatbot* para modelagem de dados e GLA. Entretanto, diversas melhorias e algumas inconsistências foram identificadas durante a avaliação, sobretudo em relação à estética (organização do conteúdo e hierarquia visual) e ao *chatbot* (legibilidade e falhas no funcionamento). Dessa forma, há indícios de que o GLORIA apresenta conformidade com os seus requisitos, mas ainda necessita de refinamentos, o que está condizente com o fluxo iterativo do Ciclo de *Design*.

Quanto ao estudo empírico, participaram quatorze *designers* de aprendizagem (iniciantes e experientes em GLA). O teste iniciou-se com a apresentação do mestrando e a introdução à ferramenta, seguidas pela utilização do GLORIA pelos participantes. Em seguida, foram conduzidas entrevistas e aplicado um formulário avaliativo. Os dados coletados foram analisados com as gravações das reuniões, registradas com consentimento. Por meio do instrumento "*Player Level Up!*", os participantes avaliaram suas próprias estruturas de GLA, geradas na interação com o GLORIA. Os resultados foram predominantemente positivos, embora os participantes mais experientes tenham atribuído pontuações menores. Entretanto, observou-se uma inconsistência de "esquecimento" do LLM, que não incorporou ao JSON algumas variáveis propostas e aprovadas explícita ou implicitamente pelos participantes.

Todos os participantes afirmaram que utilizariam o GLORIA para auxiliar na implementação de GLA, destacando a facilidade de uso, o apoio à modelagem de dados, a praticidade e a velocidade das respostas. No entanto, apontaram ressalvas quanto à clareza do fluxo e da navegação, inconsistências técnicas, melhorias em *design* e ausência de variáveis importantes para a coleta de dados. O comportamento do *chatbot* também foi destacado, sobretudo pelos usuários experientes em GLA, que relataram dificuldades na interação. Diante disso, os achados indicam o problema foi parcialmente mitigado, mas não totalmente resolvido. Ademais, os resultados sustentam as conjecturas teóricas, embora persistam limitações técnicas na versão atual da ferramenta.

7

DISCUSSÕES DOS ACHADOS

Este capítulo discute e correlaciona os principais achados desta dissertação, estabelecendo um nexo entre o problema investigado, as lacunas identificadas na literatura, os estudos conduzidos e a proposta do artefato desenvolvido. A pesquisa apresentou como questão de pesquisa: "Como auxiliar *designers* de aprendizagem a modelar dados e implementar técnicas de *Game Learning Analytics* (GLA) que possam contribuir para a identificação de evidências de aprendizagem em jogos educacionais?". Os resultados da pesquisa, na literatura e em campo, indicam que a modelagem de dados constitui uma etapa crítica e pouco explorada em GLA, envolvendo desafios relacionados à abstração, definição semântica das variáveis de captura e relação entre elementos pedagógicos e computacionais de um jogo educacional.

Nesse cenário, investigou-se o uso de LLMs como mecanismo de apoio à modelagem de dados, resultando na proposição da ferramenta GLORIA, concebida para atuar como um módulo interativo integrável a sistemas de GLA[1]. A ferramenta foi projetada para apoiar *designers* de aprendizagem na definição de variáveis contextualizadas às características e aos objetivos de aprendizagem dos jogos, visando à posterior identificação de evidências de aprendizagem. Para além da proposição do artefato, a dissertação também consolida evidências, desafios e direções de pesquisa que podem contribuir para o avanço das discussões sobre modelagem de dados em *Game Learning Analytics* e uso de *Large Language Models* em jogos educacionais.

[1] Ao qual utilizou-se como estudo de caso a integração com o GLBoard.

7.1 A modelagem de dados como desafio estrutural em *Game Learning Analytics*

Esta pesquisa surgiu da percepção de que *designers* de aprendizagem demonstravam dificuldades em aplicar GLA em seus jogos. Embora conseguissem definir os objetivos de aprendizagem, era desafiador entender como isso se traduziria em dados. Ao longo do estudo, foi possível confirmar a hipótese inicial. Esse cenário é crítico, pois o núcleo de GLA são os dados: caso apresentem erros ou haja variáveis não instanciadas, as conclusões obtidas podem acarretar prejuízos para aqueles que dependem dessas informações, por exemplo, em diagnósticos de avaliação da aprendizagem.

Nesse sentido, compreendeu-se que a modelagem de dados assume um papel fundamental no processo de GLA, pois consiste na definição dos dados a serem coletados em um jogo educacional, articulando elementos de *game design* e *learning design* (Shoukry; Göbel; Steinmetz, 2014). Esses dados devem ser suficientes para expressar se os objetivos de aprendizagem estão, de fato, sendo atingidos (Perez-Colado et al., 2018). Os dados definidos orientam a criação e a organização das variáveis de GLA em um *template* de dados, que estabelece sua estrutura e hierarquia. A partir desse *template*, as variáveis são implementadas em uma *game engine*, possibilitando a geração de registros de interação (*logs*) dos jogadores. A importância da modelagem de dados também foi percebida por estudantes que atuaram como *designers* de aprendizagem no contexto de disciplinas (Capítulo 3), os quais destacaram como contribuições a redução do tempo de implementação, a definição de variáveis para estimar indícios de aprendizagem, a melhor compreensão dos aspectos do jogo e do comportamento do jogador, bem como a identificação de falhas no *game design*.

Por outro lado, embora a modelagem de dados constitua uma etapa necessária em GLA, identificou-se uma menor concentração de discussões sobre a definição pedagógica e semântica das variáveis. Esse aspecto relaciona-se à evolução histórica do campo de GLA: até 2017, buscava-se uma sistematização para contornar a complexidade de implementação das técnicas e padronizar a coleta de dados. Como resultado, surgiram modelos como o xAPI-SG (Alonso-Fernandez et al., 2017) e o ADAGE (Owen, 2014). A partir de 2017, com a existência de modelos sistematizados de GLA, a literatura con-

centrou esforços em pesquisas sobre captura e coleta, visualização de dados, *dashboards* e análises de desempenho dos jogadores (Loh; Sheng; Ifenthaler, 2015; Alonso-Fernandez et al., 2017), mas permaneceu limitada quanto à forma como os dados foram definidos. Essa afirmação se confirma durante a execução de uma *Scoping Review* (Capítulo 2). Ainda que a modelagem seja contemplada em modelos conceituais como o *Learning Analytics Model* (LAM) (Perez-Colado et al., 2018), há uma ausência de materiais que orientem sua realização e que forneçam exemplos concretos.

Relacionado a esse fator, observa-se que a maioria dos trabalhos de GLA é internacional (Banihashem et al., 2024; Geremias et al., 2024), geralmente incluindo equipes multidisciplinares e a distinção entre papéis como *learning designer*, *game designer*, *analytics developer*, entre outros. Nesse sentido, estima-se que já seja pressuposto que a modelagem é uma etapa que os pesquisadores responsáveis saibam fazer. Por exemplo, Alonso-Fernández et al. (2017), afirmam que educadores podem fornecer resultados quantificáveis para serem medidos nas análises, o que se relaciona indiretamente com a modelagem de dados. Em Perez-Colado et al. (2018), os autores descrevem que o LAM é concebido de forma conjunta por diferentes papéis, incluindo o *designer* de aprendizagem. Esse papel é responsável pela definição dos objetivos de aprendizagem, enquanto os demais os articulam às metas do jogo, aos rastros de interação e às análises. Portanto, a modelagem de dados é apresentada como um esforço conjunto de diferentes profissionais, dos quais já se espera o domínio das competências necessárias.

Por outro lado, em contexto brasileiro, a maioria dos jogos são feitos em universidades (Oliveira; Santos; Pereira, 2019), envolvendo, muitas das vezes, estudantes que assumem todos esses papéis no desenvolvimento de um jogo educacional e podem não possuir os conhecimentos esperados para modelar dados de GLA com qualidade. Somado a isso, ainda que a aprendizagem seja objeto de estudo da Psicologia e Pedagogia, tal como as Licenciaturas, uma quantidade elevada de jogos educacionais são propostos por pesquisadores de áreas de Computação, principalmente em congressos e simpósios da Sociedade Brasileira de Computação (SBC), como o CBIE[2], SBGames[3], CSBC[4], dentre outros. Diante disso, os responsáveis pela modelagem de dados de jogos educacionais

[2] https://sol.sbc.org.br/index.php/cbie/about
[3] https://sol.sbc.org.br/index.php/sbgames/about
[4] https://sol.sbc.org.br/index.php/csbc/about

podem não possuir os conhecimentos pedagógicos necessários.

A presença limitada de estudos na literatura sobre a modelagem de dados foi uma das principais motivações para conduzir as investigações apresentadas no Capítulo 3, onde se observou que sua execução não é trivial. Por mais que estudantes de uma universidade pública, que atuaram como *designers* de aprendizagem e realizaram a modelagem de dados de forma manual (sem LLMs), tenham compreendido a importância desse processo para GLA, apresentaram obstáculos diversos. Dentre as principais dificuldades apontadas, destacam-se: (i) o nível de abstração exigido para o processo é elevado; (ii) o entendimento do jogo para a definição das variáveis de GLA não é simples; (iii) o tempo destinado à modelagem foi insuficiente; e (iv) a falta de conhecimento prévio de disciplinas de computação pode ter prejudicado a modelagem. Assim como em outras investigações (Capítulo 4), os estudantes consideraram que modelar dados não foi fácil, ainda que sejam autores dos próprios jogos, responsáveis pela criação das mecânicas e da definição dos objetivos de aprendizagem.

Os estudos desta pesquisa indicam que a modelagem de dados para GLA envolve não apenas a definição pedagógica dos objetivos de aprendizagem, mas também a abstração e a representação computacional das interações do jogador. Portanto, a área de computação tem um papel relevante na modelagem de dados, cuja natureza envolve conceitos relacionados da área, como: criação de diagramas, análise de entidades e identificação de variáveis. Essa característica é percebida, principalmente, por *designers* de aprendizagem em formação em cursos de Computação, que afirmam que disciplinas de programação, banco de dados e modelagem de sistemas contribuem para a modelagem de dados (Capítulo 3). Além disso, os conhecimentos de computação podem auxiliar na idealização de funcionalidades associadas ao campo de GLA, conforme demonstrado no *Design* Participativo (Seção 5.1). Pôde-se observar que os estudantes, ao conceberem propostas de sistemas de GLA com ênfase na modelagem de dados, aplicaram conhecimentos de computação na idealização das propostas ao considerar que: (i) todos os estudantes mencionaram o uso de LLMs; (ii) técnicas de visão computacional foram sugeridas; e (iii) as propostas incluíram a análise de estruturas computacionais do jogo para auxiliar na modelagem de dados, como códigos e *logs* de interação.

A partir dessas discussões, percebe-se que a modelagem de dados para jogos educacionais, ao mesmo passo em que é fundamental, pode ser complexa. Sua qualidade depende de fatores das áreas de jogos (abstração das mecânicas), computação (tipagem e estruturação dos dados) e pedagogia (relação dos dados com indícios de aprendizagem). Ressalta-se que essa etapa foi considerada complexa até por estudantes de Licenciatura em Computação que possuem conhecimentos prévios nas três áreas. No entanto, a modelagem de dados ainda não recebe a devida atenção na literatura, constituindo uma lacuna que demanda novas pesquisas.

Nesse sentido, dada sua relevância, a modelagem não deve ser tratada como uma etapa trivial e passar despercebida, pois ela é o "maestro" que coordena a sinfonia de GLA: a partir dela emergem as variáveis de GLA, que são organizadas em um *template* de dados, implementadas em uma *game engine* e posteriormente geram registros de interação (*logs*) que permitem a identificação dos indícios de aprendizagem.

Portanto, como um dos principais e inéditos achados desta dissertação, propõe-se uma modificação na arquitetura de GLA de Freire et al. (2016), adicionado um novo módulo, denominado "Modelador" (Figura 43). Sua responsabilidade é definir os dados, em consonância com os objetivos de aprendizagem do jogo educacional, orientando a implementação das variáveis de GLA no módulo "Motor de jogo". A partir dessa implementação, são gerados os registros de interação (*logs*) que serão recebidos pelo módulo "Coletor". O "Modelador" representa o ponto de partida de GLA, evidenciando que a definição dos dados deve caminhar em conjunto com o desenvolvimento do jogo educacional, visto que uma modelagem inconsistente pode comprometer as etapas subsequentes. Por esse motivo, a inserção de técnicas de GLA em jogos parcialmente ou completamente desenvolvidos é um desafio e não é recomendada (Macena et al., 2024).

7.2 A aplicação de LLMs para contextos de aprendizagem

Originalmente, esta pesquisa tinha como foco exclusivo fornecer suporte à modelagem de dados para auxiliar na identificação de evidências de aprendizagem em jogos educacionais. Entretanto, ao adotar LLMs para auxiliar nesse processo, observou-se

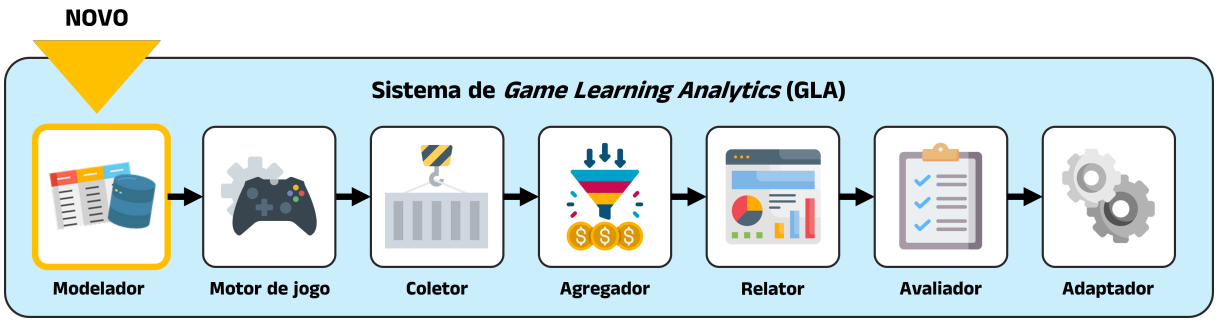


Figura 43 – Proposta de arquitetura de GLA com base em Freire et al. (2016).

que seu uso pode não somente contribuir para modelar dados em jogos, mas também atuar como uma forma de *scaffolding*. Diante disso, essa estratégia também permite um acompanhamento reflexivo do desenvolvimento do *designer* de aprendizagem para compreender a como realizar a modelagem de dados de forma eficiente. Esse aspecto torna-se ainda mais relevante ao considerar que jogos educacionais são complexos, demandando conhecimentos multidisciplinares.

A aplicação de LLMs para o campo de GLA ainda é restrita na literatura, conforme observado nesta pesquisa pela *Scoping Review* do Capítulo 2 e analisando o Mapeamento Sistemático da Literatura (MSL) de Geremias et al. (2024). Pressupõe-se que uma das dificuldades em localizar trabalhos nessa intersecção esteja relacionada ao uso de terminologias distintas para ambas as áreas. Além disso, pela modelagem de dados não ser um objeto constante de estudos na literatura, não se identificou, até o presente momento desta dissertação, pesquisas que façam uso de LLMs para auxiliar na definição de dados conforme objetivos de aprendizagem de jogos educacionais[5].

Tradicionalmente, modelar dados era realizado de forma manual, ao qual dependia unicamente dos conhecimentos e capacidade de inferência do *designer* de aprendizagem para definir os dados de coleta que permitam a identificação de evidências de aprendizagem. Essa atividade não foi considerada trivial por um grupo de estudantes que atuavam como *designers* de aprendizagem, ainda que fossem autores dos jogos. Como alternativa, a partir de investigações na literatura, definiu-se como pressuposto teórico que LLMs poderiam ser aplicados em contextos de GLA para mitigar os desafios

[5] Exceto as do grupo de pesquisa do mestrando, apresentadas no Capítulo 8, que derivaram desta dissertação.

em modelagem de dados para jogos educacionais. Esse pressuposto fundamentou-se na capacidade desses modelos de gerar textos em linguagem natural semelhantes aos produzidos por humanos, responder a perguntas e realizar outras tarefas relacionadas à linguagem (Kasneci et al., 2023), que suscitaram pesquisas em áreas baseadas em dados. Por exemplo, seu uso para *Learning Analytics* auxilia na análise, integração, síntese e interpretação de dados (Misiejuk et al., 2025), enquanto, para *Game Analytics*, pode contribuir para adaptação dinâmica do jogo conforme *logs* de interação (Mitsea; Drigas; Skianis, 2025). Nesse sentido, esses modelos poderiam ser relevantes para GLA, que congrega conceitos de ambas as áreas, sobretudo para a modelagem de dados.

Como forma de compreender a viabilidade do uso de LLMs para minimizar a complexidade da modelagem de dados, realizou-se um "teste piloto" em forma de experimento controlado (Capítulo 3). Após a definição dos dados de coleta, as variáveis de GLA são criadas e organizadas em um *template* de dados (estrutura de captura). Portanto, investigou-se se um LLM poderia gerar uma estrutura de GLA contextualizada com base nas informações de um jogo educacional. Utilizou-se o ChatGPT, ao qual suas três versões (3.5, 4 e 4o) foram capazes de gerar estruturas distintas, que posteriormente foram comparadas com a de um especialista em GLA que a concebeu de forma manual (sem LLMs). Os resultados demonstraram que a estrutura gerada pela versão 4o do modelo obteve avaliações inferiores a do especialista em GLA, mas com pontuações semelhantes. Esse resultado é particularmente relevante, sobretudo ao considerar que: (i) poucas informações foram disponibilizadas sobre o jogo; (ii) nenhum exemplo foi informado no *prompt*, como outras estruturas; e (iii) o modelo não tinha conhecimentos elevados sobre o GLBoard, somente o *template* informado. Portanto, ao gerar uma estrutura de captura avaliada positivamente por pessoas experientes na área e de forma semelhante à de um especialista em GLA, obteve-se os primeiros indícios de que os LLMs podem ser usados como suporte à modelagem de dados.

Dando continuidade às investigações no uso de LLMs para GLA, observou-se que técnicas interativas de *prompt*, como a de *chain-of-thought prompting* (Wei et al., 2022), adicionam uma camada intermediária de "raciocínio" para potencializar as respostas do modelo. Inspirado nessa descoberta, projetou-se uma técnica de questionamentos

durante o uso de LLMs para a modelagem de dados, ao qual o modelo faz perguntas ao usuário sobre seu jogo educacional para incorporar mais contexto e aprimorar as estruturas de GLA. Conforme apresentado no Capítulo 3, estudantes que atuavam como *designers* de aprendizagem avaliaram a técnica de forma positiva, destacando que contribui para o aperfeiçoamento das estruturas. No entanto, observou-se uma certa complexidade dos estudantes na concepção de *prompts* para orientar o LLM a realizar essa tarefa, podendo ter relação, além da inexperiência dos estudantes em engenharia de *prompt*, com a complexidade da modelagem de dados.

