

Agentes de Inteligência Artificial: Fundamentos, Arquiteturas e Aplicações em Ambientes Acadêmicos

Rafael Junior do Amaral Sampaio¹

¹Faculdade Donaduzzi

Avenida Max Planck, 3797 - Ed. Charles Darwin, 85919-899, Biopark - Toledo, PR

rafaeljunioramaral27@gmail.com

Abstract. *This paper analyzes AI agents for academic environments, emphasizing how they can support tasks such as writing assistance, content organization and learning tracking without replacing students' intellectual effort. We synthesize core components planning, memory, tool integration and evaluation and discuss technical and ethical risks, including hallucinations, dependence and privacy. As a practical contribution, we propose a reference architecture for educational support that incorporates human-in-the-loop controls, source verification and audit trails, and we suggest empirical metrics for evaluating educational impact.*

Resumo. *Este artigo analisa agentes de Inteligência Artificial (IA) em contextos acadêmicos, destacando de que forma esses sistemas podem apoiar atividades (assistência à escrita, organização de conteúdo, acompanhamento da aprendizagem) sem substituir o esforço intelectual do estudante. Sintetizamos componentes centrais, planejamento, memória, integração de ferramentas e avaliação, e discutimos riscos relevantes, como geração de informação não verificada, dependência cognitiva e questões de privacidade. Como contribuição prática, propomos uma arquitetura de referência para uso educacional que incorpora controle humano, verificação de fontes e trilhas de auditoria, além de indicadores para avaliação empírica.*

Palavras-chave: agentes de inteligência artificial; agentes autônomos; educação; revisão narrativa; governança pedagógica.

1. Introdução

O uso de tecnologias de IA na educação enfrenta hoje um problema prático e teórico: há muita discussão sobre o potencial das ferramentas, mas menos orientação operacional sobre como integrá-las sem reduzir o processo de aprendizagem. Em particular, faltam critérios claros de supervisão, rastreamento de fontes e responsabilidades quando se adota agentes que atuam por várias etapas do fluxo de trabalho docente-estudantil.

Neste contexto, é importante distinguir agentes de IA de chatbots tradicionais. Enquanto um chatbot costuma responder a perguntas de forma reativa, um agente proativo planeja subtarefas, mantém e recupera estado (memória), aciona ferramentas externas (buscadores, APIs, editores) e acompanha uma tarefa ao longo do tempo, coordenando múltiplas interações e verificações [Wang et al. 2023].

Essas capacidades tornam os agentes potencialmente úteis para atividades acadêmicas como levantamento bibliográfico, estruturação de rascunhos, organização de estudos e acompanhamento de progresso [OpenAI 2025]. Ao mesmo tempo, introduzem riscos, como geração de informações não verificadas, dependência cognitiva dos estudantes e ameaças à privacidade, que exigem controle, validação humana e políticas institucionais [UNESCO 2024].

O objetivo deste artigo é duplo: (i) sintetizar conceitos e padrões de arquitetura para agentes de IA; e (ii) propor um desenho de referência para aplicações educacionais que preserve a agência do estudante. A tese central defendida é que agentes de IA podem apoiar processos de ensino e aprendizagem sem substituir o esforço intelectual, desde que usados com supervisão humana, critérios de validação e mecanismos de rastreabilidade.

2. Metodologia

Este trabalho adota uma revisão narrativa e exploratória da literatura. A busca bibliográfica concentrou-se em artigos, relatórios técnicos e documentação de plataformas publicados até 2025, usando bases como arXiv, repositórios institucionais e documentação técnica de fornecedores. A seleção foi orientada por relevância para arquiteturas de agentes, uso de ferramentas e questões pedagógicas; priorizamos trabalhos que discutem desenho, riscos e diretrizes para uso educacional. As seções seguintes sintetizam as principais contribuições identificadas (seção de resultados) e discutem implicações para implementação e governança.

3. Fundamentação Teórica

Um agente de IA pode ser entendido como um sistema que recebe um objetivo, observa um contexto, decide uma ação e executa uma sequência de passos para atingir um resultado. Em implementações atuais, esse ciclo normalmente inclui:

- **Planejamento:** decomposição de objetivos em subtarefas.
- **Memória:** armazenamento de informações de curto e longo prazo.
- **Uso de ferramentas:** chamadas a APIs, buscadores, bancos de dados e ambientes de execução.
- **Avaliação:** verificação da qualidade da resposta e ajuste de rota quando necessário.

Pesquisas recentes destacam que a combinação entre raciocínio em etapas, ferramentas externas e feedback iterativo melhora a capacidade de resolver tarefas complexas [Yao et al. 2023, Wang et al. 2023].

