Sistema Tutor Inteligente Baseado em Agentes de IA para a Recuperação de Conhecimento Científico e Prático no Programa Meninas Digitais

KEVIN REHBEIN[1]; MARILTON SANCHOTENE DE AGUIAR[2]

[1]*Universidade Federal de Pelotas – krehbein@inf.ufpel.edu.br*
[2]*Universidade Federal de Pelotas – marilton@ufpel.edu.br*

1. INTRODUÇÃO

O Programa Meninas Digitais consiste em uma iniciativa da Sociedade Brasileira de Computação (SBC) que visa incentivar o ingresso de mulheres em carreiras de Tecnologia da Informação e Comunicação (TIC), mediante a congregação de projetos parceiros em todo o Brasil. Cada projeto produz conhecimentos relevantes, a exemplo de materiais de ações e de iniciativas, os quais se encontram parcialmente documentados nos anais do *Women in Information Technology* (WIT), *workshop* do Congresso da SBC. No entanto, não obstante esse expressivo volume de produção, identifica-se uma problemática: tais projetos operam de forma isolada e confrontam dificuldades em razão da distância geográfica e da ausência de mecanismos eficientes para o compartilhamento e para a reutilização do conhecimento já produzido. Essa situação acarreta retrabalho e a perda de práticas bem-sucedidas que não chegam a ser publicadas formalmente. Para mitigar essa problemática, o presente trabalho propõe o desenvolvimento de um Sistema Tutor Inteligente (ITS), fundamentado em uma arquitetura de Sistemas Multiagentes (MAS) baseada em Grandes Modelos de Linguagem (LLM). A ferramenta atuará como um copiloto, direcionada à geração de planos de projetos e de atividades personalizados para cada localidade. A solução emprega os LLMs como núcleo de processamento para agentes especializados, integrando a técnica de Geração Aumentada por Recuperação (RAG) para assegurar a precisão e a coerência nas respostas. As contribuições esperadas compreendem uma ferramenta funcional para a rede Meninas Digitais, capaz de democratizar o acesso a materiais de qualidade, de reduzir o retrabalho e de acelerar a curva de aprendizado de novos projetos em todo o território nacional. Além disso, a pesquisa objetiva demonstrar a viabilidade da construção de um sistema agêntico de Inteligência Artificial (IA) compatível com as restrições orçamentárias universitárias.

A Inteligência Artificial (IA) constitui o campo do conhecimento dedicado a projetar e a desenvolver sistemas capazes de simular a inteligência e as habilidades humanas. Por sua vez, o conceito de agente, originário da filosofia e incorporado à computação, define entidades artificiais que percebem o ambiente, tomam decisões e agem de forma autônoma para alcançar objetivos específicos. Com o avanço das pesquisas em IA, o conceito de agente consolidou-se na área para descrever entidades dotadas de comportamento inteligente, caracterizadas por atributos como autonomia, reatividade, proatividade e capacidade de interação social. Em virtude disso, a investigação e o desenvolvimento técnico de agentes passaram a ocupar uma posição de destaque nas discussões da comunidade científica (Xi et al., 2025).

Com o desenvolvimento dos LLMs, estabeleceu-se um novo paradigma na implementação de agentes de IA. Esses modelos demonstraram capacidades notáveis que os qualificam para atuar como o "cérebro" de agentes

especializados. Entre essas habilidades, destacam-se a interação em linguagem natural com os seres humanos, a aquisição de conhecimento, a execução de instruções, o raciocínio, o planejamento e a utilização de ferramentas (Xi et al., 2025). As competências excepcionais alcançadas pelos LLMs nesses domínios alinharam-se substancialmente às expectativas humanas a respeito de agentes aptos a perceber o ambiente e a tomar decisões (Li; Wang; Zeng; Wu; Yang, 2024).

Adicionalmente, a proposição de sistemas multiagentes possibilitou uma nova forma de aproveitar a inteligência coletiva inter-agêntica, sem descartar as especialidades individuais de cada agente. Assim, emergiu uma nova e mais sofisticada arquitetura de inteligência autônoma, pautada na colaboração e na comunicação entre agentes com diferentes identidades para alcançar objetivos e executar tarefas mais complexas (Li; Wang; Zeng; Wu; Yang, 2024).