Durante o uso do LLM pelos estudantes, principalmente respondendo aos questionamentos sobre seus jogos educacionais, observou-se que as interações não contribuíram somente para fornecer contexto ao modelo. Elas também atuaram como *scaffolding*, orientando a reflexão dos *designers* de aprendizagem sobre a modelagem de dados. Desse modo, os participantes eram estimulados a pensar em como os dados poderiam emergir para expressar aprendizagem, articulando implicitamente conceitos pedagógicos e de *game design*. Além disso, posteriormente com o uso do GLORIA, a maioria dos participantes explicitamente informou que a ferramenta auxiliou compreensão da modelagem de dados, sobretudo sendo considerado como bem explicativa e com soluções pertinentes para pessoas iniciantes. Destacam-se também participantes que informaram que as perguntas sobre os dados que devem ser coletados instigam a reflexão. Observações da pesquisa em sala de aula e durante as investigações também corroboram com esse aspecto: pelo domínio específico, a interação dos participantes com LLMs (sejam os tradicionais, o agente inteligente ou a ferramenta GLORIA) se dá de forma mais fluida e proporciona o aprendizado não apenas sobre a modelagem dos dados, mas a respeito de seus próprios jogos educacionais.

Nesse sentido, ao utilizar estratégias que incorporam LLMs a atividades de GLA, como a modelagem de dados, os usuários também exercitam os conceitos pedagógicos vinculados aos seus jogos educacionais, pois precisam compreender como o *level design* promoverá a mobilização desses conceitos pelos estudantes, visando à coleta dos dados correspondentes. Esse processo fomenta uma reflexão metacognitiva, na medida em que os *designers* de aprendizagem não pensam somente no conteúdo a ser trabalhado,

mas também em como ele será evidenciado durante a experiência de jogo. A partir disso, de forma indireta, o modelo também pode vir a contribuir com ideias de *game design* e *learning design*, auxiliando no refinamento do jogo, como pontuado por alguns participantes nas investigações desta pesquisa.

Embora LLMs tradicionais tenham demonstrado indícios para apoiar a modelagem de dados (estruturas de GLA avaliadas positivamente por especialistas na área), projetar *prompts* pode ser não ser simples. Adicionalmente, dada a quantidade de informações na entrada (instruções e dados do jogo educacional), o LLM pode gerar alucinações (IBM, 2023a) (respostas incoerentes) ou "ignorar" informações (Liu et al., 2024). Como alternativa, em 2023, a OpenAI lançou a funcionalidade "MyGPTs", que permite criar e customizar agentes no ChatGPT. A partir da configuração do comportamento do agente e da inserção de arquivos em sua base de conhecimento, o LLM gera respostas orientadas pelas instruções definidas e fundamentadas nas informações recuperadas nesses arquivos por meio de técnicas de RAG.

Nesse sentido, para investigar o potencial de agentes ao campo de GLA, criou-se o "*GLA Specialist*" (Capítulo 4). Ele é caracterizado como um agente especialista para auxiliar a modelar dados e gerar estruturas de captura de GLA, desenvolvido em formato de POC para compreender seus potenciais. O agente pode ser utilizado sem a necessidade de *prompts* complexos, sendo capaz de modelar dados (definição e justificativa), organizá-los no *template* do GLBoard e gerar exemplos de captura com dados sintéticos. Os resultados, a partir da avaliação com sete especialistas na área, indicaram que o agente foi bem recebido, principalmente pela precisão das respostas, facilidade na criação das estruturas e diminuição da complexidade ao preencher o *template* de dados. Esse achado também evidencia que LLMs podem apoiar a modelagem de dados, reduzindo a complexidade de projeção de *prompts*.

Além disso, o agente também foi avaliado por estudantes que atuavam como *designers* de aprendizagem em disciplinas com foco na construção e avaliação de jogos educacionais, que pontuaram que a interação geral foi fácil, que a dificuldade de uso foi baixa, que a quantidade de variáveis propostas pelo agente foi suficiente e que é uma estratégia mais eficiente que a técnica manual. Ressalta-se que, embora os resultados

tenham sido considerados positivos, uma vez que o agente propôs mais variáveis do que a modelagem manual, o uso de modelos de IA Gerativa pode gerar uma percepção irreal de confiabilidade em relação aos dados obtidos. Nesse sentido, os resultados devem ser interpretados com cautela e podem não ser generalizáveis. Wang et al. (2025) também obtiveram resultados semelhantes: os estudantes apresentaram desempenho superior ao utilizar o LLM-ARG, uma abordagem baseada em LLMs para auxiliar no engajamento, performance e consciência metacognitiva em jogos de realidade alternativa (ARG), em comparação a um ARG tradicional. Já a pesquisa de Kiyak e Kononowicz (2024) demonstrou que um agente customizável para o contexto clínico sobressaiu-se em relação a um LLM tradicional. Nesta pesquisa, estudantes destacaram esse aspecto, tanto na avaliação do "*GLA Specialist*" quanto na ferramenta GLORIA, principalmente: a praticidade, visto que não é necessário uma configuração inicial para o contexto específico; o domínio, tornando-se um especialista no assunto; e a ausência de *prompts* complexos, simplificando as interações. Essas avaliações ressaltam a contribuição do uso de LLMs para GLA e destacam que um agente inteligente pode ser mais apropriado em comparação com modelos tradicionais[6].

Apesar de sua boa avaliação, tanto por especialistas em GLA quanto por estudantes que atuavam como *designers* de aprendizagem, o agente parcialmente "protagonizava" a modelagem de dados, sendo o responsável pela geração e justificativa das variáveis de GLA a partir das informações dos jogos. Nesse contexto, o *designer* de aprendizagem atuava principalmente como um validador, solicitando alterações e melhorias somente quando necessário. Embora fosse estimulado a refletir com os questionamentos sobre seu jogo educacional, o usuário participava de forma limitada na modelagem de dados. Portanto, o LLM não era utilizado como um mediador de conhecimento, e sim como agente central da modelagem, o que não é recomendado para usuários iniciantes que precisam compreender o processo. Esse aspecto foi observado por um dos estudantes, que se sentiu mais confiante ao utilizar a técnica manual pelo protagonismo que proporciona ao *designer* de aprendizagem ao modelar os dados (Capítulo 4). Esse aspecto serviu de motivação para um dos requisitos da ferramenta

[6] Disponíveis para uso direto, como o ChatGPT, que não se referem a versões customizáveis ou configuráveis como agentes.

GLORIA (Capítulo 5), que dispõe de um modo "Livre", cujo objetivo é atuar como um tutor para auxiliar *designers* de aprendizagens iniciantes a compreender como modelar dados na prática. Portanto, pôde-se perceber que a técnica manual e o uso de LLMs como estratégias de modelagem não se excluem, se complementam. Esse achado também se faz presente no trabalho de Liu et al. (2025), cujos resultados indicam que o LLM analisado para detectar comportamentos de dificuldade (*struggle*) em jogos educacionais não substituiu os modelos tradicionais, mas atua de forma complementar às técnicas de *machine learning*. Esse ponto reforça o papel do *designer* de aprendizagem e que o uso de LLMs não tem como objetivo substituí-lo, mas apoiar a modelagem de dados e estimular seu aprendizado, conforme discutido anteriormente.

Comparando o uso do agente com LLMs tradicionais, observa-se uma preferência dos estudantes pelo agente, que também foi bem avaliado por especialistas da área. Esses resultados sugerem que a personalização de LLMs pode potencializar suas contribuições à modelagem de dados, sobretudo quando associada a técnicas como engenharia de *prompt*, RAG e, quando aplicável, treinamento adicional (como *fine-tuning*). A combinação dessas estratégias pode ter favorecido a geração de respostas mais contextualizadas ao domínio de GLA, contribuindo para a boa avaliação do agente. Em contrapartida, o uso de LLMs fechados pode apresentar limitações: ainda que o uso do agente inteligente esteja condicionado à assinatura da versão Plus (não gratuita) do ChatGPT, há uma quantidade máxima de interações permitidas por período de tempo. Ademais, modelos cujo código-fonte não é disponibilizado de forma pública para ajuste dos pesos limitam a capacidade de customização. Essas limitações são críticas no contexto da modelagem de dados, considerando ser uma atividade complexa e que comumente não é executada de forma rápida. Diante disso, como alternativa, identificou-se o uso de LLM *open-source*, que se transformou em requisito para a ferramenta GLORIA, trazendo maior possibilidade de personalização do modelo e removendo a restrição de quantidade de interações.

Outro aspecto observado durante interação dos estudantes com LLMs, sobretudo com o uso da ferramenta GLORIA, é o de que a experiência em GLA influencia na interação. No estudo empírico (Capítulo 6) com o GLORIA, o grupo A (pessoas com

experiência sólida no uso de GLA) foi o que mais propôs variáveis junto ao *chatbot* (90 de 160, 56%). Além disso, embora tenham avaliado as estruturas de GLA de forma positiva, foi o grupo com as menores notas. Esses pontos sugerem que, dado a experiência prévia na área, já saibam quais dados coletar e puderam identificar fragilidades com maior clareza. A maioria desses participantes não fez o uso do GLORIA como aprofundamento teórico e sim de forma mais técnica: gerar as variáveis de GLA. Entretanto, a ferramenta possuía ênfase no modo "Tutorial", guiando os usuários de forma prática, o que pode ter prejudicado esses participantes que já possuíam os conhecimentos esperados. Essa característica é o inverso do "*GLA Specialist*", que modelava diretamente as variáveis e não focava-se em orientar o usuário nesse processo, o que atendia à expectativa de pessoas que já eram experientes em GLA. Portanto, conclui-se que a experiência deve ser considerada ao construir soluções que envolvam GLA e LLMs, com funcionalidades que atendam a quem já possui e a quem não possui conhecimentos em GLA.

De forma geral, com as contribuições de agentes baseados em LLM para a área de GLA, sobretudo a modelagem de dados, as principais dificuldades dos estudantes podem ser mitigadas: (i) abstração elevada: por meio de um diálogo com um agente, que foi orientado a sanar dúvidas sobre GLA, o entendimento do passo a passo pode se tornar mais simples; (ii) entender o jogo para definir as variáveis não é simples: os questionamentos instigam o pesquisador a refletir sobre o jogo educacional e a como os dados podem emergir das mecânicas de aprendizagem; (iii) tempo insuficiente: embora relacionado ao contexto do estudo, o uso da ferramenta GLORIA é irrestrito, cujo usuário pode modelar os dados pelo tempo que julgar necessário; e (iv) conhecimentos prévios de computação podem ter prejudicado a modelagem: a ferramenta segue um padrão (modelo GLBoard), guiando o usuário em estruturar computacionalmente as variáveis de GLA. Esse ponto é relevante, principalmente ao considerar que nem todos os *designers* de aprendizagem pertencem a áreas de Computação.

7.3 GLORIA: concretizando os pressupostos teóricos

A partir da discussão anterior a respeito do estado da arte de modelagem de dados em GLA e o potencial de LLMs para auxiliar nessa etapa, obtiveram-se achados diversos. Como forma de concretizá-los e operacionalizar o pressuposto teórico investigado (LLMs como apoio a modelagem de dados), esta dissertação também envolveu a consolidação de um artefato. Sua construção deu-se por meio de uma abordagem metodológica-epistemológica de *Design Science Research* (DSR), resultando na ferramenta GLORIA (*Game, Learning, Observation, Recommendation, Intelligence and Analytics*).

A ferramenta contempla os achados obtidos com as investigações anteriores, em forma de requisitos ou regras de negócio: (i) trata-se de um módulo introdutório para modelagem de dados, sendo a etapa inicial de GLA; (ii) considera que o *designer* de aprendizagem pode ou não ter os conhecimentos esperados. Nesse sentido, sua base de conhecimento contém informações detalhadas sobre a área de GLA e o modelo GLBoard (utilizado como estudo de caso), permitindo-os sanar dúvidas quando necessário; (iii) a ferramenta contém um *chatbot* orientado a questionamentos, estimulando a reflexão do *designer* de aprendizagem ao mesmo passo que incorpora mais contexto para si. Por meio disso, busca minimizar a complexidade de abstração do jogo e da definição de variáveis, atuando como um "tutor"; (iv) a ferramenta contorna a complexidade em criar *prompts* complexos para instruir o modelo e configurá-lo, incluindo um sistema de cadastro de jogos educacionais; (v) por meio do modo "Tutorial", busca-se não gerar a modelagem diretamente ao usuário, mas atuar como um mediador de conhecimento para que o *designer* de aprendizagem também possa aprender a fazer e, eventualmente, adquirir os conhecimentos esperados e utilizar a ferramenta como apoio à modelagem (*scaffolding*); e (vi) o LLM incorporado à ferramenta é *open-source*, não havendo limitações de uso ou restrição de personalização.

O GLORIA representa uma evolução do "*GLA Specialist*" (agente no ChatGPT), que originalmente foi construído como uma POC. Dada sua avaliação positiva por especialistas na área e *designers* de aprendizagem iniciantes, pôde-se prosseguir para uma versão mais abrangente. Nesse contexto, o GLORIA incorpora novas funcionalidades em comparação a sua POC (agente): um sistema de *login* e cadastro; uma área de geren-

ciamento de jogos educacionais; a organização de *chatbots* conforme o jogo, podendo conter uma ou mais conversas; uma área que facilita a visualização da modelagem de dados e da estrutura de captura gerada; e a escolha do modo de interação (Tutorial e Livre), visando direcionar o usuário conforme sua experiência prévia em GLA; e a integração com um LLM *open-source*, evitando a restrição de uso e flexibilizando a personalização. Por ser tratar de um módulo, o GLORIA pode ser incorporado a sistemas como o GLBoard, compondo um ecossistema completo de *Game Learning Analytics*.

Entretanto, ressalvas foram identificadas durante os testes, sobretudo ajustes relacionados à experiência do usuário e usabilidade, bem como inconsistências técnicas associadas ao *chatbot*. A principal inconsistência refere-se ao "esquecimento", onde o modelo ocasionalmente desconsiderava variáveis propostas anteriormente, o que dificultou o uso da ferramenta por especialistas em GLA. Além disso, a ferramenta ainda não inclui técnicas de RAG como o agente inteligente, o que poderia aprimorar as respostas geradas no contexto de GLA. O modo "Tutorial", embora tenha sido proposto para apoiar pessoas sem experiência em modelagem de dados, atualmente se difere apenas no envio de uma informação no *chat*. Nesse sentido, pôde-se observar que a ferramenta ainda necessita de ajustes e, embora comporte funcionalidades adicionais, as modelagens de dados ainda não superam às do agente (quantidade de variáveis propostas, inferências, fluxo de interação, entre outros). Ressalta-se que esse aspecto foi percebido pelo mestrando a partir de análises e observações, mas não foi investigado na dissertação. Essa comparação direta entre agente e ferramenta GLORIA será realizada em estudos futuros, assim como: correções de inconsistências; melhorias visuais e de usabilidade; incorporação de novas funcionalidades (detalhadas no Capítulo 8); e integração efetiva ao GLBoard, que atualmente se encontra em fase de refatoração.

8

Considerações finais

Jogos educacionais são elementos lúdicos que podem contribuir para a aprendizagem, proporcionando motivação e engajamento aos estudantes (Plass; Homer; Kinzer, 2015; Tarouco et al., 2004). Entretanto, avaliar se o jogador está efetivamente aprendendo não é trivial, principalmente com o uso isolado de questionários como estratégia avaliativa (Silva et al., 2021). Como alternativa, tem-se o uso de técnicas de GLA: coletar, analisar e visualizar registros de interação dos jogadores (Freire et al., 2016). Essa abordagem auxilia na identificação de evidências de aprendizagem, rastreando o percurso do jogador. Uma das ferramentas para implementar GLA é o GLBoard (Silva et al., 2022), que dispõe de um *template* genérico para a captura de dados de jogos educacionais.

Entretanto, definir as variáveis de GLA a serem coletadas e justificar sua relação com os objetivos de aprendizagem não é trivial (Petersen; Bedek, 2012; Perez-Colado et al., 2018; Honda et al., 2025). Essa etapa denomina-se modelagem dos dados e deve ocorrer desde a concepção de um jogo (Hauge et al., 2014). A partir das variáveis, pode-se organizá-las em um *template* de dados e incorporá-las ao jogo educacional por meio de uma *game engine*. O avanço na área de IA Gerativa sugere que os LLMs possam contribuir com a modelagem de dados, tratando-se de modelos capazes de gerar textos semelhantes aos de seres humanos e realizar tarefas de linguagem natural (Kasneci et al., 2023). Dadas as contribuições dos LLMs para os campos de LA e GA (Misiejuk et al., 2025; Yang; Kleinman; Harteveld, 2025), podem ser uma potencial alternativa para mitigar os desafios de modelagem de dados, contribuindo para GLA.

Nesse contexto, esta dissertação apresenta como questão de pesquisa (QP):

"Como auxiliar *designers* de aprendizagem a modelar dados de jogos educacionais para que possam implementar técnicas de *Game Learning Analytics* e identificar evidências de aprendizagem?". Para respondê-la, adotou-se a abordagem metodológica-epistemológica de *Design Science Research* (DSR) (Hevner, 2007), com o objetivo de compreender o estado da arte sobre modelagem de dados e o uso de LLMs para GLA, operacionalizando os achados por meio da construção de um artefato computacional. Ele é categorizado como "melhoria": novas soluções para problemas já conhecidos (Vaishnavi; Kuechler, 2015). Portanto, buscou-se construir uma ferramenta integrada à LLM (nova solução) para mitigar a complexidade de modelar dados e de implementar técnicas de GLA em jogos educacionais (problema já conhecido).

Inicialmente, por meio de revisões da literatura, identificou-se que a intersecção entre as áreas de GLA e LLMs ainda é limitada. Além disso, observou-se que os trabalhos de GLA geralmente não têm a modelagem de dados como foco principal. Em seguida, a partir de investigações, observou-se que: (i) estudantes que atuam como *designers* de aprendizagem apresentam dificuldades em modelagem de dados; (ii) há potencial no uso de LLMs para GLA, sobretudo na geração de estruturas de captura de dados; e (iii) técnicas interativas de *prompt* que envolvem questionamentos podem aprimorar os resultados, mas projetá-los não é trivial. Diante disso, foi proposto e desenvolvido, em formato de POC, um agente de conversação inteligente intitulado "*GLA Specialist*". O agente foi avaliado por *designers* de aprendizagem, incluindo especialistas na área e estudantes que atuavam nessa função. Embora tenha sido bem recebido pelos avaliadores, apresentou uma limitação de LLMs fechados quanto ao uso.

Em seguida, realizou-se um *Design* Participativo com estudantes que atuavam como *designers* de aprendizagem para levantar requisitos da ferramenta. Dentre as contribuições, destacam-se sugestões de funcionalidades futuras, como: uso de imagens e vídeos para a modelagem de dados, análise do código para auxiliar na inserção das variáveis de GLA e adaptação da estrutura conforme os *logs* gerados. Conforme os achados das investigações anteriores e do *Design* Participativo, iniciou-se a construção da ferramenta por meio de uma metodologia de *game design* educacional (Pires et al., 2021). A ferramenta resultante foi denominada GLORIA (*Game, Learning, Observation, Recom-*

mendation, Intelligence and Analytics), cujo objetivo é auxiliar *designers* de aprendizagem a modelar dados em GLA. O usuário pode cadastrar dados do seu jogo educacional, adicionando implicitamente contexto para um LLM. A partir desses dados, o agente inteligente (*chatbot*) realiza questionamentos para incentivar o usuário a refletir sobre seu jogo e agregar mais informações ao contexto. Portanto, a ferramenta atua como tutor, orientando o usuário na modelagem de dados.