4. Resultados

4.1. Arquiteturas de Agentes de IA

Arquiteturas de agentes geralmente seguem um loop de percepção, decisão e ação. Um padrão comum inclui os módulos: *orquestrador*, *memória*, *executores de ferramentas* e *avaliador*. O orquestrador define o próximo passo; a memória recupera contexto; os executores interagem com sistemas externos; e o avaliador mede se o objetivo foi atingido.

Guias técnicos recentes para desenvolvimento de agentes enfatizam práticas como limites de autonomia, validação de saída, logs de auditoria e mecanismo de aprovação humana para ações sensíveis [OpenAI 2025].

4.2. Proposta de Aplicação em Ambiente Acadêmico

Como aplicação, propomos um agente de apoio à produção de artigos com quatro etapas: (1) definição de tema e escopo, (2) levantamento bibliográfico inicial, (3) estruturação das seções e (4) revisão final orientada por checklist.

Para reduzir riscos, o fluxo deve adotar princípios de *human in the loop*: todo trecho gerado passa por revisão do estudante, e todas as referências devem ser verificadas manualmente antes da submissão. Essa abordagem busca equilibrar produtividade e confiabilidade, em linha com diretrizes éticas para uso educacional de IA [UNESCO 2024].

No desenho proposto, o agente não substitui o estudante, mas atua como tutor técnico para tarefas instrumentais. O processo de aprendizagem continua centrado na pessoa, especialmente em atividades de interpretação, argumentação e tomada de decisão. Em termos práticos, o estudante deve registrar:

- qual pergunta foi feita ao agente;
- qual resposta foi recebida;
- quais trechos foram aceitos, modificados ou descartados;
- quais fontes foram verificadas antes de uso final.

Esse registro funciona como evidência de autoria intelectual e de processo de aprendizagem, reduzindo a percepção de “terceirização” do trabalho.

4.3. IA na Educação: Tabu, Riscos e Oportunidades

O uso de IA em escolas gera resistência por motivos legítimos, como risco de plágio, potencial de dependência cognitiva e desigualdades de acesso, e por isso exige respostas institucionais equilibradas. A proibição ampla pode empurrar práticas para a clandestinidade, enquanto políticas bem desenhadas podem permitir usos pedagógicos controlados.

Políticas eficazes diferenciam tipos de atividade: em exercícios formativos o agente pode atuar como tutor oferecendo pistas e exemplos; em avaliações somativas, o uso deve ser restringido ou mediado; em trabalhos de pesquisa, o agente pode apoiar levantamento inicial desde que todas as fontes sejam verificadas manualmente.

Do ponto de vista pedagógico, agentes podem oferecer ganhos concretos quando integrados ao desenho das atividades:

- **Feedback rápido:** retorno imediato sobre erros comuns de escrita e estrutura lógica.
- **Aprendizagem personalizada:** explicações em diferentes níveis de profundidade para cada perfil de estudante.
- **Apoio metacognitivo:** sugestões para planejar estudo, revisar entendimento e monitorar progresso.
- **Inclusão:** suporte linguístico e reformulação de texto para estudantes com necessidades específicas.

Esses benefícios aparecem quando o agente é integrado ao planejamento didático, e não usado de forma isolada. Em outras palavras, a tecnologia precisa estar a serviço de objetivos de aprendizagem explicitamente definidos.

5. Diretrizes de Implementação para Escolas e Docentes

Para viabilizar uso responsável de agentes de IA na educação básica e superior, recomendam-se cinco diretrizes de implementação.

Diretriz 1: Política institucional de uso. A escola deve publicar um documento simples com permissões, restrições e exemplos de uso adequado por tipo de atividade.

Diretriz 2: Transparência obrigatória. Estudantes devem declarar quando utilizaram IA, em quais etapas e com qual finalidade.

Diretriz 3: Avaliação orientada a processo. Parte da nota deve refletir planejamento, justificativas e iterações de melhoria, e não apenas o texto final.

Diretriz 4: Formação docente contínua. Professores precisam de formação prática sobre desenho de prompts, verificação de fontes e critérios de avaliação com IA.

Diretriz 5: Segurança e privacidade. Deve-se evitar inserção de dados pessoais sensíveis em plataformas externas e priorizar ferramentas com controles de proteção de dados.

Essas diretrizes ajudam a converter uma adoção espontânea e desorganizada em uma estratégia pedagógica consistente, auditável e eticamente sustentável [UNESCO 2024, OpenAI 2025, Ministerio da Ciencia, Tecnologia e Inovacao 2021, CETIC.br—NIC.br 2023].