Partindo desse pressuposto, a implementação de um MAS com o objetivo de atuar como um tutor inteligente apresenta-se como uma alternativa tecnológica moderna e adequada para solucionar a lacuna de desagregação do conhecimento e o retrabalho pedagógico observados na rede de projetos associados ao Meninas Digitais.

No entanto, os LLMs são modelos suscetíveis a fenômenos de geração de texto que divergem da realidade, denominados alucinações. Para contornar essas limitações, uma abordagem eficaz consiste na utilização do RAG, que combina mecanismos de recuperação para produzir respostas coerentes e ricas em contexto. Nessa técnica, o sistema extrai fragmentos relevantes para a consulta a partir de fontes externas fornecidas (Li; Wang; Zeng; Wu; Yang, 2024). Embora o RAG solucione parcialmente esses problemas, o seu funcionamento baseia-se na similaridade semântica pura, o que pode torná-lo menos eficaz em contextos educacionais, nos quais as relações conceituais são indispensáveis (Dong; Yuan; Chen; Cheng; Wen, 2025). Por essa razão, a Geração Aumentada por Recuperação baseada em Grafos de Conhecimento (KG-RAG) desponta como uma solução viável ao organizar o conhecimento em nós e arestas (entidades e relações), o que garante a manutenção das conexões entre os conceitos pedagógicos (Zhang et al., 2025).

2. METODOLOGIA

O desenvolvimento do trabalho está organizado em etapas relacionadas à investigação, implementação e avaliação do sistema proposto. Inicialmente, foi realizada uma revisão da literatura sobre agentes de IA baseados em LLMs, Sistemas Multiagentes, Sistemas Tutores Inteligentes, RAG e KG-RAG. Também foram estudados *frameworks* para desenvolvimento de agentes, com destaque para LangGraph e CrewAI, além dos LLMs disponíveis considerando as restrições orçamentárias do contexto universitário. Essa etapa serviu de base para as decisões relacionadas à arquitetura e às tecnologias utilizadas.

Na segunda etapa a arquitetura foi projetada e prototipada. Foram definidos agentes especializados responsáveis por diferentes funções do sistema, incluindo um agente orquestrador, um agente recuperador, um agente pedagógico e um agente responsável pela geração de planos. A comunicação entre esses agentes e a estratégia de raciocínio foram definidas de acordo com os requisitos identificados na revisão da literatura, com enfoque no paradigma ReAct. O *framework* LangGraph foi escolhido para o desenvolvimento, pois propicia maior flexibilidade e controle da orquestração dos agentes, mantendo um estado

compartilhado entre eles. No que diz respeito aos LLMs, o modelo Gemini 3.5 Flash foi utilizado, via API da Google AI Studio, por cumprir os critérios orçamentários estabelecidos. O artefato resultante dessa etapa consiste em um protótipo inicial, destinado a validar a arquitetura.

A terceira etapa está em andamento e consiste na preparação da base de conhecimento. Serão coletados e pré-processados os anais do WIT disponíveis publicamente e os materiais relacionados ao Programa Meninas Digitais por meio de processos de extração de texto e segmentação semântica. Na sequência, identificam-se as entidades pedagógicas e as suas respectivas relações com o auxílio de LLMs, de modo a organizá-las em um grafo de conhecimento mediante a utilização de ferramentas como o Neo4j.

A quarta etapa tem o propósito de integrar a técnica KG-RAG à arquitetura MAS. O grafo elaborado na etapa antecedente é conectado a um mecanismo de recuperação que alia a busca por similaridade semântica a uma travessia estruturada do grafo. Esse recurso é disponibilizado aos agentes por intermédio de chamadas de ferramentas. A qualidade da recuperação é validada antes da integração ao sistema em sua totalidade.

Para mitigar o risco de instabilidades sistêmicas previamente à execução dos experimentos, a metodologia incorpora uma fase dedicada a testes técnicos e à validação da engenharia de *software*. Essa etapa transicional abrange a execução de testes unitários e de integração, visando assegurar a robustez da comunicação entre os agentes e a correta interoperabilidade entre o módulo de recuperação e os LLMs. Apenas após a certificação da estabilidade técnica do protótipo, o sistema prosseguirá para a fase de experimentação formal, o que garante que a avaliação empírica incida sobre um artefato tecnológico funcional e confiável.