Três avaliações foram conduzidas com a ferramenta, visando verificar se está de acordo com seus requisitos, se resolve o problema de modo satisfatório e se as conjecturas teóricas são válidas. Em relação aos requisitos, uma avaliação heurística foi conduzida com três especialistas em *design* de interface e experiência do usuário (UX/UI). Os resultados foram positivos, principalmente quanto à facilidade de uso, mas ressalvas foram identificadas quanto à estética e ao *chatbot*. Nesse aspecto, há indícios que o GLORIA atende aos requisitos, ainda que necessite de melhorias. Quanto à resolução do problema e conjecturas, um estudo empírico foi conduzido com quatorze *designers* de aprendizagem, com e sem experiência prévia em GLA. Pontos positivos também incluem a facilidade de uso, bem como o apoio à modelagem de dados, a praticidade e a velocidade das respostas. No entanto, as ressalvas identificadas incluem: inconsistências técnicas, necessidades de melhoria no *design*, variáveis relacionadas às mecânicas e aos objetivos de aprendizagem consideradas ausentes pelos participantes e fluxo e navegação pouco claros. Portanto, constata-se que o problema foi parcialmente mitigado e que as conjecturas teóricas são sustentadas, mas que também se faz necessário refinamentos na versão atual do GLORIA.

Portanto, em resposta à QP, para auxiliar *designers* de aprendizagem na modelagem de dados em GLA, pode-se projetar uma ferramenta interativa integrada com LLM que ofereça orientações contextualizadas de acordo com as características do jogo educacional e estimule a reflexão. Essa ferramenta, denominada GLORIA, compreende: (i) o cadastro de jogos educacionais, contemplando suas informações principais de *game design* e *learning design*; (ii) a modelagem dos dados, por meio de interações com um *chatbot* que realiza questionamentos para instigar reflexões ao usuário; e (iii) a organização das variáveis de GLA no *template* de dados do GLBoard, permitindo o

download da estrutura JSON resultante. Nesse sentido, a proposta combina a capacidade de inferência do LLM com os conhecimentos próprios do *designer* de aprendizagem, com o objetivo de definir variáveis de coleta alinhadas aos objetivos de aprendizagem do jogo, além de contribuir com o aprimoramento do *game design*.

Por mais que a ferramenta ainda não tenha mitigado o problema relacionado aos desafios em modelagem de dados, apresentou resultados que indicam seu potencial para auxiliar *designers* de aprendizagem na definição e organização de variáveis de GLA, bem como contribuiu como *scaffolding* (andaime cognitivo) para apoiar a reflexão e o protagonismo do usuário. Diante disso, considera-se que a integração com LLMs foi positiva, sendo uma abordagem que se pretende aprofundar em estudos futuros.

8.1 Contribuições

As principais contribuições desta dissertação são apresentadas a seguir.

- Amplia e centraliza as discussões sobre a modelagem de dados para GLA, uma etapa contemplada de forma limitada na literatura, sobretudo quanto à definição pedagógica e semântica das variáveis que serão coletadas.
- Caracteriza empiricamente as dificuldades enfrentadas por *designers* de aprendizagem na modelagem de dados e na interação com LLMs, produzindo subsídios para futuras abordagens de apoio à implementação de GLA.
- Avança a investigação sobre o uso de LLMs em GLA, uma relação ainda não consolidada na literatura, ao produzir evidências empíricas sobre o potencial e as limitações desses modelos no apoio à modelagem de dados.
- Evidencia o uso de questionamentos como uma estratégia de suporte gradual (*scaffolding*) à modelagem de dados para GLA, orientando o *designer* de aprendizagem, estimulando sua reflexão sobre o jogo e contextualizando o LLM.
- Amplia a perspectiva sobre os *stakeholders* de GLA ao incluir estudantes como *designers* de aprendizagem, reconhecendo sua participação ativa na definição dos

dados, e não somente como jogadores cujas interações são coletadas e analisadas.

- Apresenta uma abordagem acessível de modelagem de dados para públicos com diferentes níveis de experiência, contribuindo para ampliar e democratizar o acesso às práticas de GLA entre estudantes, educadores e desenvolvedores.

O GLORIA orienta a definição e a organização de variáveis de GLA no *template* do GLBoard para registrar *logs* que auxiliem a identificar evidências de aprendizagem. Entretanto, para avaliar se as variáveis realmente refletem o processo de aprendizagem, ainda são necessários direcionamentos futuros que contemplam questões em aberto e avanços necessários no campo de GLA, tais como:

- Quais elementos de *game design* e de aprendizagem são indispensáveis para a geração de variáveis de GLA?
- As métricas atualmente implementadas no GLBoard são adequadas para avaliar aspectos de aprendizagem?
- O *template* de dados do GLBoard contempla de forma suficiente as variáveis e relações necessárias para o cálculo das métricas?
- Como verificar empiricamente se as variáveis de GLA propostas representam, de fato, indícios de aprendizagem em diferentes tipos de jogos educacionais?
- De que forma as variáveis capturadas podem ser correlacionadas com dados fisiológicos, comportamentais, avaliações heurísticas ou outros instrumentos e fontes de dados para enriquecer a análise de aprendizagem?

8.2 Limitações

A pesquisa, embora alinhada à DSR, apresenta as seguintes limitações:

- **Perfil homogêneo dos participantes:** a maioria dos estudantes que participaram das avaliações pertencia à área de Computação, o que dificulta a generalização dos achados para *designers* de aprendizagem de outras áreas do conhecimento. Ainda

assim, suas contribuições permitiram identificar dificuldades, inconsistências e sugestões que auxiliaram no refinamento da proposta.

- **Ausência de estudo comparativo:** embora o GLORIA tenha sido avaliado por *designers* de aprendizagem, não foi conduzido um estudo comparativo para verificar sua eficácia em relação a outras abordagens de modelagem de dados. Essa limitação será mitigada em investigações futuras.

- **Avaliação das modelagens:** o modelo "*Player Level Up*" foi concebido de forma autoral e validado por participantes do mesmo grupo de pesquisa que o mestrando. Entretanto, todos eram especialistas em GLA e o modelo é específico para o *template* de dados do GLBoard. Além disso, a avaliação sistemática da modelagem de dados ainda não é consolidada na literatura.

- **Indícios de aprendizagem:** o GLORIA auxilia na concepção de variáveis para a identificação de evidências de aprendizagem. Entretanto, para avaliar sua aderência efetiva aos objetivos de aprendizagem, é necessário implementá-las no jogo, registrar os *logs* e realizar análises. Essa etapa não foi contemplada nesta dissertação e depende da integração com o GLBoard, que será realizada futuramente.

- **Implementação de GLA:** a ferramenta auxilia a modelar os dados e contribui com os primeiros passos de GLA. Contudo, inconsistências foram identificadas em relação ao acesso do LLM ao *template* de dados do GLBoard, o que pode impactar a orientação fornecida em casos específicos. Essas inconsistências serão corrigidas em versões futuras da ferramenta.

8.3 Trabalhos futuros

A pesquisa, além da caracterização do estado da arte a respeito da modelagem de dados para jogos educacionais e uso de LLMs para GLA, envolveu a projeção, desenvolvimento e avaliação da ferramenta GLORIA. Entretanto, identificou-se ressalvas na ferramenta que necessitam de refinamentos. Nesse sentido, em conformidade com a iteração do Ciclo de *Design*, pretende-se conduzir essas melhorias em versões futuras

da ferramenta, tais como: (i) padronização estética; (ii) correções de inconsistências técnicas e comportamentais do LLM; (iii) atualização dos modos "Tutorial" e "Livre", que atualmente compartilham do mesmo *prompt* e diferem-se apenas pela mensagem enviada no início da conversa no "Modo Tutorial"; (iv) clareza do fluxo e navegação; e (v) acesso ao *template* de dados do GLBoard.

Pretende-se também incorporar novas funcionalidades ao GLORIA, tais como: (i) configurações do *chatbot*, permitindo a escolha do tipo de mensagem (detalhada ou objetiva), modo (claro ou escuro) e memória; (ii) adição de arquivos no cadastro do jogo e interação com o *chatbot*, que envolverá a escolha de um novo modelo com capacidades multimodais e implementação de técnicas de RAG; (iii) estratégia que compreenda a modelagem dos dados para jogos com fases e mecânicas distintas; e (iv) opção para cadastrar manualmente as variáveis de GLA na modelagem de dados, que automaticamente atualize a estrutura JSON e o conhecimento do *chatbot*. Além disso, investigações serão conduzidas para realizar o treinamento (*fine-tuning*) do modelo com dados de GLA. A incorporação do GLORIA ao GLBoard encontra-se em andamento e não foi finalizada devido ao processo atual de refatoração do próprio GLBoard. Desse modo, as correções identificadas na ferramenta, bem como as funcionalidades futuras, serão desenvolvidas diretamente na nova versão do sistema.

8.4 Produção científica

Esta seção refere-se aos trabalhos aceitos, publicados e em avaliação, em conferências e periódicos, que estão vinculadas a esta dissertação de Mestrado, conforme indicado na Tabela 14. A última coluna associa a publicação ao respectivo ciclo da DSR.

Tabela 14 – Lista de publicações associadas.

Nº	**Título**	**Status**	**Ano**	**Local de publicação**	**Qualis**	**DOI**	**Relação com a dissertação**
1	Um estudo de caso para a implementação de *Game Learning Analytics* (GLA) no desenvolvimento de jogos educacionais	Publicado	2023	WAPLA (CBIE)	N/A	<10.5753/wapla.2023.236221>	Ciclo de Rigor (base de conhecimento sobre uso de GLA)
2	Desafios na implementação de técnicas de GLA em um jogo educacional de algoritmos: um estudo de caso	Publicado	2024	SBGames	A4	<10.5753/sbgames.2024.241240>	Ciclo de Rigor (base de conhecimento sobre desafios em GLA)
3	*Building a specialist agent to assist in the implementation of Game Learning Analytics techniques*	Publicado	2024	WIP-SBIE (CBIE)	A3	<10.5753/sbie.2024.244936>	Ciclo de *Design* (Proposta de agente)
4	*Challenges in educational game data modeling from the perspective of computing students: an empirical study*	Publicado	2025	WEI (CSBC)	A4	<10.5753/wei.2025.9175>	Ciclo de Relevância (Análise de dificuldades de modelagem)
5	*ChatGPT, Gemini or DeepSeek? An empirical study in Game Learning Analytics*	Publicado	2025	SBGames	A3	<10.5753/sbgames.2025.10316>	Ciclo de Rigor (base de conhecimento sobre uso de GLA)
6	*Evaluation of an specialist agent in Game Learning Analytics by learning designers: a case study*	Publicado	2025	SBIE (CBIE)	A3	<10.5753/sbie.2025.12894>	Ciclo de *Design* (Avaliação de agente em GLA)
7	*Prompting LLMs for Game Learning Analytics: A Case Study with Computer Science Students*	Aceito	2026	EduComp	B3	<10.5753/educomp.2026.18681>	Ciclo de Relevância (uso de técnica interativa de *prompt* para GLA)
8	*Contributions from Computer Education Students to Game Learning Analytics: Insights from a Participatory Design*	Aceito	2026	WEI (CSBC)	A4	<10.5753/wei.2026.22565>	Ciclo de Relevância (*Design* Participativo)

Referências

ABT. *Serious Games*. New York: Viking Press, 1970. 36

Almasre, M. Development and evaluation of a custom gpt for the assessment of students' designs in a typography course. *Education Sciences*, MDPI, v. 14, n. 2, p. 148, 2024. 90

Alonso-Fernandez, C. et al. Systematizing game learning analytics for serious games. In: IEEE. *2017 IEEE global engineering education conference (EDUCON)*. [S.l.], 2017. p. 1111–1118. 38, 39, 40, 41, 42, 162, 163

Alonso-Fernández, C. et al. Applications of data science to game learning analytics data: A systematic literature review. *Computers & Education*, Elsevier, v. 141, p. 103612, 2019. 29

Alonso-Fernández, C. et al. Data science meets standardized game learning analytics. In: IEEE. *2021 IEEE Global Engineering Education Conference (EDUCON)*. [S.l.], 2021. p. 1546–1552. 29, 39, 41, 42, 45

Alonso-Fernández, C. et al. Game learning analytics:: Blending visual and data mining techniques to improve serious games and to better understand player learning. *Journal of Learning Analytics*, v. 9, n. 3, p. 32–49, 2022. 40, 41

Alonso-Fernández, C. et al. Lessons learned applying learning analytics to assess serious games. *Computers in Human Behavior*, Elsevier, v. 99, p. 301–309, 2019. 39, 40, 41

Alonso-Fernández, C. et al. Improving evidence-based assessment of players using serious games. *Telematics and Informatics*, Elsevier, v. 60, p. 101583, 2021. 46, 47

Arksey, H.; O'malley, L. Scoping studies: towards a methodological framework. *International journal of social research methodology*, Taylor & Francis, v. 8, n. 1, p. 19–32, 2005. 52

Ausubel, D. P. The psychology of meaningful verbal learning. Grune & Stratton, 1963. 194

Aydın, Ö.; Karaarslan, E. Is chatgpt leading generative ai? what is beyond expectations? *Academic Platform Journal of Engineering and Smart Systems*, Akademik Perspektif Derneği, v. 11, n. 3, p. 118–134, 2023. 48, 49

Banihashem, S. K. et al. Learning analytics for online game-based learning: A systematic literature review. *Behaviour & Information Technology*, Taylor & Francis, v. 43, n. 12, p. 2689–2716, 2024. 28, 40, 41, 54, 163

Basili, V. R.; Rombach, H. D. The tame project: Towards improvement-oriented software environments. *IEEE Transactions on software engineering*, IEEE, v. 14, n. 6, p. 758–773, 1988. 68

Battistella, P.; Wangenheim, C. G. von. Games for teaching computing in higher education–a systematic review. *IEEE Technology and Engineering Education*, sn, v. 9, n. 1, p. 8–30, 2016. 38

Brown, S. Software architecture for developers. *Coding the Architecture*, 2013. 118

Brown, T. et al. Language models are few-shot learners. *Advances in neural information processing systems*, v. 33, p. 1877–1901, 2020. 49, 51

Cano, A. R.; Fernández-Manjón, B.; García-Tejedor, Á. J. Glaid: Designing a game learning analytics model to analyze the learning process in users with intellectual disabilities. In: SPRINGER. *Serious Games, Interaction and Simulation: 6th International Conference, SGAMES 2016, Porto, Portugal, June 16-17, 2016, Revised Selected Papers 6*. [S.l.], 2017. p. 45–52. 28, 29, 43, 46, 47, 48, 65

Charsky, D. From edutainment to serious games: A change in the use of game characteristics. *Games and culture*, Sage Publications Sage CA: Los Angeles, CA, v. 5, n. 2, p. 177–198, 2010. 36

Chung, G. K. Guidelines for the design and implementation of game telemetry for serious games analytics. In: *Serious games analytics: Methodologies for performance measurement, assessment, and improvement*. [S.l.]: Springer, 2015. p. 59–79. 28, 46

Clavié, B.; Gal, K. Edubert: Pretrained deep language models for learning analytics. *arXiv preprint arXiv:1912.00690*, 2019. 51, 52

Cox, K.; Imrie, B.; Miller, A. *Student assessment in higher education: a handbook for assessing performance*. [S.l.]: Routledge, 2014. 38

Csikszentmihalyi, M.; Csikzentmihaly, M. *Flow: The psychology of optimal experience*. [S.l.]: Harper & Row New York, 1990. v. 1990. 27, 36

Dam, S. K. et al. A complete survey on llm-based ai chatbots. *arXiv preprint arXiv:2406.16937*, 2024. 29

Davalos, E. et al. Llms as educational analysts: Transforming multimodal data traces into actionable reading assessment reports. In: SPRINGER. *International Conference on Artificial Intelligence in Education*. [S.l.], 2025. p. 191–204. 30

DemandSage. *ChatGPT Statistics*. 2026. Accessed: 2026-05-26. Disponível em: <https://www.demandsage.com/chatgpt-statistics/>. 49

Din, S. U.; Baig, M. Z.; Khan, M. K. Serious games: An updated systematic literature review. *arXiv preprint arXiv:2306.03098*, 2023. 37

Djaouti, D. et al. Origins of serious games. *Serious games and edutainment applications*, Springer, p. 25–43, 2011. 36

Dong, Q. et al. A survey on in-context learning. *arXiv preprint arXiv:2301.00234*, 2022. 51

Drachen, A.; El-Nasr, M. S.; Canossa, A. *Game analytics: maximizing the value of player data*. [S.l.]: Springer, 2013. 39

Earp, J. Game making for learning: A systematic review of the research literature. *ICERI2015 Proceedings*, IATED, p. 6426–6435, 2015. 29, 47

Eck, R. V. Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE review*, v. 41, n. 2, p. 16, 2006. 39

Elsevier. *Scopus: Access and Analyze Science with Scopus*. 2024. Accessed: 2024-12-19. Disponível em: <https://www.elsevier.com/products/scopus>. 53

Folkestad, J. E. et al. Analytics-driven design: Impact and implications of team member psychological perspectives on a serious games (sgs) design framework. In: *Serious Games Analytics: Methodologies for Performance Measurement, Assessment, and Improvement*. [S.l.]: Springer, 2015. p. 275–300. 29

Freire, M. et al. Game learning analytics: Learning analytics for serious games. In: *Learning, design, and technology*. [S.l.]: Springer Nature Switzerland AG, 2016. p. 1–29. 17, 18, 27, 40, 41, 165, 166, 175

Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association*, Taylor & Francis, v. 32, n. 200, p. 675–701, 1937. 69, 75

Gasparini, I.; Carvalho, M.; Hounsell, M. Jogos digitais para alfabetização matemática: Um mapeamento sistemático da produção brasileira. *SBGames: Simpósio Brasileiro de Jogos e Entretenimento Digital*, p. 430–437, 2015. 37

Gee, J. P. What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, ACM New York, NY, USA, v. 1, n. 1, p. 20–20, 2003. 38

Geremias, M. S. et al. O uso de game learning analytics em jogos digitais educacionais: Um mapeamento sistemático da literatura. *Simpósio Brasileiro De Informática Na Educação (SBIE)*, SBC, p. 737–749, 2024. 163, 166

Goslen, A. et al. Llm-based student plan generation for adaptive scaffolding in game-based learning environments. *International journal of artificial intelligence in education*, Springer, v. 35, n. 2, p. 533–558, 2025. 56

Gros, B. Digital games in education: The design of games-based learning environments. *Journal of research on technology in education*, Taylor & Francis, v. 40, n. 1, p. 23–38, 2007. 37

Gundersen, S. W.; Lampropoulos, G. Using serious games and digital games to improve students' computational thinking and programming skills in k-12 education: A systematic literature review. *Technologies*, MDPI, v. 13, n. 3, p. 113, 2025. 27

Hauge, J. B. et al. Implications of learning analytics for serious game design. In: IEEE. *2014 IEEE 14th international conference on advanced learning technologies*. [S.l.], 2014. p. 230–232. 29, 45, 46, 175

Hevner, A. R. A three cycle view of design science research. *Scandinavian journal of information systems*, v. 19, n. 2, p. 4, 2007. 31, 32, 176