6. Cenário de Uso: Apoio à Produção de Artigos

Como cenário ilustrativo, considere uma turma de ensino superior desenvolvendo artigos curtos. O agente pode apoiar quatro momentos.

No primeiro momento, auxilia na delimitação de tema, sugerindo recortes mais específicos e perguntas de pesquisa viáveis. No segundo, apoia a busca inicial por literatura, propondo palavras-chave e estrutura de fichamento. No terceiro, apoia organização de seções, coerência argumentativa e revisão linguística. No quarto, ajuda a montar checklist de conformidade com o template de submissão.

Mesmo nesse cenário, o critério central permanece: o estudante é responsável por entender, justificar e defender cada afirmação escrita. O agente atua como acelerador de processo, não como autor principal. Esse modelo tende a aumentar produtividade sem abrir mão de rigor acadêmico.

7. Desafios e Limitações

Apesar do potencial, há limitações importantes: geração de referências inexistentes, erros factuais, dependência excessiva do modelo e dificuldades de reproducibilidade. Também há desafios de privacidade, principalmente quando dados sensíveis são enviados para serviços externos.

Assim, recomenda-se: definição clara do papel do agente, trilhas de auditoria, checagem automática de citações e políticas institucionais de uso responsável.

Outro ponto crítico é o risco de dependência cognitiva. Se toda etapa de planejamento e resolução for terceirizada ao agente, o estudante pode reduzir o esforço de elaboração própria. Para mitigar esse efeito, atividades avaliativas devem exigir

explicação oral, comparação de alternativas e reflexão sobre as decisões tomadas ao longo do processo.

Também é necessário considerar desigualdades de acesso. Instituições com infraestrutura limitada podem enfrentar dificuldades para adotar ferramentas de forma equitativa. Nesses casos, políticas de laboratório, horários assistidos e ferramentas de menor custo são medidas importantes para evitar ampliação de disparidades.

No contexto brasileiro, essas medidas podem ser alinhadas a marcos de estratégia digital e de proteção de dados, garantindo que iniciativas educacionais com IA respeitem princípios de segurança, finalidade e transparência [Ministerio da Ciencia, Tecnologia e Inovacao 2021, Brasil 2018]. Estudos nacionais sobre tecnologia em educação também ajudam a contextualizar barreiras e oportunidades de implementação [CETIC.br—NIC.br 2023, SBC OpenLib 2023].

8. Conclusão

Agentes de IA ampliam capacidades de apoio a tarefas acadêmicas ao combinar raciocínio em etapas, memória contextual e integração com ferramentas. Contudo, este artigo sustentou que seu valor pedagógico depende de requisitos operacionais: supervisão humana, critérios explícitos de validação, rastreamento de interações e políticas institucionais que protejam privacidade e equidade.

Como trabalho futuro, propõe-se avaliar empiricamente a arquitetura sugerida em turmas reais, mensurando indicadores como ganho de tempo, melhoria na qualidade textual, fidedignidade das referências e efeitos sobre a autonomia dos estudantes. Essas avaliações são essenciais para validar se os agentes atingem seu objetivo de apoiar, e não substituir, o esforço intelectual.

Em síntese, a adoção responsável de agentes em ambientes acadêmicos exige uma combinação entre autonomia estudantil, mediação docente e governança institucional, condicionada a práticas de verificação e transparência.

Referências

- Brasil (2018). Lei no 13.709, de 14 de agosto de 2018 (lei geral de proteção de dados pessoais). Presidência da República. Acesso em: 26 abr. 2026.
- CETIC.br—NIC.br (2023). Pesquisa tic educação. Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação. Acesso em: 26 abr. 2026.
- Ministerio da Ciencia, Tecnologia e Inovacao (2021). Estratégia brasileira de inteligência artificial (ebia). Governo Federal. Acesso em: 26 abr. 2026.
- OpenAI (2025). Building agents. OpenAI Platform Documentation. Acesso em: 26 abr. 2026.
- SBC OpenLib (2023). Tecnologia qr code como ferramenta para o gerenciamento de frequência em ambientes educacionais. SBC OpenLib. Acesso em: 26 abr. 2026.
- UNESCO (2024). Guidance for generative ai in education and research. UNESCO. Acesso em: 26 abr. 2026.
- Wang, L., Liu, W., He, Y., Wang, S., et al. (2023). A survey on large language model based autonomous agents. arXiv. Acesso em: 26 abr. 2026.

Yao, S., Zhao, J., Yu, D., Du, N., et al. (2023). React: Synergizing reasoning and acting in language models. arXiv. Acesso em: 26 abr. 2026.