Adicionalmente, no que tange à etapa final de avaliação, a aferição das métricas de desempenho será conduzida com o suporte de *frameworks* de avaliação automatizada consolidados na literatura para sistemas RAG. Ferramentas específicas para esse propósito, a exemplo do Ragas ou do TruLens, serão empregadas para viabilizar a análise quantitativa da fidelidade contextual, da relevância das respostas e da completude do contexto recuperado. A adoção de tais instrumentos metodológicos assegura maior precisão à validação experimental, minimiza o viés humano na aferição das métricas e confere reprodutibilidade à análise comparativa entre o sistema proposto e a abordagem de referência.

3. RESULTADOS E DISCUSSÃO

O protótipo inicial permitiu validar a estrutura de comunicação entre os agentes e a utilização de LLM como núcleo de processamento. Tornou-se possível assim, demonstrar a viabilidade de construção do sistema proposto, carecendo ainda de outras etapas para que os experimentos sejam conduzidos e resultados robustos sejam obtidos.

Até o momento, foi realizada a revisão bibliográfica e definido o escopo da arquitetura do sistema. Foram especificados quatro agentes especializados: orquestrador, recuperador, pedagógico e gerador de planos. A arquitetura foi prototipada utilizando o *framework* LangGraph, escolhido por proporcionar maior controle sobre a orquestração e o compartilhamento de estado entre os agentes. O modelo Gemini 3.8 Flash foi adotado considerando os critérios orçamentários

estabelecidos para o projeto, atendendo às necessidades do problema proposto e apresentando qualidade satisfatória nas inferências realizadas.

Atualmente, encontra-se em desenvolvimento a preparação da base de conhecimento, incluindo a coleta e o processamento dos anais do WIT e de materiais do Programa Meninas Digitais.

A próxima etapa consiste na construção do grafo de conhecimento e na integração do mecanismo KG-RAG ao protótipo. Posteriormente, serão realizados testes técnicos e experimentos comparativos para avaliar a qualidade da recuperação e das respostas geradas pelo sistema.

4. CONCLUSÕES

O trabalho em andamento apresenta uma abordagem para apoiar a recuperação e a reutilização do conhecimento produzido pelo Programa Meninas Digitais por meio da integração de LLMs, Sistemas Multiagentes e KG-RAG. A principal inovação consiste na utilização conjunta dessas técnicas em um Sistema Tutor Inteligente direcionado ao apoio de educadores e colaboradores.

Até o momento, a arquitetura multiagente foi projetada e prototipada, estabelecendo a base para a integração do conhecimento estruturado. As etapas seguintes permitirão concluir o desenvolvimento e avaliar experimentalmente a solução proposta.

5. REFERÊNCIAS BIBLIOGRÁFICAS

DONG, C.; YUAN, Y.; CHEN, K.; CHENG, S.; WEN, C. How to Build an Adaptive AI Tutor for Any Course Using Knowledge Graph-Enhanced Retrieval-Augmented Generation (KG-RAG). In: INTERNATIONAL CONFERENCE ON EDUCATIONAL AND INFORMATION TECHNOLOGY (ICEIT), 14., 2025. **Proceedings. . .** IEEE, 2025. p.152–157.

LI, X.; WANG, S.; ZENG, S.; WU, Y.; YANG, Y. A survey on LLM-based multi-agent systems: workflow, infrastructure, and challenges. **Vicinagearth**, Berlin/Heidelberg, Germany, v.1, n.1, Oct. 2024.

XI, Z. et al. The rise and potential of large language model based agents: a survey. **Science China Information Sciences**, Berlin/Heidelberg, Germany, v.68, n.2, Jan. 2025.

ZAWACKI-RICHTER, O.; MARíN, V. I.; BOND, M.; GOUVERNEUR, F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? **International Journal of Educational Technology in Higher Education**, Berlin/Heidelberg, Germany, v.16, n.1, Oct. 2019.

ZHANG, Q. et al. A Survey of Graph Retrieval-Augmented Generation for Customized Large Language Models. **arXiv**, Cornell University, 2025.