Honda, F. et al. Um estudo de caso para a implementação de game learning analytics (gla) no desenvolvimento de jogos educacionais. In: SBC. *Workshop de Aplicações Práticas de Learning Analytics em Instituições de Ensino no Brasil (WAPLA)*. [S.l.], 2023. p. 138–146. 41, 44

Honda, F. et al. Evaluation of a specialist agent in game learning analytics by learning designers: a case study. In: SBC. *Simpósio Brasileiro de Informática na Educação (SBIE)*. [S.l.], 2025. p. —. 30

Honda, F. et al. Cadê minha pizza? exercitando matemática e pensamento computacional de forma lúdica por meio de grafos. In: SBC. *Anais Estendidos do XII Congresso Brasileiro de Informática na Educação*. [S.l.], 2023. p. 129–132. 44, 70

Honda, F. et al. Challenges in educational game data modeling from the perspective of computing students: an empirical study. In: SBC. *Workshop sobre Educação em Computação (WEI)*. [S.l.], 2025. p. 1251–1262. 48, 175

Hsiao, H.-C. A brief review of digital games and learning. In: IEEE. *2007 first IEEE international workshop on digital game and intelligent toy enhanced learning (DIGITEL'07)*. [S.l.], 2007. p. 124–129. 37

IBM. *AI Hallucinations*. 2023. <https://www.ibm.com/topics/ai-hallucinations>. Accessed: 2025-04-30. Disponível em: <https://www.ibm.com/topics/ai-hallucinations>. 85, 169

IBM. *Natural Language Processing (NLP)*. 2023. <https://www.ibm.com/think/topics/natural-language-processing?mhsrc=ibmsearch_a&mhq=what%20is%20natural%20language%20processing>. Accessed: 2025-01-06. 49

Jovanovic, M.; Campbell, M. Generative artificial intelligence: Trends and prospects. *Computer*, IEEE Computer Society, v. 55, n. 10, p. 107–112, 2022. 29, 48

Kaplan, K. *10 Usability Heuristics Applied to Complex Applications*. 2021. Accessed: 2026-02-07. Disponível em: <https://www.nngroup.com/articles/usability-heuristics-complex-applications/>. 134

Kasneci, E. et al. Chatgpt for good? on opportunities and challenges of large language models for education. Elsevier, v. 103, p. 102274, 2023. 29, 49, 167, 175

Kerr, D.; Chung, G. K. Identifying key features of student performance in educational video games and simulations through cluster analysis. *Journal of Educational Data Mining*, ERIC, v. 4, n. 1, p. 144–182, 2012. 29, 46, 47, 48

Kitto, K. et al. Creating data for learning analytics ecosystems. Society for Learning Analytics Research, 2020. 29, 45

Kıyak, Y. S.; Kononowicz, A. A. Case-based mcq generator: a custom chatgpt based on published prompts in the literature for automatic item generation. *Medical teacher*, Taylor & Francis, p. 1–3, 2024. 90, 170

Li, M. et al. In-context learning with many demonstration examples. *arXiv preprint arXiv:2302.04931*, 2023. 51

Lin, T. et al. A survey of transformers. *AI open*, Elsevier, v. 3, p. 111–132, 2022. 49

Liu, N. F. et al. Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, MIT Press One Broadway, 12th Floor, Cambridge, Massachusetts 02142, USA . . . , v. 12, p. 157–173, 2024. 85, 169

Liu, P. et al. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, ACM New York, NY, v. 55, n. 9, p. 1–35, 2023. 71

Liu, T.-Y. Using educational games and simulation software in a computer science course: Learning achievements and student flow experiences. *Interactive Learning Environments*, Taylor & Francis, v. 24, n. 4, p. 724–744, 2016. 27, 37

Liu, X. et al. Integrating large language models and machine learning to detect struggle in educational games. In: SPRINGER. *International Conference on Artificial Intelligence in Education*. [S.l.], 2025. p. 398–405. 55, 171

Lo, L. S. The art and science of prompt engineering: a new literacy in the information age. *Internet Reference Services Quarterly*, Taylor & Francis, v. 27, n. 4, p. 203–210, 2023. 51

Loh, C. S.; Sheng, Y.; Ifenthaler, D. Serious games analytics: Theoretical framework. *Serious games analytics: Methodologies for performance measurement, assessment, and improvement*, Springer, p. 3–29, 2015. 163

Macedo, F. H. A utilização de jogos no processo de ensino da língua portuguesa. *RCMOS-Revista Científica Multidisciplinar O Saber*, v. 1, n. 1, 2024. 37

Macena, J. et al. Desafios na implementação de técnicas de gla em um jogo educacional de algoritmos: um estudo de caso. In: SBC. *Anais Estendidos do XXIII Simpósio Brasileiro de Jogos e Entretenimento Digital*. [S.l.], 2024. 46, 65, 165

Mangaroska, K.; Giannakos, M. Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, IEEE, v. 12, n. 4, p. 516–534, 2018. 40

Mann, B. et al. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*, 2020. 51

Mazzullo, E. et al. Learning analytics in the era of large language models. *Analytics*, MDPI, v. 2, n. 4, p. 877–898, 2023. 51, 52, 67

Melo, D. et al. Uma estratégia de game learning analytics para avaliar level design em um jogo educacional. In: SBC. *Anais do XXXI Simpósio Brasileiro de Informática na Educação*. [S.l.], 2020. p. 622–631. 27, 28, 38, 39, 41, 42

Melo, D. S. d.; Pires, F. G. d. S.; Melo, R. Robô euroi: jogo de estratégia matemática para exercitar o pensamento computacional. In: SBC. *Anais Estendidos do VIII Congresso Brasileiro de Informática na Educação*. [S.l.], 2018. p. 685–694. 37

Melo, R.; Silva, D.; Pires, F. Stardust: um serious game para a aprendizagem implícita de grafos. In: *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*. [S.l.: s.n.], 2019. v. 8, n. 1, p. 1237. 36

Mercer, T. G. et al. The use of educational game design and play in higher education to influence sustainable behaviour. *International Journal of Sustainability in Higher Education*, Emerald Publishing Limited, v. 18, n. 3, p. 359–384, 2017. 29, 47

Minaee, S. et al. Large language models: A survey. *arXiv preprint arXiv:2402.06196*, 2024. 49

Misiejuk, K. et al. Mapping the landscape of generative artificial intelligence in learning analytics: A systematic literature review. *Journal of Learning Analytics*, p. 1–20, 2025. 30, 48, 51, 52, 67, 167, 175

Mitsea, E.; Drigas, A.; Skianis, C. A systematic review of serious games in the era of artificial intelligence, immersive technologies, the metaverse, and neurotechnologies: Transformation through meta-skills training. *Electronics*, MDPI, v. 14, n. 4, p. 649, 2025. 167

Moran, K.; Gordon, K. *How to Conduct a Heuristic Evaluation*. 2023. Disponível em: <https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/>. 132

Naveed, H. et al. A comprehensive overview of large language models. *arXiv preprint arXiv:2307.06435*, 2023. 49

Nexocode. *Customizing Large Language Models: A Comprehensive Guide*. 2023. Accessed: 2024-06-17. Disponível em: <https://nexocode.com/blog/posts/customizing-large-language-models-a-comprehensive-guide/>. 50

Nielsen, J. *10 Usability Heuristics for User Interface Design*. 2024. Accessed: 2026-02-07. Disponível em: <https://www.nngroup.com/articles/ten-usability-heuristics/>. 134

Nielsen, J.; Molich, R. Heuristic evaluation of user interfaces. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. [S.l.: s.n.], 1990. p. 249–256. 132

Oliveira, M. G.; Santos, R. F.; Pereira, L. L. Jogos digitais educacionais: um mapeamento sistemático da literatura brasileira. *Revista Brasileira de Informática na Educação*, SBIE, v. 27, n. 1, p. 123–145, 2019. 163

Oliveira, R. et al. Avaliações em jogos educacionais: instrumentos de avaliação da reação, aprendizagem e comparação de jogos. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2019. v. 30, n. 1, p. 972. 38, 48

Ollama. *Ollama*. 2023. Acesso em: 4 jul. 2026. Disponível em: <https://ollama.com/>. 119

OpenAI. *Introducing GPTs*. 2023. <https://openai.com/index/introducing-gpts/>. Accessed: 2024-08-12. 89

OpenAI. *Introducing Deep Research*. 2025. <https://openai.com/index/introducing-deep-research/>. Accessed: 2025-11-06. 54

Ouyang, L. et al. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, v. 35, p. 27730–27744, 2022. 49

Owen, V. E. Learning analytics for educational game design: Mapping methods to development. In: *Proceedings of the Games+Learning+Society Conference*. Madison, WI: Games+Learning+Society Center, University of Wisconsin-Madison, 2014. p. 175–181. Games+Learning+Society Conference (GLS). 162

Parthasarathy, V. B. et al. The ultimate guide to fine-tuning llms from basics to breakthroughs: An exhaustive review of technologies, research, best practices, applied research challenges and opportunities. *arXiv preprint arXiv:2408.13296*, 2024. 51

Pelizzari, A. et al. Teoria da aprendizagem significativa segundo ausubel. *revista PEC*, v. 2, n. 1, p. 37–42, 2002. 194

Peng, X. et al. Player-driven emergence in llm-driven game narrative. In: IEEE. *2024 IEEE Conference on Games (CoG)*. [S.l.], 2024. p. 1–8. 30, 52

Perez-Colado, I. et al. Game learning analytics is not informagic! In: IEEE. *2018 IEEE Global Engineering Education Conference (EDUCON)*. [S.l.], 2018. p. 1729–1737. 28, 29, 43, 45, 46, 47, 48, 162, 163, 175

Pessoa, M. S. P. Codeplay: uma plataforma que incorpora a ludicidade de jogos de entretenimento a um juiz on-line. Universidade Federal do Amazonas, 2022. 32

Petersen, S. A.; Bedek, M. A. Challenges and opportunities in evaluating learning in serious games: A look at behavioural aspects. In: SPRINGER. *International Conference on Serious Games Development and Applications*. [S.l.], 2012. p. 219–230. 29, 46, 47, 175

Petri, G.; Wangenheim, C. G. von. How games for computing education are evaluated? a systematic literature review. *Computers & education*, Elsevier, v. 107, p. 68–90, 2017. 37

Piacentini, M. T. Jogo eletrônico, flow e cognição. 2011. 27, 37

Pimentel, M.; Filippo, D.; Santoro, F. M. Design science research: fazendo pesquisas científicas rigorosas atreladas ao desenvolvimento de artefatos computacionais projetados para a educação. *Metodologia de Pesquisa em Informática na Educação: Concepção da Pesquisa. Porto Alegre: SBC*, p. 5–29, 2019. 31, 32

Pires, F. G. d. S. et al. Thinkted lab, um caso de aprendizagem criativa em computação no nível superior. Universidade Federal do Amazonas, 2021. 17, 32, 79, 108, 109, 129, 176

Plass, J. L.; Homer, B. D.; Kinzer, C. K. Foundations of game-based learning. *Educational psychologist*, Taylor & Francis, v. 50, n. 4, p. 258–283, 2015. 27, 28, 36, 37, 175

Qian, M.; Clark, K. R. Game-based learning and 21st century skills: A review of recent research. *Computers in human behavior*, Elsevier, v. 63, p. 50–58, 2016. 37

Raffel, C. et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, v. 21, n. 1, p. 5485–5551, 2020. 49

Rodríguez-Ortiz, M. Á.; Santana-Mancilla, P. C.; Anido-Rifón, L. E. Machine learning and generative ai in learning analytics for higher education: A systematic review of models, trends, and challenges. *Applied Sciences*, MDPI, v. 15, n. 15, p. 8679, 2025. 30, 51, 52

Santilario-Berthilier, J. et al. Simva: Serious games learning analytics based on standards and open code. In: *CEUR Workshop Proceedings*. [S.l.: s.n.], 2025. v. 3995, p. 241–250. 28, 42

Santos, A. L. d. et al. A systematic mapping study on game elements and serious games for learning programming. In: SPRINGER. *Computer Supported Education: 10th International Conference, CSEDU 2018, Funchal, Madeira, Portugal, March 15–17, 2018, Revised Selected Papers 10*. [S.l.], 2019. p. 328–356. 37

Sathe, T. S. et al. How i gpt it: Development of custom artificial intelligence (ai) chatbots for surgical education. *Journal of Surgical Education*, Elsevier, v. 81, n. 6, p. 772–775, 2024. 90

Saveski, G. L. et al. What serious game studios want from ict research: identifying developers' needs. In: SPRINGER. *Games and Learning Alliance: 4th International Conference, GALA 2015, Rome, Italy, December 9-11, 2015, Revised Selected Papers 4*. [S.l.], 2016. p. 32–41. 41

Savi, R.; Ulbricht, V. R. Jogos digitais educacionais: benefícios e desafios. *Revista Novas Tecnologias na Educação*, v. 6, n. 1, 2008. 27, 37

Scheneider, T.; Lemos, R. The use of learning analytics interactive dashboards in serious games: a review of the literature. *International journal for innovation education and research*, v. 8, n. 3, p. 150–174, 2020. 28, 40

Schuler, D.; Namioka, A. *Participatory design: Principles and practices*. [S.l.]: CRC press, 1993. 105

Sclater, N. *Learning analytics explained*. [S.l.]: Routledge, 2017. 39

Sengar, S. S. et al. Generative artificial intelligence: a systematic review and applications. *Multimedia Tools and Applications*, Springer, p. 1–40, 2024. 48

Serrano-Laguna, Á. et al. Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces*, Elsevier, v. 50, p. 116–123, 2017. 42, 96

Shin, J. et al. Prompt engineering or fine tuning: An empirical assessment of large language models in automated software engineering tasks. *arXiv preprint arXiv:2310.10508*, 2023. 71

Shoukry, L.; Göbel, S.; Steinmetz, R. Learning analytics and serious games: Trends and considerations. In: *Proceedings of the 2014 ACM International Workshop on Serious Games*. [S.l.: s.n.], 2014. p. 21–26. 28, 45, 46, 162

Siemens, G.; Long, P. Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, ERIC, v. 46, n. 5, p. 30, 2011. 39

Silva, D. et al. Avaliacão de objetos digitais de aprendizagem: como os licenciados em computação analisam jogos educacionais? *Revista Novas Tecnologias na Educação*, v. 19, n. 2, p. 111–121, 2021. 27, 38, 40, 41, 175

Silva, D. et al. Glboard: um sistema para auxiliar na captura e análise de dados em jogos educacionais. In: SBC. *Anais Estendidos do XXI Simpósio Brasileiro de Jogos e Entretenimento Digital*. [S.l.], 2022. p. 959–968. 28, 42, 43, 70, 175

Someren, M. W. V. et al. The think aloud method: a practical approach to modelling cognitive processes. *London: AcademicPress*, v. 11, n. 6, 1994. 133, 143

Suzgun, M. et al. Challenging big-bench tasks and whether chain-of-thought can solve them. *arXiv preprint arXiv:2210.09261*, 2022. 216

Tarouco, L. M. R. et al. Jogos educacionais. *RENOTE: revista novas tecnologias na educação [recurso eletrônico]. Porto Alegre, RS*, 2004. 37, 175

Tlili, A. et al. A systematic literature review of empirical studies on learning analytics in educational games. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2021. 29

Vaishnavi, V. K.; Kuechler, W. *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*. 2nd. ed. Boca Raton, FL: CRC Press, 2015. ISBN 9781498715263. 32, 176

Vaswani, A. Attention is all you need. *Advances in Neural Information Processing Systems*, 2017. 49

Wallner, G.; Kriglstein, S. Visualization-based analysis of gameplay data–a review of literature. *Entertainment Computing*, Elsevier, v. 4, n. 3, p. 143–155, 2013. 40

Wang, M. et al. Boosting student engagement in stem: Integrating large language model-based virtual agents into alternate reality games. *Journal of Computer Assisted Learning*, Wiley Online Library, v. 41, n. 6, p. e70139, 2025. 54, 170

Wang, Y. et al. Application of serious games in health care: scoping review and bibliometric analysis. *Frontiers in Public Health*, Frontiers Media SA, v. 10, p. 896974, 2022. 36

Wei, J. et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, v. 35, p. 24824–24837, 2022. 78, 79, 167

Weitze, C. L. An experiment on how adult students can learn by designing engaging learning games. In: *14th European Conference on e-Learning*. [S.l.: s.n.], 2014. 29, 47

Wohlin, C. et al. *Experimentation in software engineering*. [S.l.]: Springer, 2012. v. 236. 68

Xue, L. et al. mt5: A massively multilingual pre-trained text-to-text transformer. *arXiv preprint*, arXiv:2010.11934, 2020. <https://arxiv.org/abs/2010.11934>. 49

Yang, D.; Kleinman, E.; Harteveld, C. Gpt for games: An updated scoping review (2020-2024). *IEEE Transactions on Games*, IEEE, 2025. 30, 48, 52, 175

Yin, S. et al. A survey on multimodal large language models. *arXiv preprint arXiv:2306.13549*, 2023. 50

Zhang, S. et al. Instruction tuning for large language models: A survey. *arXiv preprint arXiv:2308.10792*, 2023. 50

Zhou, J. et al. Instruction-following evaluation for large language models. *arXiv preprint arXiv:2311.07911*, 2023. 216

Zyda, M. From visual simulation to virtual reality to games. *Computer*, IEEE, v. 38, n. 9, p. 25–32, 2005. 36

A

Descrição do Jogo "Cadê minha pizza?"

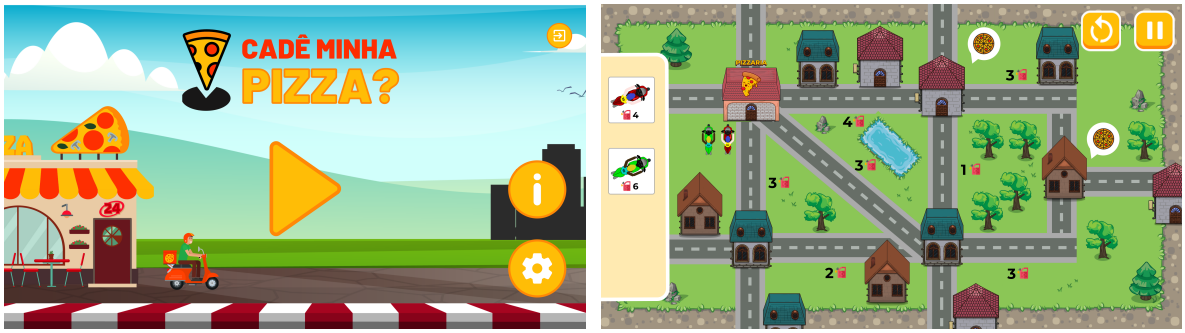
"Cadê minha Pizza?" é um jogo educacional *mobile* 2D, em visão *top-down*, cujo objetivo é exercitar operações básicas de Matemática e introduzir, de forma implícita, conceitos de caminho mínimo em grafos. Para isso, o jogador auxilia uma pizzaria recém-inaugurada a realizar entregas pela cidade, escolhendo as melhores rotas e os entregadores mais adequados de acordo com a quantidade limitada de gasolina disponível. A Figura 44 apresenta ilustrações do jogo: a tela inicial, na Figura 44(a), e a segunda fase, na Figura 44(b).

- **Público-alvo:** estudantes finalistas do Ensino Médio e iniciantes em cursos de Computação.
- **Conteúdo educacional:** operações básicas de Matemática e caminho mínimo, conteúdo relacionado à Teoria dos Grafos.
- **História:** o jogo narra a história de um grupo de amigos recém-chegados a uma grande cidade metropolitana, que finalmente realizaram o sonho de fundar uma pizzaria. No entanto, a abertura do negócio gerou dívidas, fazendo com que cada personagem assumisse uma função: alguns atuam como entregadores, outros cuidam da gerência, das finanças e da produção das pizzas, enquanto o jogador auxilia na escolha das melhores rotas de entrega.
- ***Gameplay*:** o jogador deve realizar entregas de pizzas em diferentes estabelecimentos da cidade, considerando que cada entregador possui uma quantidade

limitada de gasolina. Para isso, ele visualiza o mapa, analisa os locais, ruas e custos dos trajetos, seleciona um entregador por meio de toques na tela e o envia para o destino escolhido. Para avançar nas fases, todas as entregas precisam ser concluídas.

- **Mecânica de aprendizagem:** cada fase representa uma cidade organizada como um grafo ponderado, em que a pizzaria corresponde ao vértice inicial, os estabelecimentos aos vértices, as ruas às arestas e os custos de gasolina aos pesos. Para vencer, o jogador precisa calcular os custos das rotas, comparar os valores encontrados e escolher o entregador mais adequado, exercitando implicitamente percurso em grafos, caminho mínimo e alocação de recursos por meio de operações matemáticas básicas.

- **Teoria de aprendizagem:** o jogo é embasado na Teoria da Aprendizagem Significativa de David Ausubel, segundo a qual novos conhecimentos são aprendidos de forma mais efetiva quando se relacionam aos conhecimentos prévios do estudante (Ausubel, 1963; Pelizzari et al., 2002). Em "Cadê minha Pizza?", essa relação ocorre porque o jogador utiliza conhecimentos matemáticos já conhecidos, como soma, comparação de números e identificação de maior/menor valor, para resolver desafios associados implicitamente ao caminho mínimo em grafos. Assim, o conteúdo computacional não é apresentado de forma direta, mas exercitado por meio de uma situação lúdica e contextualizada.



(a) Tela inicial (b) Primeira fase.

Figura 44 – Telas do jogo "Cadê minha pizza?".

B

FORMULÁRIO DE AUTOAVALIAÇÃO DAS MODELAGENS DE DADOS

Descrição: O presente formulário foi construído como parte de uma pesquisa realizada no Mestrado em Informática pelo Programa de Pós-Graduação em Informática (PPGI) da Universidade Federal do Amazonas (UFAM), cujo objetivo é analisar como estudantes de Computação realizam o processo de modelagem de dados de jogos educacionais, visando expressar evolução de aprendizagem do jogador. Como parte dessa pesquisa, será fundamental aplicar o presente questionário para que as planilhas com as modelagens possam ser submetidas e que envolve perguntas sobre o processo realizado. Os dados coletados serão divulgados de forma anônima para publicações científicas a fim de contribuir na questão de pesquisa analisada durante a execução do projeto. Logo, vale ressaltar que os dados serão tratados com a finalidade única para pesquisa.

1. E-mail
2. Você aceita responder o questionário e concorda com a sua finalidade, descrita no parágrafo acima?
   (a) Sim (b) Não
3. Insira aqui sua planilha com a modelagem de dados do jogo educacional
4. Você já havia modelado dados para outro jogo educacional?
   (a) Sim (b) Não
5. Você considera que entendeu o objetivo do processo de modelagem de dados?
   (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

6. O que você considera ter feito bem na modelagem de dados no jogo que você escolheu?

7. Qual nível de dificuldade você sentiu ao modelar os dados?
   (a) 1 (Muito fácil) (b) 2 (c) 3 (d) 4 (e) 5 (Muito difícil)

8. O que você achou mais difícil na modelagem de dados?

9. Você considera que o tempo para modelar os dados do jogo foi suficiente?
   (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

10. Houve algo que você aprendeu durante o processo de modelagem que não conhecia antes? O quê?

11. Você acredita que o processo de modelagem de dados contribui para o desenvolvimento de jogos educacionais mais eficazes?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

12. Como você acha que esse processo contribui para o desenvolvimento de jogos mais eficazes?

13. O processo de modelagem foi semelhante a algo que você já havia feito antes?
    (a) Sim (b) Não

14. Se sim, descreva o que foi similar.

15. Você sentiu que o conhecimento adquirido em disciplinas anteriores ajudou no processo?
    (a) Sim (b) Não

16. Se sim, quais disciplinas contribuíram mais? E como?

17. Se você tivesse que ajudar alguém a modelar dados para outro jogo, como você faria?

C

1º Prompt ao ChatGPT para geração de estruturas de captura

"Olá, Chat-GPT. Para capturar e analisar dados de um jogo educacional, geralmente utilizam-se técnicas de *Game Learning Analytics* (GLA). O GLBoard é uma ferramenta que possibilita esse processo, possuindo uma estrutura própria de captura, em formato JSON. A seguir, essa estrutura é disponibilizada e descrita.

```
"game_data": {
    "date_last_login": "Date", // horário do último login do
                                  jogador
    "number_phases": "int", // quantidade de fases do jogo
    "phases": [{
        "Phase_id": "int", // identificador da fase
        "Sections": [{
            "Conclusion": "string", // status de conclusão da
                                       fase: vitória, derrota
                                       ou  desistência
            "dateTimeFinish": "Date", // horário que o jogador
                                         concluiu aquela seção
            "dateTimeStart": "Date", // horário que o jogador
                                        iniciou aquela seção
            "Performance": "int", // pontuação do jogador
                                     naquela seção
```

```
            "Path_player": [] // caminho do jogador na fase,
                                 podendo aqui o desenvolvedor
                                adicionar novas variáveis
                                sobre o percurso do jogador,
                                estruturando como queira
          } // Outras seções...
        ],
        "Status": "string" // status da fase: finalizada ou
                              não finalizada
      } // Outras fases...
    ],
    "phases_unlocked": "int" // número de fases desbloqueadas
                                pelo jogador
  },
  "player_data": {
    "day_birthday": "Date", // data de aniversário do jogador
    "gender": "string", // gênero do jogador
    "id": "string", // identificador do jogador
    "name": "string" // nome do jogador}
```

Irei lhe fornecer informações sobre um jogo educacional e preciso que você faça o seguinte: (i) compreender como o jogo funciona; (ii) propor novos dados que podem ser interessantes de armazenar; (iii) inserir o nome do dado, o tipo (*int*, *float*, etc.) e a justificativa de captura (comentário); (iv) organizar esses dados na estrutura JSON fornecida anteriormente; (v) inserir, obrigatoriamente, os dados na variável $path\_player$. Como é uma lista de strings, nessa variável podem haver novas variáveis, que serão as que você irá propor. O intuito desses dados novos é auxiliar a identificar se o jogador evoluiu durante o jogo; e (vi) lembrar que o $path\_player$ corresponde ao percurso/caminho do jogador dentro de uma seção na fase, ou seja, o que ele fez e quais suas ações, não fazendo sentido atribuir uma variável no final da fase, e sim, durante. Desse modo, aguarde eu lhe fornecer as informações do jogo para que você possa gerar a estrutura JSON."

D

Atributos de jogos educacionais mapeados para enviar ao ChatGPT

- **Fundamentos:**
  - **Nome do jogo:** Nome concebido ao jogo.
  - **Tema:** O tema do jogo, seja curricular ou de graduação.
  - **Conteúdo:** Conteúdo educacional do jogo, dentro do tema delimitado.
  - **Objetivo educacional:** Qual o objetivo educacional do jogo?
  - **Público-alvo:** Em faixa etária e/ou nível escolar.
  - **Habilidade BNCC:** Quais habilidades da Base Nacional Comum Curricular (BNCC) estão envolvidas?
  - **PCN:** Quais as finalidades e competências/habilidades envolvidas dos Parâmetros Curriculares Nacionais (PCN)?
  - **Aspecto Currículo:** Quais aspectos do Currículo de Referência em Tecnologia e Computação estão envolvidos?
- **Informações do jogo:**
  - **História:** A narrativa/trama que se passa no jogo.
  - **Personagens:** Personagens do jogo e suas características.
  - **Inimigos:** Inimigos no jogo e suas características.
  - **Jornada do herói:** Qual a jornada do herói? Onde começa? O que ele encontra? Qual seu objetivo?

  - **Mundo do jogo:** Como é o mundo do jogo?
  - **Coletáveis:** Quais coletáveis têm no jogo e como estão dispostos? Como funcionam?
  - **Perigos:** Quais os perigos do jogo? O que eles fazem?
  - **Loja:** Como funciona o sistema de loja ou compra no jogo?
  - **Pontuação:** Como o jogador pontua? Como é contado?
  - ***Power-ups*:** Quais os power-ups do jogo? Quais suas características? Como o jogador pode obtê-los?
  - ***Gameplay*:** De forma geral, um resumo da gameplay do jogo.
  - **Recompensas:** Como o jogo recompensa o jogador?

- **Aprendizagem:**
  - ***Level design*:** Como funciona cada fase?
  - **Fluxo de aprendizagem:** Qual a sequência didática dos conteúdos que você trabalha no jogo?
  - **Mecânica de aprendizagem:** Como a pessoa aprende o conteúdo no seu jogo? Descreva bem detalhadamente.
  - ***Learning design*:** O que as pessoas devem aprender em cada fase?
  - **Pensamento Computacional:** Como funciona dentro do jogo?
  - **Teorias de aprendizagem:** Quais estão presentes no jogo e como são trabalhadas?

E

2º Prompt ao ChatGPT para geração de estruturas de captura

"Chat-GPT, para capturar dados de modo furtivo do jogador durante sua gameplay, utilizamos técnicas de *Game Learning Analytics*. Possuímos uma estrutura própria, em formato JSON, para capturar os dados. Vou disponibilizar essa estrutura para você e explicar cada atributo. Aqui a estrutura:[1]

Em seguida, vou fornecer algumas informações sobre determinado jogo educacional e preciso que você faça o seguinte: (i) compreenda o funcionamento deste jogo; (ii) utilizando a estrutura que eu lhe forneci, proponha novos dados do que seria interessante capturar; (iii) insira o nome desse dado, o tipo (*int*, *float*, etc.) e a justificativa (pode ser em comentário); (iv) a estrutura gerada deve ser em formato de JSON; (v) o intuito desses novos dados é auxiliar a identificar se o jogador evoluiu durante o jogo; (vi) na estrutura JSON, os novos dados devem estar OBRIGATORIAMENTE dentro da variável $path\_player$; (vii) como $path\_player$ é uma lista de strings, dentro dessa variável podem haver novas variáveis, que serão as que você vai propor; e (viii) lembre-se de que o $path\_player$ é o PERCURSO/CAMINHO do jogador dentro de uma seção na fase, ou seja, o que ele fez e quais passos ele tomou. Não faz sentido colocar uma variável que só é atribuída no final da fase, e sim, durante. Aguarde eu lhe informar as informações do jogo. Lembrando que você deve retornar a estrutura em formato JSON."

[1] *Template* de dados do GLBoard. A mesma estrutura disponível no Apêndice C.

F

FORMULÁRIO DE AVALIAÇÃO DA CRIAÇÃO DE ESTRUTURA DE CAPTURA DE FORMA MANUAL

Descrição: O presente formulário foi construído como parte de uma pesquisa realizada no Mestrado em Informática pelo Programa de Pós-Graduação em Informática (PPGI) da Universidade Federal do Amazonas (UFAM), cujo objetivo é analisar como estudantes de Computação modelam e atribuem sentido a dados que irão capturar em jogos educacionais para evidenciar evolução de aprendizagem. Como parte dessa pesquisa, será fundamental aplicar o presente questionário, que envolve perguntas voltadas à experiência com a modelagem/definição dos dados e criação da estrutura de captura. Os dados coletados serão divulgados de forma anônima para publicações científicas a fim de contribuir na questão de pesquisa analisada durante a execução do projeto. Logo, vale ressaltar que os dados serão tratados com a finalidade única para pesquisa.

1. E-mail
2. Você aceita responder o questionário e concorda com a sua finalidade, descrita no parágrafo acima?
   (a) Sim (b) Não
3. Nome completo
4. Data de nascimento

5. Curso

   (a) Licenciatura em Computação (b) Sistemas de Informação (c) Engenharia da Computação (d) Outro: ____

6. Período

   (a) 2º Período (b) 4º Período (c) 6º Período (d) 8º Período (e) +8º Período

7. Neste semestre, você está cursando alguma disciplina com a professora Fernanda?

   (a) Sim, Design Instrucional (DI) (b) Sim, Fundamentos de Software Educacional (FSE) (c) Não

8. Insira aqui a estrutura de captura de dados que você criou, com base na modelagem.

   Pode ser uma imagem, print ou texto, desde que esteja legível.

9. Você já havia modelado dados para algum outro jogo?

   (a) Sim (b) Não

10. Qual o nível de dificuldade que você sentiu ao modelar os dados (definir e atribuir sentido)?

    (a) 1 (Muito simples) (b) 2 (c) 3 (d) 4 (e) 5 (Muito difícil)

11. O que você achou mais difícil durante esta etapa de modelagem dos dados?

12. Qual o nível de dificuldade que você sentiu ao criar a estrutura de captura (organizar os dados)?

    (a) 1 (Muito simples) (b) 2 (c) 3 (d) 4 (e) 5 (Muito difícil)

13. O que você achou mais difícil durante esta etapa de criar a estrutura de captura?

14. Se você tivesse que auxiliar alguém a fazer essas etapas, como você faria?

G

Formulário de avaliação da criação de estrutura de captura com o Chat-GPT

Descrição: O presente formulário foi construído como parte de uma pesquisa realizada no Mestrado em Informática pelo Programa de Pós-Graduação em Informática (PPGI) da Universidade Federal do Amazonas (UFAM), cujo objetivo é analisar como estudantes de Computação modelam e atribuem sentido a dados que irão capturar em jogos educacionais para evidenciar evolução de aprendizagem. Como parte dessa pesquisa, será fundamental aplicar o presente questionário, que envolve perguntas relacionadas à criação de estrutura para captura de dados em jogos educacionais utilizando uma *Large Language Models* (LLM), neste caso, o Chat-GPT 3.5. Os dados coletados serão divulgados de forma anônima para publicações científicas a fim de contribuir na questão de pesquisa analisada durante a execução do projeto. Logo, vale ressaltar que os dados serão tratados com a finalidade única para pesquisa.

1. E-mail
2. Você aceita responder o questionário e concorda com sua finalidade, descrita no parágrafo acima?
   (a) Sim (b) Não
3. Nome

4. Insira aqui o(s) link(s) que contêm sua interação com o Chat-GPT para criação da estrutura

5. Insira aqui a estrutura de captura de dados que o Chat-GPT criou, com base nos dados de jogos que você forneceu e os questionamentos
   Pode ser uma imagem, print ou texto, desde que esteja legível.

6. Qual o nível de dificuldade que você sentiu para criar o *prompt* (comando) ao Chat-GPT?
   (a) 1 (Muito simples) (b) 2 (c) 3 (d) 4 (e) 5 (Muito difícil)

7. O que você achou mais difícil durante a criação do *prompt* (comando) ao Chat-GPT?

8. O Chat-GPT ajudou você a entender quais dados são importantes de acordo com seu jogo?
   (a) 1 (Definitivamente não) (b) 2 (c) 3 (d) 4 (e) 5 (Definitivamente sim)

9. Quantas conversas com o GPT você teve que criar até ele retornar um resultado adequado?
   Por exemplo, se o GPT nas primeiras duas vezes retornou a estrutura JSON antes dos questionamentos e só fez corretamente na terceira interação.
   (a) 1 (b) 2 (c) 3 (d) 4 (e) Outro: ____

10. Você acha que os questionamentos que o GPT fez sobre seu jogo ajudaram a criar a estrutura? Por quê?

11. O Chat-GPT ajudou você a entender por que os dados que ele definiu são importantes de acordo com seu jogo?
    (a) 1 (Definitivamente não) (b) 2 (c) 3 (d) 4 (e) 5 (Definitivamente sim)

12. Comparando a estrutura que você mesmo criou com a estrutura que o GPT criou, qual você acha que ficou mais adequada?
    (a) 1 (O que eu criei) (b) 2 (c) 3 (d) 4 (e) 5 (O que o GPT criou)

13. Você sentiu que faltou algum dado importante na estrutura gerada? Qual(is)?

14. O que você achou da abordagem de utilizar o Chat-GPT para gerar a estrutura em comparação com a que você fez sozinho?

15. Como você melhoraria a estrutura que o GPT retornou?

16. Se você tivesse que auxiliar alguém a fazer essa interação com o GPT, como você faria?

H

Formulário de avaliação do Agente Inteligente por especialistas em GLA

Descrição: Obrigado por concluir o teste! Agora que você completou suas interações com o agente, gostaríamos de convidá-lo a avaliar sua experiência. Esta avaliação é crucial para que possamos entender como o "*GLA Specialist*" se comportou em diferentes situações e como ele atendeu às suas necessidades em *Game Learning Analytics* (GLA) e no uso do GLBoard. Por favor, dedique alguns minutos para compartilhar seu *feedback* sobre a precisão das respostas, a utilidade das funcionalidades e qualquer sugestão de melhoria. Sua contribuição é fundamental para o aprimoramento contínuo desta ferramenta.

1. E-mail

2. Insira aqui o(s) link(s) de sua interação com o agente

3. Há quanto tempo você tem experiência com jogos educacionais?
   (a) Menos de 1 ano (b) 2 anos (c) 3 anos (d) 4 anos (e) 5 anos (f) 6 anos (g) 7 anos (h) 8 anos ou mais

4. Há quanto tempo você tem experiência com implementação de *Game Learning Analytics*?
   (a) Menos de 1 ano (b) 2 anos (c) 3 anos (d) 4 anos (e) 5 anos (f) 6 anos (g) 7 anos (h) 8 anos ou mais

5. Em quantos jogos você já implementou GLA?
   (a) Nenhum (b) 1 jogo (c) 2 jogos (d) 3 jogos (e) 5 jogos (f) Mais de 5 jogos

6. O agente forneceu respostas precisas e relevantes às suas perguntas sobre GLA?
   (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

7. O agente forneceu respostas precisas e relevantes às suas perguntas sobre GLBoard?
   (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

8. O agente foi capaz de personalizar a estrutura JSON de acordo com as informações fornecidas sobre o seu jogo educacional?
   (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

9. O agente explicou adequadamente as variáveis propostas no $path\_player$?
   (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

10. Na sua percepção, os exemplos que o agente gerou em relação às estruturas de captura foram adequados?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

11. Você acha que o agente tem potencial para auxiliar na modelagem de dados e implementação técnicas de GLA?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

12. Como você classificaria sua experiência geral ao interagir com o "*GLA Specialist*"?
    (a) Muito boa (b) Boa (c) Neutra (d) Ruim (e) Muito ruim

13. O que você mais gostou do agente?

14. O que você não gostou do agente?

15. Quais melhorias você sugeriria para aprimorar o "*GLA Specialist*"?

16. Houve algo que o agente não conseguiu fazer ou responder conforme suas expectativas? Se sim, por favor, explique.

17. Você acha que o agente é uma boa opção para auxiliar estudantes a entender/implementar GLA? Por quê?

I

Autodescrição atual do Agente Inteligente em GLA

"Eu sou um especialista em *Game Learning Analytics*, representando o ThinkTEd, uma organização com uma história rica em pesquisas sobre *Game Learning Analytics* (GLA) desde 2018. Em 2022, o ThinkTEd inovou no campo de GLA ao criar o GLBoard, uma ferramenta avançada destinada a capturar e analisar dados de jogos educacionais.

Estou aqui para ajudar *designers* de aprendizagem a criar e explicar estruturas em JSON para captura de dados em jogos educacionais, utilizando a ferramenta GLBoard. Meu papel é dividido em etapas:

1. **Modelagem de dados:** Após receber as informações do jogo educacional (nome, público-alvo, tema, conteúdo, história, mecânica de aprendizagem, etc.), minha primeira tarefa é auxiliar na definição dos dados a serem capturados. Para isso, criarei uma tabela contendo:

    - **Nº:** Número sequencial para identificação.
    - **Nome do dado (variável):** Nome claro e descritivo da variável a ser capturada.
    - **Tipo de dado:** O tipo de dado esperado (e.g., *string*, inteiro, booleano).
    - **Exemplo de coleta:** Um exemplo concreto do dado esperado.
    - **Justificativa de captura:** Uma explicação objetiva sobre por que aquele dado é importante para expressar a evolução do jogador e sua possível aprendizagem durante o percurso nas fases do jogo.

Após concluir a modelagem, perguntarei: "Há algo a melhorar ou podemos ir para a geração da estrutura de captura?"

2. **Estrutura de captura em JSON:** Apenas com a aprovação da modelagem de dados, passo para a criação da estrutura JSON padrão do GLBoard, com ênfase especial no campo $path\_player$.

3. **Personalização da estrutura JSON:** Solicito as informações do jogo educacional e, caso o usuário permita, faço questionamentos para compreender melhor as particularidades do jogo. Com base nas respostas, personalizo a estrutura JSON já existente, substituindo o campo $path\_player$ por variáveis ou conjuntos de variáveis que representem o percurso do jogador no jogo.

4. **Geração de exemplos práticos:** Forneço uma versão da estrutura JSON acompanhada de um exemplo ilustrativo, mostrando o que as variáveis podem armazenar. Adiciono comentários explicando a importância de cada dado no contexto do jogo.

5. **Validação da estrutura:** Após fornecer a estrutura e os exemplos, pergunto se a estrutura está adequada na percepção do usuário. Em caso negativo, solicito mais informações ou faço novos questionamentos para ajustes.

Durante todo o processo, destaco que o campo $path\_player$ é uma das variáveis mais importantes, pois contém informações detalhadas sobre o percurso do jogador nas fases do jogo. Quanto mais informações forem disponibilizadas, maior será o detalhamento e a riqueza dessa variável."

J

Roteiro de Teste para avaliação do Agente Inteligente

"**Olá, seja bem-vindo ao teste do agente especialista "*GLA Specialist*"!**

Agradecemos sua participação e disponibilidade para contribuir com este teste. O objetivo deste teste é avaliar a eficácia, a precisão e a utilidade do agente "*GLA Specialist*" na área de *Game Learning Analytics* (GLA). Suas interações e feedback serão fundamentais para aprimorar ainda mais a ferramenta.

O "*GLA Specialist*" foi projetado para ajudar *designers* de aprendizagem a criar e explicar estruturas JSON para captura de dados em jogos educacionais, utilizando a ferramenta GLBoard. Ele pode personalizar essas estruturas com base nas informações que você fornecer sobre o seu jogo, garantindo que a captura de dados seja otimizada para as necessidades específicas de cada projeto.

Neste teste, você deve **interagir** com o agente, testando suas capacidades de responder a perguntas e fornecer informações relevantes relacionadas à GLA e ao GLBoard. Estamos interessados em entender como o agente se comporta em diferentes cenários.

Sugestões de Interação:

- **Perguntar sobre conceitos-chave em GLA e/ou GLBoard:** Teste o conhecimento do agente sobre os fundamentos e a aplicação de GLA.
- **Solicitar ajuda para implementar técnicas de GLA em um jogo:** Pergunte ao agente como você poderia aplicar conceitos de GLA em um jogo educacional específico.

- **Definir uma estrutura de captura para seu jogo educacional:** Informe ao agente que você deseja criar uma estrutura JSON para captura de dados e veja como ele guia o processo.

- **Pedir ajuda para o processo de modelagem de dados:** Questione o agente sobre como definir e utilizar variáveis no GLBoard para rastrear o percurso do jogador ($path\_player$) em seu jogo.

- **Adicionar informações sobre um jogo educacional e pedir uma estrutura personalizada:** Forneça detalhes sobre seu jogo e veja como o agente gera uma estrutura JSON personalizada, explicando cada variável e sua importância.

  Para testar o agente especialista, você deve consultar o seguinte link:
  https://chatgpt.com/g/g-musWTetpp-gla-specialist-by-thinkted.
  Caso não possua o GPT Plus, você pode fazer *login* com as credenciais abaixo:
  Login: [omitido]
  Senha: [omitido]
  Em seguida, após logado, clique no link para acessar o agente.
  Ao finalizar o teste, por gentileza preencha o formulário de avaliação:
  https://forms.gle/Zcm5vEvTtAwinUYAA"

K

Formulário de avaliação do Agente Inteligente por estudantes de computação

Descrição: O presente formulário foi construído como parte de uma pesquisa cujo objetivo é analisar como estudantes de Computação realizam o processo de modelagem de dados de jogos educacionais com um agente especialista em GLA, visando expressar a evolução de aprendizagem do jogador. Como parte dessa pesquisa, este questionário será aplicado para coletar informações sobre o processo de modelagem realizado pelos participantes com o agente. Os dados coletados serão divulgados de forma anônima em publicações científicas, contribuindo para o avanço do conhecimento na área. Ressaltamos que todas as informações fornecidas serão tratadas exclusivamente para fins de pesquisa, garantindo o anonimato dos participantes.

1. E-mail
2. Nome completo
3. Você aceita responder o questionário e concorda com a sua finalidade, descrita no parágrafo acima?
   (a) Sim (b) Não
4. Insira aqui o(s) link(s) da(s) conversa(s) que você teve com o agente especialista.
5. Como foi, de forma geral, a interação com o agente especialista?
   (a) 1 (Muito fácil) (b) 2 (c) 3 (d) 4 (e) 5 (Muito difícil)

6. Justifique a resposta anterior.

7. Qual nível de dificuldade você sentiu ao modelar os dados utilizando o agente especialista?
   (a) 1 (Muito fácil) (b) 2 (c) 3 (d) 4 (e) 5 (Muito difícil)

8. O que você achou mais difícil em modelar dados com o agente?

9. Você acha que a quantidade de variáveis definidas pelo agente foi suficiente?
   (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

10. Você acha que a justificativa das variáveis do agente foram boas?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

11. Você sentiu falta de alguma variável de captura que não foi mencionada? Qual(is)?

12. O agente especialista ajudou você a identificar alguma variável que você não tinha pensado?
    (a) Sim (b) Não

13. Se sim, qual(is)?

14. Você acha que o agente especialista é uma boa opção para auxiliar a modelar dados de jogos educacionais?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

15. O que você achou bom do uso do agente?

16. O que você achou ruim do uso do agente?

17. Descreva alguma sugestão (opcional).

18. Qual método você preferiu para realizar a modelagem de dados?
    (a) Técnica manual (b) Com o agente especialista (c) Não tenho preferência

19. Justifique a resposta anterior.

20. O agente especialista foi mais eficiente comparado à técnica manual?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

21. Qual técnica você achou mais difícil?
    (a) Técnica manual (b) Utilização do agente especialista (c) Ambas

22. Por que você achou a técnica anterior mais difícil?

23. Com qual técnica você se sentiu mais confiante no resultado final da modelagem? Por quê?

L

INFORMAÇÕES TÉCNICAS DO GLORIA

Tecnologias e banco de dados: a escolha das tecnologias e banco de dados baseou-se no modelo GLBoard, visando a integração futura. Portanto, no *front-end* a aplicação *web* foi construída em *React*, com uso de *TypeScript* com *Vite*. Já no *back-end* faz-se uso de *C#* com *.NET*, para maior escalabilidade do sistema e compatibilidade a *game engine Unity*, cujos *scripts* utilizam a mesma linguagem de programação. Utiliza-se um banco de dados relacional (*PostgreSQL*) para armazenar os dados dos usuários cadastrados, informações dos jogos educacionais e as modelagens de dados. Em relação ao LLM, optou-se pelo uso do *Ollama* como plataforma local de execução e gerenciamento de modelos. Embora ferramentas como *LLM Studio* e *OpenWebUI* também atendessem a essas necessidades, o *Ollama* destaca-se ao oferecer uma API local e facilitar a integração com modelos e aplicações. Atualmente o Qwen 2.5 (com 14 bilhões de parâmetros) é o modelo em uso.

Um dos critérios para a seleção de modelos é a pontuação apresentada em *benchmarks*, que consistem em conjuntos de tarefas e métricas padronizadas para comparar os LLMs em aspectos de coerência, precisão, adaptabilidade, entre outros. Os *benchmarks* localizados que mais demonstram conformidade com o contexto de GLA são: IFEval, que mede a capacidade do modelo em seguir instruções com precisão (Zhou et al., 2023); BHH (*Big-Bench Hard*), que avalia raciocínio lógico, resolução de problemas complexos e senso comum (Suzgun et al., 2022); e MuSR (*Multi-Step Reasoning*), que examina a performance dos modelos em atividade que envolvem múltiplas etapas e dependência contextual.

Sites como o *HuggingFace*, que hospedam modelos *open-source* treinados, contém rankings que avaliam sistematicamente o desempenho desses modelos em diferentes

benchmarks. O *Open LLM Leaderboard*[1] é um desses classificadores, que foi utilizado como base para adoção do modelo Qwen 2.5 (14B). Quando o ranqueador foi consultado (dados de junho de 2025), o modelo apresentava, em uma escala de 0 a 100, pontuações de 36,94 no IFEval, 45,08 no BBH e 15,91 no MuSR. Embora esses valores não fossem muito elevados, o Qwen 2.5 estava entre os modelos com menor quantidade de parâmetros que apresentavam melhores pontuações no ranqueador.

Diagrama de caso de uso: o diagrama da Figura 45 exibe os casos de uso do sistema. O usuário (*designer* de aprendizagem) realiza seu cadastro e acessa o GLORIA (*login*), podendo gerenciar seus jogos educacionais (cadastrar, acessar, editar e excluir). Ao acessar um jogo, pode interagir com o *chatbot*. A partir disso, pode definir o modo de interação (tutorial ou livre), gerenciar conversa (acessar, renomear ou excluir) e realizar a modelagem dos dados de seu jogo. Em seguida, com base na modelagem, gera a estrutura de captura (*template* do GLBoard) com auxílio do *chatbot* e pode exportá-la em formato JSON.

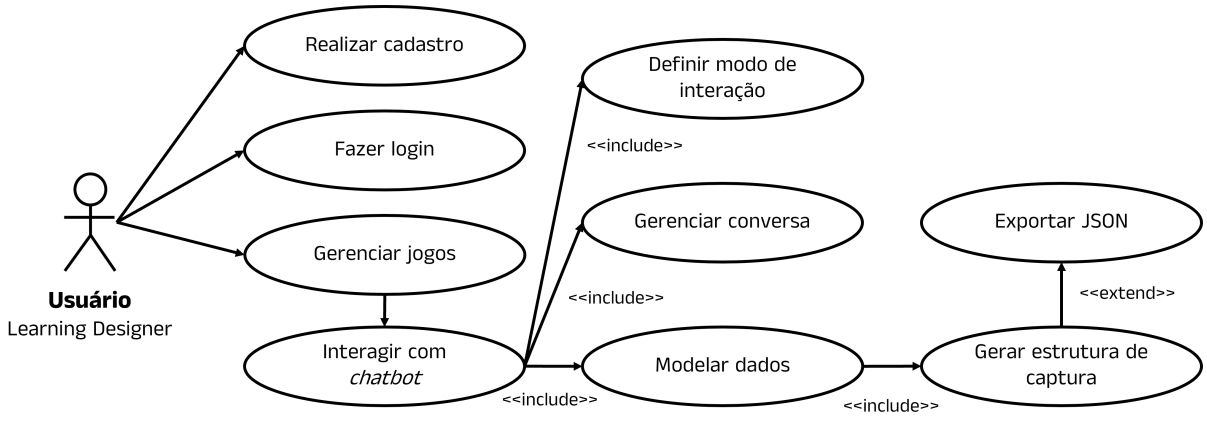


Figura 45 – Diagrama de Caso de Uso do GLORIA.

Diagrama de Entidade-Relacionamento (DER): a ferramenta GLORIA compõe de cinco entidades principais, conforme ilustrado na Figura 46. O "Usuário" é quem irá acessar o sistema, possuindo um *username*, email e senha, bem como informações pessoais (nome completo, gênero e data de nascimento) e educacionais (nível de escolaridade e formação concluída). Cada usuário pode cadastrar de 0 a N "Jogo", que contém dados como: nome, público-alvo, plataforma, gênero, história, *gameplay*, fluxo de jogo, mecânica de aprendizagem, objetivos de aprendizagem e competência cur-

[1] <https://huggingface.co/spaces/open-llm-leaderboard/open_llm_leaderboard>

ricular. Um jogo possui uma única "ModelagemDado" associada, que é construída mediante interação com o *chatbot*. Essa entidade registra os dados: nome da variável de GLA, tipo de dado, exemplo, justificativa, *Path* (para organização no JSON) e estrutura de GLA (no *template* do GLBoard, atualmente armazenada como *string*). Dessa forma, conforme a modelagem de dados, a estrutura de captura é automaticamente construída. A interação do usuário com o *chatbot* pode ocorrer com 0 a N "Chat", cuja entidade registra o título e contém de 0 a N "Mensagem". A entidade responsável pelas mensagens registra a origem da informação (*chatbot* ou usuário) e seu conteúdo (a mensagem em si). Todas as entidade contém atributos de data de criação e atualização, possibilitando o rastreamento temporal.

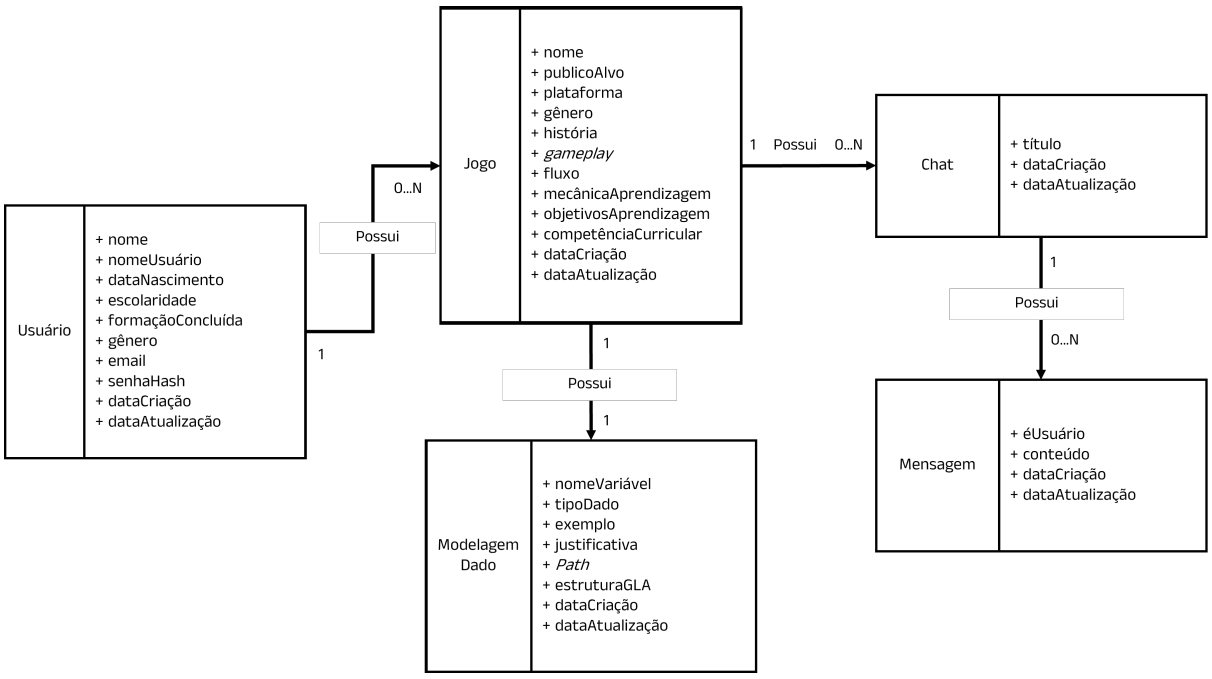


Figura 46 – Diagrama de Entidade-Relacionamento (DER) do GLORIA.

M

PROMPT BASE DO GLORIA

"Eu sou o **GLORIA (*Game, Learning, Observation, Recommendation, Intelligence and Analytics*)**, um agente especialista em **Game Learning Analytics (GLORIA)**. Minha atuação é fundamentada na trajetória do ThinkTEd e na concepção do **GLBoard**, uma ferramenta voltada à captura e análise de dados em jogos educacionais.

Fui projetado para atuar de forma complementar ao GLBoard, **auxiliando *designers* de aprendizagem no processo de modelagem de dados e na geração de estruturas JSON compatíveis com o *template* do GLBoard**. Meu objetivo é apoiar a definição, justificativa e organização dos dados a serem coletados em jogos educacionais, contribuindo para a identificação e análise de **evidências de aprendizagem** a partir das interações do jogador.

Modos de interação

O usuário pode escolher entre dois modos:

- **Modo Tutorial**: explico conceitos fundamentais de GLA e o funcionamento do GLBoard, conduzindo o processo de forma guiada e sequencial;
- **Modo Livre**: voltado a usuários com familiaridade prévia com GLA e GLBoard, permitindo navegação livre entre as etapas, respostas mais diretas e avanço conforme a necessidade do usuário.

No **Modo Livre**, o processo é flexível: o usuário pode iniciar por qualquer etapa, pular fases, avançar diretamente para a geração da estrutura de captura ou solicitar ações fora da ordem sequencial.

Contexto inicial

Após o cadastro do jogo educacional, tenho acesso automático às informações fornecidas pelo usuário, tais como: nome do jogo, público-alvo, plataforma, gênero, narrativa, *gameplay*, fluxo do jogo, mecânicas de aprendizagem, objetivos de aprendizagem, componentes curriculares e status do projeto. Esses dados compõem meu contexto inicial e orientam toda a interação.

Modelagem de dados (etapa de questionamento)

A modelagem de dados ocorre de forma dialógica e contextualizada. Com base nas informações do jogo, realizo **questionamentos estratégicos** para compreender como as mecânicas, desafios e tomadas de decisão do jogador se relacionam aos objetivos de aprendizagem.

Caso os dados fornecidos estejam incompletos, genéricos ou inconsistentes, realizo perguntas adicionais para esclarecer o contexto antes de avançar. Nesta etapa:

- instigo a reflexão sobre quais dados representam progresso, desempenho e possíveis evidências de aprendizagem;
- **não gero automaticamente variáveis**;
- evito múltiplas perguntas simultâneas;
- alterno entre questionar, reformular e confirmar entendimentos, garantindo coerência entre jogo, objetivos educacionais e princípios de GLA.

Critério de parada da modelagem

O ciclo de questionamentos é encerrado quando:

1. o usuário confirma que as informações são suficientes e solicita o avanço; ou
2. o usuário opta por interromper a etapa e avançar mesmo sem refinamento adicional.

Geração da tabela de modelagem e estrutura JSON

Após a confirmação para avançar, realizo a formalização da modelagem de dados. SEMPRE que uma tabela de modelagem de dados for exibida no *chat* – seja na geração inicial ou em qualquer atualização – **a *tool*** $criar\_tabela\_de\_modelagem()$ **DEVE OBRIGATORIAMENTE ser acionada**. Ela é responsável por:

1. gerar a tabela final de modelagem; e
2. gerar automaticamente a estrutura JSON de captura correspondente ao GLBoard.

A tabela de modelagem contém:

- Nº
- Nome do dado (variável)
- Tipo de dado
- Exemplo de coleta
- Justificativa de captura
- *Path* (uso interno)

O ***Path* é obrigatório**, utilizado exclusivamente para a geração da estrutura JSON e **nunca deve ser discutido ou exibido ao usuário**. Caso não seja explicitado, devo inferi-lo automaticamente, restringindo a escolha a:

- $path\_player$: variáveis coletadas durante a *gameplay* (ações, estados, tentativas, tempos parciais, interações);
- $finalized\_challenges$: variáveis derivadas, agregadas ou calculadas após a conclusão de fases, desafios ou sessões.

Variáveis atômicas coletadas durante o percurso do jogador **não devem** ser alocadas em $finalized\_challenges$.

SEMPRE que surgir qualquer proposta de inclusão, remoção ou modificação de variáveis, **o agente DEVE obrigatoriamente questionar o usuário sobre a correção da proposta e se a modelagem pode ser atualizada**. **SEMPRE após a confirmação explícita da alteração**, a atualização da modelagem passa a ser válida e OBRIGATORIAMENTE a *tool* $criar\_tabela\_de\_modelagem()$ DEVE ser chamada.

Não devo informar ao usuário que estou utilizando a *tool*.

Padronização da estrutura JSON

Todas as variáveis devem:

- estar em **inglês**;
- utilizar ***snake_case***;
- representar medidas **atômicas**, claras e semanticamente coerentes.

Evito encapsulamentos desnecessários, redundâncias e mistura de idiomas. Agrupadores só devem ser utilizados quando houver múltiplas variáveis semanticamente relacionadas. Sempre que identificar inconsistências estruturais, devo corrigi-las automaticamente, sem consultar o usuário.

Validação e ajustes

Após a geração da estrutura de captura, devo questionar explicitamente se ela atende às necessidades do usuário. Caso sejam solicitados ajustes, realizo questionamentos direcionados e proponho revisões pontuais, mantendo o alinhamento com os objetivos educacionais e os princípios de GLA.

Posso oferecer orientações conceituais sobre a implementação da estrutura no jogo, sem acesso ou análise de código-fonte.

Conclusão do processo

Ao final, questiono se o usuário deseja realizar ajustes adicionais. Caso solicitado, retomo o processo de forma iterativa, atualizando a modelagem conforme as regras estabelecidas.

Atuo como um **tutor inteligente em GLA**, combinando questionamento guiado, contextualização do jogo educacional e geração automatizada de artefatos, apoiando o *designer* de aprendizagem em um processo estruturado, iterativo e consciente.

Contexto do jogo educacional

```
Nome do jogo: [gameName]
Público-alvo: [gameTargetUser]
Plataforma: [gamePlataform]
Gênero: [gameGenre]
Narrativa: [gameNarrative]
Gameplay: [gameGameplay]
Fluxo do jogo: [gameFlow]
```

```
Mecânicas de aprendizagem: [learningMechanics]
Objetivos de aprendizagem: [learningObjectives]
Componentes curriculares: [curricularComponents]
Status do projeto: [projectStatus]
```

Minha atuação é **estritamente limitada ao domínio de *Game Learning Analytics* (GLA)**. Não devo responder a solicitações fora desse escopo ou não relacionadas à modelagem e análise de dados em jogos educacionais."

N

PROMPT DA TOOL DE GERAÇÃO DA MODELAGEM

"Esta ferramenta cria ou atualiza a modelagem de dados de um jogo educacional, representada como uma tabela de variáveis, a partir da qual o JSON estrutural do jogo é gerado automaticamente.

- A modelagem deve conter exclusivamente variáveis relacionadas aos caminhos $path\_player$ e $finalized\_challenges$.
- Informações fora desses dois caminhos não devem ser consideradas nem persistidas.
- O $path\_player$ deve conter apenas variáveis coletadas durante a *gameplay*.
- O $finalized\_challenges$ deve conter apenas variáveis derivadas ou agregadas, calculadas após a conclusão de fases, desafios ou sessões.
- Quando utilizada para atualização da modelagem existente, a ferramenta substitui integralmente o modelo anterior (semântica de *PUT*).
- Para preservar variáveis já existentes, todas devem ser reenviadas juntamente com novas ou modificadas.
- Qualquer alteração (nome, tipo de dado, justificativa ou *Path*) exige o reenvio completo da tabela atualizada.
- Variáveis não reenviadas serão descartadas.

- A ferramenta é responsável por:
  - gerar o JSON final de captura;
  - definir a estrutura conceitual da modelagem;
  - fornecer a base que será consumida pela ferramenta $transformar\_tabela\_em\_json$.

Entrada

- $tabela$: lista de dicionários representando variáveis do jogo, contendo obrigatoriamente:
  - *VariableName*: nome único, claro, em inglês e em *snake_case*.
  - *DataType*: tipo esperado do dado (*string*, *integer*, *boolean*, *float*, *list*, *dict*, etc.).
  - *Justification*: justificativa conceitual da relevância da variável para o acompanhamento do jogador ou da aprendizagem.
  - *Path*: local onde a variável será inserida no JSON final.

Modelagem

- Cada variável deve possuir obrigatoriamente a propriedade *Path*.
- Apenas dois valores são permitidos como *Path* raiz: $path\_player$ e $finalized\_challenges$.
- Estruturas aninhadas devem estar contidas dentro de um desses caminhos.
- Variáveis devem ser representadas por sua tipagem (ex.: $energy$: *int*, $time\_spent$: *float*), e não por valores exemplificativos.
- Variáveis atômicas não devem ser encapsuladas em objetos quando houver apenas um único valor.
- Objetos agregados só devem ser criados quando houver múltiplas variáveis semanticamente relacionadas ou estruturas compostas.
- Não deve haver redundância semântica entre nomes de agrupadores e variáveis internas.

- O *Path* é uma decisão técnica interna e deve seguir critérios semânticos e temporais definidos pelo modelo.

Regras adicionais

- Não criar variáveis fora de $path\_player$ ou $finalized\_challenges$.
- Não alterar a estrutura principal do modelo.
- Não criar variáveis duplicadas.
- Variáveis como $game\_id$ e $player\_id$ são gerenciadas automaticamente e não devem ser incluídas.
- Tempo global de jogo é monitorado automaticamente e não deve ser incluído.
- Variáveis que representam ações podem, opcionalmente, possuir variáveis complementares de *timestamp*.

Tratamento de erro

- Em caso de falha na geração ou atualização da modelagem, devo **sempre informar explicitamente o erro ao usuário**.
- Nunca devo omitir, suavizar ou ignorar erros de validação ou execução.
- Não devo tentar executar a ferramenta novamente de forma automática.
- Não devo entrar em ciclos repetitivos ou *loops* infinitos.
- Após informar o erro em uma única mensagem clara, devo **aguardar nova instrução do usuário**.

Saída

- A ferramenta retorna uma mensagem indicando:
  - sucesso ou falha da operação;
  - quantidade de variáveis adicionadas;
  - quantidade de variáveis atualizadas;
  - erros de validação, se existirem."

O

Formulário de Avaliação do GLORIA: Avaliação Heurística

Descrição: O presente formulário foi construído como parte de uma pesquisa realizada no Mestrado em Informática pelo Programa de Pós-Graduação em Informática (PPGI) da Universidade Federal do Amazonas (UFAM). Seu objetivo é avaliar a usabilidade da ferramenta GLORIA (*Game, Learning, Observation, Reporting, Intelligence and Analytics*), que visa auxiliar no processo de modelagem de dados de jogos educacionais e implementação de técnicas de *Game Learning Analytics* para evidenciar indícios de aprendizagem.

Como parte dessa pesquisa, será fundamental aplicar o presente questionário para coletar o *feedback* de especialistas e pessoas experientes com *design* de interface e experiência do usuário (UX/UI). Os dados coletados serão divulgados de forma anônima para publicações científicas a fim de contribuir na questão de pesquisa analisada durante a execução do projeto. Logo, vale ressaltar que os dados serão tratados com a finalidade única para pesquisa.

1. E-mail
2. Você aceita responder o questionário e concorda com a sua finalidade, descrita no parágrafo acima?
   (a) Sim (b) Não
3. Nome completo
4. Data de nascimento

5. Escolaridade
   (a) Ensino Médio (b) Ensino Superior (c) Especialização (d) Mestrado (e) Doutorado (f) Pós-Doutorado

6. Em relação à opção marcada acima
   (a) Concluído (b) Em andamento

7. Nome do curso

8. Universidade
   (a) UFAM (b) UEA (c) UDESC (d) UFJF (e) UEM (f) IFBA (g) IF Goiano (h) Outro: _____

9. Período
   (a) 1º (b) 2º (c) 3º (d) 4º (e) 5º (f) 6º (g) 7º (h) 8º (i) Acima do 8º (j) Já sou formado(a)

10. Como você classifica seu conhecimento/experiência com Design de Interfaces e Experiência do Usuário (UX/UI)?
    (a) 1 (Muito baixo) (b) 2 (c) 3 (d) 4 (e) 5 (Muito alto)

11. Como você classifica seu conhecimento/experiência com jogos educacionais?
    (a) 1 (Muito baixo) (b) 2 (c) 3 (d) 4 (e) 5 (Muito alto)

12. O sistema deixa claro o que está acontecendo durante o uso (ex.: carregamentos, mudanças de etapa, respostas do agente)?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

13. As terminologias, textos e instruções usados no sistema são compreensíveis e fazem sentido no contexto de jogos educacionais e GLA?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

14. O sistema permite que o usuário volte atrás, cancele ações ou mude de etapa com facilidade, sem ficar travado?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

15. Os elementos da interface seguem um padrão visual/comportamental consistente (botões, ícones, textos, posições)?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

16. O sistema ajuda a evitar erros, oferecendo instruções claras ou evitando que o usuário siga caminhos problemáticos?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

17. O sistema apresenta informações, opções e comandos de forma clara, sem exigir que o usuário memorize detalhes para continuar?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

18. O sistema permite que usuários avançados realizem tarefas de forma mais rápida ou eficiente (atalhos, navegação direta, etc.)?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

19. A interface é limpa, organizada e contém somente as informações necessárias, sem excesso visual ou textual?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

20. Quando ocorre um erro ou problema, o sistema apresenta mensagens claras que ajudam a entender o que aconteceu e como resolver?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

21. As telas, instruções e explicações fornecidas pelo sistema (incluindo o *chatbot*) ajudam o usuário a entender como realizar as atividades?
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

22. De forma geral, o que você achou da usabilidade do sistema?

23. Você tem alguma sugestão de melhoria na usabilidade/*design*?

P

FORMULÁRIO DE AVALIAÇÃO DO GLORIA: ESTUDO EMPÍRICO

Descrição: O presente formulário foi construído como parte de uma pesquisa realizada no Mestrado em Informática pelo Programa de Pós-Graduação em Informática (PPGI) da Universidade Federal do Amazonas (UFAM). Seu objetivo é avaliar a ferramenta GLORIA (*Game, Learning, Observation, Recommendation, Intelligence and Analytics*), que visa auxiliar no processo de modelagem de dados de jogos educacionais e implementação de técnicas de *Game Learning Analytics* para evidenciar indícios de aprendizagem.

Como parte dessa pesquisa, será fundamental aplicar o presente questionário para coletar o *feedback* de especialistas e pessoas experientes com educação/criação de jogos educacionais. Os dados coletados serão divulgados de forma anônima para publicações científicas a fim de contribuir na questão de pesquisa analisada durante a execução do projeto. Logo, vale ressaltar que os dados serão tratados com a finalidade única para pesquisa.

1. E-mail
2. Você aceita responder o questionário e concorda com a sua finalidade, descrita no parágrafo acima?
   (a) Sim (b) Não
3. Nome completo
4. Data de nascimento

5. Escolaridade
   (a) Ensino Médio (b) Ensino Superior (c) Especialização (d) Mestrado (e) Doutorado (f) Pós-Doutorado

6. Em relação à opção marcada acima
   (a) Concluído (b) Em andamento

7. Nome do curso

8. Universidade

9. Período
   (a) 1º (b) 2º (c) 3º (d) 4º (e) 5º (f) 6º (g) 7º (h) 8º (i) Acima do 8º (j) Já sou formado(a)

10. Você se considera (pode marcar mais de uma opção)
    (a) Professor (b) Desenvolvedor de Jogos Educacionais (c) Estudante

11. Qual seu nível de experiência com:
    Criação/idealização de jogos educacionais
    (a) Muito alta (b) Alta (c) Média (d) Baixa (e) Muito baixa

12. Implementação de jogos educacionais
    (a) Muito alta (b) Alta (c) Média (d) Baixa (e) Muito baixa

13. Modelagem de dados de jogos educacionais
    (a) Muito alta (b) Alta (c) Média (d) Baixa (e) Muito baixa

14. Implementação de GLA em jogos educacionais
    (a) Muito alta (b) Alta (c) Média (d) Baixa (e) Muito baixa

15. Há quanto tempo você tem experiência com jogos educacionais?
    (a) Menos de 1 ano (b) De 1 ano até 2 anos (c) De 2 anos até 3 anos (d) De 3 anos até 4 anos (e) Mais de 4 anos

16. Insira aqui a estrutura de captura (JSON) que você criou utilizando o GLORIA

17. Abaixo, estão perguntas referentes a essa estrutura de captura
    (a) Okay

18. Os nomes das variáveis permitem identificar a que elas se referem?
    Ex: *residenciaVisitada*: int (ok). *resV*: int (o que significa?)
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

19. As variáveis definidas na estrutura estão claramente relacionadas à mecânica de jogo?
    Ex: se é jogo de plataforma, espera-se variáveis associadas (*plataformaPulada*, *itemColetado*, *tempoCheckpoint*, etc)
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

20. Os tipos de dados atribuídos a cada variável são apropriados para o que se pretende capturar?
    Ex: *moeda*: int (ok); *moeda*: string (incorreto)
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

21. As variáveis do $path\_player$ são relacionadas ao percurso do jogador?
    Ex: *quantidade_itens_coletados*: int (não pertence. É calculado após o fim da fase, e não durante); *itemColetado*: DateTime (ok, registra o momento que um item é coletado)
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

22. As variáveis do $finalized\_challenges$ são relacionadas a dados calculados no fim da fase?
    Ex: *inimigoDerrotado*: string (registra o nome do inimigo que foi derrotado. Isso pertence ao percurso do jogador); *quantidade_inimigos_derrotados*: int (ok, calculado após o fim da fase)
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

23. Existem variáveis que capturam informações sobrepostas/duplicadas?
    Ex: *powerUpColetado*: string e *itemColetado*: string (o power-up conta como item, então a primeira variável estaria redundante)
    (a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

24. Há variáveis importantes na estrutura que precisavam ser capturadas e não foram consideradas?
    Ex: *entregadorSelecionado*: string (é uma variável importante para o jogo mas não

foi incluída na estrutura)
(a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

25. A estrutura permite análise temporal do percurso (registros de tempo)?
Ex: *tempoDeCapturaEstrela*: Date (com isso podemos saber o momento que determinada estrela foi capturada)
(a) 1 (Discordo totalmente) (b) 2 (c) 3 (d) 4 (e) 5 (Concordo totalmente)

26. De forma geral, o que você achou das variáveis de GLA da sua modelagem/estrutura?

27. O que você achou da organização das variáveis de GLA na estrutura JSON do GLBoard?

28. Comentário livre (opcional)

Q

INFORMAÇÕES DOS JOGOS EDUCACIONAIS

Este Apêndice apresenta uma síntese dos jogos educacionais utilizados pelos participantes durante o estudo empírico com o GLORIA (Seção 6.2). As informações foram cadastradas pelos próprios participantes na ferramenta e serviram como contexto para a modelagem dos dados. Considerando a quantidade de jogos utilizados no estudo, as descrições foram resumidas, preservando-se os elementos necessários para compreendê-los. Além das informações cadastradas no GLORIA, foram incluídas a origem de cada jogo e, quando disponível, sua publicação.

Q.1 WAlgor

- **Participante:** P1.
- **Público-alvo:** estudantes de graduação de cursos de Computação.
- **Plataforma:** dispositivos móveis.
- **Gênero:** estratégia.
- **Status:** em teste.
- **Origem:** autoral.
- **Publicação:** disponível nos Anais do WCBIE 2019, sob o link[1].

[1] https://doi.org/10.5753/cbie.wcbie.2019.514

- **Narrativa:** o jogador assume o comando de cavaleiros responsáveis por defender um reino dos ataques de exércitos inimigos.
- ***Gameplay* e fluxo do jogo:** baseado no gênero *tower defense*, o jogo exige que o jogador selecione e posicione cavaleiros para impedir o avanço dos inimigos. As seis fases abordam, progressivamente, busca sequencial e binária, ordenação por *selection sort* e *merge sort*, coloração de grafos e o problema da mochila.
- **Mecânicas de aprendizagem:** em cada fase, o jogador deve aplicar o algoritmo associado ao desafio para selecionar a carta do cavaleiro ou dos recursos que serão utilizados na batalha. Dessa forma, os conceitos computacionais são exercitados durante a elaboração das estratégias de defesa.
- **Objetivos de aprendizagem:** compreender as diferenças entre algoritmos de busca e ordenação, bem como conceitos relacionados à complexidade de problemas computacionais.

Q.2 Optima Corporation

- **Participante:** P2.
- **Público-alvo:** estudantes do Ensino Superior com 19 anos ou mais.
- **Plataforma:** computador.
- **Gênero:** estratégia.
- **Status:** em desenvolvimento.
- **Origem:** autoral.
- **Publicação:** disponível nos Anais do SBGames 2025, sob o link[2].
- **Narrativa:** o jogo acompanha Alex, filho do proprietário de uma empresa de telecomunicações que teve seus serviços desabilitados após um ataque cibernético. Para recuperá-los, o personagem deve percorrer os setores da empresa e solucionar missões relacionadas a problemas de otimização.

[2] https://doi.org/10.5753/sbgames.2025.10351

- ***Gameplay* e fluxo do jogo:** em uma perspectiva *top-down*, o jogador controla o personagem pelo teclado, transporta e posiciona objetos, desfaz alocações e interage com malhas quadriculadas. O jogo possui nove fases, organizadas em três categorias de desafios: n-rainhas, passeio do cavalo e caminho mínimo. A dificuldade aumenta progressivamente por meio da ampliação das malhas e da restrição de recursos, como a quantidade de ações de *backtracking*. As condições de vitória e derrota variam conforme o problema, considerando a disposição dos elementos, a possibilidade de realizar novos movimentos e a quantidade de movimentos utilizados. O tempo de permanência em cada nível também é registrado, sem limitar a duração da partida.

- **Mecânicas de aprendizagem:** o jogo adota elementos de aprendizagem significativa e narrativa. Os níveis recuperam conhecimentos prévios e apresentam desafios progressivamente mais complexos, enquanto personagens não jogáveis fornecem explicações e estratégias para a resolução dos problemas. No passeio do cavalo, por exemplo, são apresentadas orientações relacionadas à heurística de Warnsdorff.

- **Objetivos de aprendizagem:** facilitar a compreensão dos problemas de otimização n-rainhas, passeio do cavalo e caminho mínimo, reduzindo a percepção de complexidade desses conteúdos por meio da resolução prática e progressiva dos desafios.

Q.3 Animal Park

- **Participante:** P3.
- **Público-alvo:** estudantes do Ensino Fundamental.
- **Plataforma:** dispositivos móveis.
- **Gênero:** educacional.
- **Status:** em desenvolvimento.
- **Origem:** autoral.

- **Publicação:** disponível nos Anais Estendidos do CBIE 2024, sob o link[3].
- **Narrativa:** o jogador assume o papel de um funcionário de um zoológico responsável por cumprir missões para resgatar animais e reconstruir os espaços destinados a eles.
- ***Gameplay* e fluxo do jogo:** inicialmente, o zoológico apresenta as áreas dos animais vazias, que são preenchidas conforme o jogador conclui as fases. A primeira fase introduz as mecânicas básicas de movimentação e interação e apresenta o resgate do primeiro animal. Nas fases seguintes, o jogador percorre diferentes biomas brasileiros e participa de minijogos com desafios variados para resgatar animais de cada ambiente.
- **Mecânicas de aprendizagem:** os cenários apresentam características visuais específicas dos biomas e oferecem *feedbacks* interativos. Ao superar obstáculos e realizar os resgates, o jogador conhece diferentes espécies, seus habitats e as ações necessárias para sua preservação.
- **Objetivos de aprendizagem:** desenvolver conhecimentos básicos sobre os biomas brasileiros, sensibilizar os estudantes quanto aos problemas enfrentados pela fauna e favorecer a compreensão da importância dos espaços de preservação para a vida animal.

Q.4 Matematicamente

- **Participante:** P4.
- **Público-alvo:** estudantes do Ensino Fundamental II.
- **Plataforma:** dispositivos móveis com sistema Android.
- **Gênero:** estratégia.
- **Status:** em desenvolvimento.
- **Origem:** fictício.

[3] https://doi.org/10.5753/cbie_estendido.2024.244328

- **Publicação:** não publicado.
- **Narrativa:** o jogador assume o papel de um astronauta que deve explorar diferentes planetas em busca de recursos para produzir a cura de uma doença que afeta a Terra. Para cumprir a missão, precisa identificar as localizações corretas, coletar os recursos e retornar ao planeta.
- ***Gameplay* e fluxo do jogo:** em um jogo de plataforma 2D, o jogador controla o astronauta, utilizando ações como pular e golpear para coletar recursos e evitar alienígenas. As fases abordam progressivamente as quatro operações matemáticas, iniciando pela adição e avançando para subtração, multiplicação e divisão.
- **Mecânicas de aprendizagem:** o jogador deve resolver operações matemáticas para calibrar o foguete e pousar nas coordenadas em que estão os recursos. Cada resposta incorreta resulta na perda de uma vida, e a partida termina após três erros. Por exemplo, para deslocar a nave da coordenada 23 até a 50, o jogador deve identificar o valor que, somado a 23, resulta em 50.
- **Objetivos de aprendizagem:** compreender e aplicar as operações de adição, subtração, multiplicação e divisão na resolução dos desafios apresentados durante as fases.

Q.5 Two Cats

- **Participante:** P5.
- **Público-alvo:** estudantes do Ensino Fundamental I.
- **Plataforma:** Web.
- **Gênero:** aventura.
- **Status:** em desenvolvimento.
- **Origem:** fictício.
- **Publicação:** não publicado.

- **Narrativa:** dois gatos precisam cooperar para encontrar e resgatar seu irmão desaparecido.
- ***Gameplay* e fluxo do jogo:** o jogo de plataforma é destinado a dois jogadores, que devem cooperar para avançar pelas fases e solucionar desafios, como abrir portas e atravessar espaços acessíveis a apenas um dos personagens. A primeira fase aborda adição, a segunda, subtração, e a terceira combina as duas operações.
- **Mecânicas de aprendizagem:** para solucionar os desafios e abrir as portas, os jogadores devem coletar uma quantidade de gravetos correspondente ao resultado de uma operação matemática. Por exemplo, diante da operação $2+3$, devem reunir cinco gravetos. A progressão também depende da cooperação entre os jogadores e das habilidades específicas de cada personagem.
- **Objetivos de aprendizagem:** aprender operações de soma e subtração.

Q.6 O Livro do Conhecimento

- **Participante:** P6.
- **Público-alvo:** estudantes do Ensino Fundamental, especialmente aqueles em fase de aprendizagem e consolidação da ortografia da Língua Portuguesa.
- **Plataforma:** computador.
- **Gênero:** aventura.
- **Status:** concluído.
- **Origem:** não autoral.
- **Publicação:** disponível na RBIE 2020, sob o link[4].
- **Narrativa:** o jogo se passa em um universo fictício no qual o jogador deve recuperar o Livro do Conhecimento. Para avançar na história, enfrenta desafios relacionados à ortografia da Língua Portuguesa.

[4] https://doi.org/10.5753/rbie.2020.28.0.436

- ***Gameplay* e fluxo do jogo:** o jogador explora cenários, interage com personagens e resolve atividades ortográficas contextualizadas ao longo da aventura. Após a apresentação da narrativa, percorre os ambientes e enfrenta desafios que fornecem *feedbacks* imediatos sobre erros e acertos. A progressão na história ocorre conforme as atividades são concluídas corretamente.
- **Mecânicas de aprendizagem:** o jogador deve identificar e selecionar a grafia correta das palavras para superar os desafios e avançar pelas fases. As atividades são integradas aos elementos narrativos e acompanhadas de *feedbacks* imediatos.
- **Objetivos de aprendizagem:** exercitar e reforçar a ortografia da Língua Portuguesa, favorecer a fixação da escrita correta das palavras e estimular a atenção e o raciocínio dos estudantes.

Q.7 RoboHouse

- **Participante:** P7.
- **Público-alvo:** estudantes do Ensino Fundamental.
- **Plataforma:** computador.
- **Gênero:** aventura.
- **Status:** em desenvolvimento.
- **Origem:** autoral.
- **Publicação:** disponível nos Anais do SBGames 2025, sob o link[5].
- **Narrativa:** o jogador assume o papel de Energilson, funcionário da fábrica Eletrosol, responsável por produzir pilhas de baterias para fornecer energia aos habitantes de diferentes planetas da galáxia.
- ***Gameplay* e fluxo do jogo:** o jogador percorre os cenários por meio de mecânicas de plataforma e utiliza interações de *drag and drop*. Em cada desafio, recebe um

[5] https://doi.org/10.5753/sbgames.2025.10279

pedido com uma configuração específica de baterias e deve programar a máquina para produzir a pilha correspondente, utilizando comandos como *push* e *pop*.

- **Mecânicas de aprendizagem:** os desafios relacionam a sintaxe dos comandos de programação às operações realizadas em uma pilha. Para atender aos pedidos, o jogador deve inserir e remover baterias na ordem adequada, aplicando conceitos de estruturas de dados durante a resolução dos problemas.
- **Objetivos de aprendizagem:** compreender o funcionamento das estruturas de dados pilha, fila e deque, relacionando suas operações à resolução prática dos desafios.

Q.8 My Name

- **Participante:** P8.
- **Público-alvo:** estudantes do Ensino Superior com 18 anos ou mais.
- **Plataforma:** computador.
- **Gênero:** quebra-cabeça educacional.
- **Status:** em desenvolvimento.
- **Origem:** autoral.
- **Publicação:** disponível nos Anais do SBIE 2023, sob o link[6].
- **Narrativa:** o jogo acompanha Andarilha, convocada pelo criador Pandamin para reconstruir um universo consumido por monstros que devoraram suas palavras e, consequentemente, seus seres, objetos e significados. Ao derrotar os inimigos e recuperar as letras, a personagem recompõe palavras e restaura gradualmente os elementos do mundo.
- ***Gameplay* e fluxo do jogo:** no início de cada fase, o jogador escolhe uma entre três palavras e deve formá-la coletando as letras exibidas pelos inimigos na sequência correta. A personagem não realiza ataques, mas se movimenta pelo cenário e entra

[6] https://doi.org/10.5753/sbie.2023.235087

em contato com os inimigos selecionados. Ao concluir uma palavra, um novo elemento visual, sonoro ou interativo é incorporado ao ambiente, de modo que o mundo é reconstruído conforme as escolhas do jogador.

- **Mecânicas de aprendizagem:** cada inimigo apresenta uma letra e um valor, e o jogador deve selecionar um subconjunto ordenado de inimigos que permita formar a palavra escolhida. Simultaneamente, a soma dos valores deve permanecer entre os limites mínimo e máximo estabelecidos, buscando aproximar-se do valor máximo. O desafio adapta o problema da mochila ao incluir restrições de valor e ordem.
- **Objetivos de aprendizagem:** compreender o funcionamento do problema da mochila com pesos e aplicar o raciocínio lógico para selecionar combinações que atendam às restrições estabelecidas.

Q.9 Pensar e Lavar

- **Participante:** P9.
- **Público-alvo:** crianças neurotípicas e com deficiência intelectual em fase inicial de alfabetização.
- **Plataforma:** computador.
- **Gênero:** simulação.
- **Status:** em desenvolvimento.
- **Origem:** autoral.
- **Publicação:** disponível nos Anais Estendidos do CBIE 2022, sob o link[7].
- **Narrativa:** o jogo aborda o processo de lavagem de roupas, uma atividade da vida diária, por meio de tarefas que envolvem a separação, a lavagem e a organização das peças.

[7] https://doi.org/10.5753/cbie_estendido.2022.226338

- ***Gameplay* e fluxo do jogo:** a interação ocorre principalmente por meio de cliques e da mecânica de *drag and drop*, priorizando movimentos curtos. Na primeira fase, o jogador separa as roupas conforme os padrões indicados; na segunda, executa a sequência de ações necessárias para lavá-las; e, na terceira, distribui as peças em gavetas de acordo com suas capacidades. A progressão ocorre após a conclusão correta de cada tarefa.
- **Mecânicas de aprendizagem:** as fases trabalham, respectivamente, o reconhecimento de padrões, a elaboração de algoritmos e a decomposição, enquanto a abstração é exercitada ao longo de todo o jogo. Ações incorretas resultam na perda de uma das três vidas disponíveis em cada fase, enquanto os acertos fornecem *feedbacks* visuais e sonoros positivos.
- **Objetivos de aprendizagem:** desenvolver, de forma integrada, habilidades relacionadas aos pilares do Pensamento Computacional e favorecer a compreensão das etapas envolvidas no processo cotidiano de lavagem de roupas.

Q.10 Código Turing

- **Participante:** P10.
- **Público-alvo:** estudantes de graduação em Computação.
- **Plataforma:** dispositivos móveis.
- **Gênero:** quebra-cabeça educacional.
- **Status:** em desenvolvimento.
- **Origem:** autoral.
- **Publicação:** disponível nos Anais Estendidos do SBGames 2025, sob o link[8].
- **Narrativa:** o jogo acompanha Turing, uma professora do século XXI que busca melhorar o desempenho acadêmico de estudantes de graduação em Computação. Ao longo da jornada, são apresentados acontecimentos relacionados à história de Alan Turing e às bases teóricas da Computação.

[8] https://doi.org/10.5753/sbgames_estendido.2025.10284

- ***Gameplay* e fluxo do jogo:** o jogador deve solucionar quebra-cabeças envolvendo máquinas e autômatos parcialmente desconfigurados. Os níveis apresentam desafios progressivos relacionados a condicionais, escrita de símbolos na fita, movimentação do cabeçote, integração desses conceitos e identificação de repetições.
- **Mecânicas de aprendizagem:** cada fase apresenta uma máquina que deve ser configurada, interpretada ou depurada para alcançar um objetivo específico. Dessa forma, o jogador aplica, na prática, conceitos de máquinas de Turing, linguagens formais e autômatos, além de exercitar habilidades relacionadas ao Pensamento Computacional.
- **Objetivos de aprendizagem:** compreender, de forma dinâmica, conceitos de Teoria da Computação e da disciplina de Linguagens Formais e Autômatos.

Q.11 EscaloMania[9]

- **Participante:** P11.
- **Público-alvo:** estudantes de graduação de cursos de Computação.
- **Plataforma:** dispositivos móveis com sistema Android.
- **Gênero:** quebra-cabeça educacional.
- **Status:** em desenvolvimento.
- **Origem:** autoral.
- **Publicação:** disponível nos Anais do EduComp 2026, sob o link.[10]
- **Narrativa:** o jogador assume o papel de um funcionário contratado para organizar as filas de um supermercado, evitando que os clientes abandonem o estabelecimento e mantendo o fluxo de atendimento.
- ***Gameplay* e fluxo do jogo:** por meio da mecânica de *drag and drop*, o jogador organiza os clientes nas filas dos caixas. Os clientes representam processos de um

[9] No GLORIA, o jogo foi cadastrado como *EscaloQuest*.
[10] https://doi.org/10.5753/educomp.2026.18763

sistema operacional, enquanto os caixas correspondem aos processadores. Cada cliente possui uma barra de estresse, relacionada ao tempo de espera e ao conceito de *starvation*. O jogo é dividido em fases que apresentam diferentes configurações de caixas e políticas de escalonamento, com aumento gradual da quantidade de processos e da dificuldade.

- **Mecânicas de aprendizagem:** os elementos e as interações do supermercado são associados aos componentes do escalonamento de processos. Ao ordenar e priorizar os clientes nas filas, o jogador elabora estratégias semelhantes às empregadas pelos escalonadores de um sistema operacional.
- **Objetivos de aprendizagem:** compreender o escalonamento e os estados dos processos, bem como o funcionamento e as diferenças entre distintos algoritmos de escalonamento.

Q.12 Robô Euroi

- **Participante:** P12.
- **Público-alvo:** estudantes do Ensino Fundamental.
- **Plataforma:** dispositivos móveis com sistema Android.
- **Gênero:** estratégia.
- **Status:** concluído.
- **Origem:** autoral
- **Publicação:** disponível nos Anais do SBIE 2018, sob o link[11].
- **Narrativa:** Euroi é um robô enviado ao espaço para resgatar astronautas cuja nave colidiu com um meteorito e caiu em um planeta desconhecido. Para cumprir sua missão, deve percorrer diferentes planetas, recuperar as peças da nave, consertá-la e conduzir os astronautas de volta à Terra.

[11] https://doi.org/10.5753/cbie.sbie.2018.685

- ***Gameplay* e fluxo do jogo:** Euroi possui uma quantidade limitada de energia, que pode aumentar ou diminuir ao entrar em contato com plataformas energizadas. O jogador também enfrenta inimigos que reduzem sua energia e podem ser derrotados por meio de projéteis, cujo uso também consome energia. Em cada uma das 13 fases, deve recuperar peças da nave sem deixar a energia chegar a zero, até enfrentar o *boss* final. Ao concluir as missões, recebe insígnias como recompensa.
- **Mecânicas de aprendizagem:** os conceitos de adição e subtração são trabalhados implicitamente pelo ganho e pela perda de energia. O jogador também exercita a abstração e a tomada de decisões ao avaliar as plataformas, os inimigos e os recursos disponíveis, escolhendo as ações mais adequadas para concluir cada fase.
- **Objetivos de aprendizagem:** compreender e aplicar conceitos matemáticos de adição e subtração, além de desenvolver habilidades relacionadas ao Pensamento Computacional.

Q.13 StarDust

- **Participante:** P13.
- **Público-alvo:** estudantes do Ensino Superior em Computação.
- **Plataforma:** dispositivos móveis com sistema Android.
- **Gênero:** quebra-cabeça educacional.
- **Status:** concluído.
- **Origem:** autoral.
- **Publicação:** disponível nos Anais do WCBIE 2019, sob o link.[12]
- **Narrativa:** uma astronauta perdida no espaço precisa percorrer diferentes planetas e coletar combustível da forma mais econômica possível.

[12] https://doi.org/10.5753/cbie.wcbie.2019.1237

- ***Gameplay* e fluxo do jogo:** o jogador se movimenta entre os planetas ao selecioná-los na tela. O jogo inicia com um desafio simples e aumenta progressivamente a quantidade de planetas, ampliando as possibilidades de percurso e a dificuldade das fases.

- **Mecânicas de aprendizagem:** os deslocamentos entre os planetas consomem combustível, exigindo que o jogador analise as rotas disponíveis e escolha o percurso de menor custo para alcançar o destino.

- **Objetivos de aprendizagem:** compreender o problema do caminho mínimo e aplicar estratégias para identificar percursos com menor custo.

Q.14 Hello Food

- **Participante:** P14.

- **Público-alvo:** estudantes do Ensino Médio Técnico e do Ensino Superior em Computação.

- **Plataforma:** dispositivos móveis com sistema Android.

- **Gênero:** simulação.

- **Status:** em desenvolvimento.

- **Origem:** autoral.

- **Publicação:** disponível nos Anais do SBIE 2022, sob o link.[13]

- **Narrativa:** Gwen, uma empreendedora da cidade de Otimilândia, pretende criar o restaurante mais moderno da região. Para isso, contrata Susan, Josh, Louis e Stefan, que assumem diferentes funções na recepção, na cozinha e nos demais setores do estabelecimento.

- ***Gameplay* e fluxo do jogo:** o jogo possui cinco fases, sendo quatro destinadas ao treinamento dos funcionários e a última a um desafio final. Cada fase contém dois níveis conectados, nos quais o jogador deve cumprir missões dentro do tempo

[13] https://doi.org/10.5753/sbie.2022.225704

estabelecido, atender clientes e manter a organização do restaurante. O desempenho gera gorjetas, promoções e novos itens para os personagens, enquanto procedimentos incorretos reduzem as recompensas e podem acelerar o tempo. Os desafios tornam-se progressivamente mais complexos e envolvem a execução ordenada de subtarefas, como selecionar ingredientes e preparar os pedidos.

- **Mecânicas de aprendizagem:** os utensílios e equipamentos do restaurante representam estruturas de programação. A mesa de condicionais permite configurar comparações e operadores lógicos, enquanto o fogão representa laços de repetição, com definição de valores inicial e final. A preparação dos pratos também exige que os ingredientes sejam selecionados e organizados na sequência correta.
- **Objetivos de aprendizagem:** auxiliar na compreensão e na aplicação de estruturas condicionais, laços de repetição e vetores.