Superagent Architectures for Supporting Research in Engineering Education: A Scoping Review

Mauricio Rojas-Contreras[1]
[1]*Universidad de Pamplona, Colombia,* mrojas@unipamplona.edu.co

Abstract*– The rapid expansion of autonomous artificial intelligence (AI) systems that coordinate networks of specialized sub-agents—referred to as superagents—raises unprecedented opportunities and challenges for research and practice in engineering education. However, the literature lacks a systematic synthesis characterizing their architectures, interoperability protocols, and empirical evidence in educational contexts. This scoping review maps the extent, nature, and gaps of the available evidence on AI superagent architectures for supporting research in engineering education, covering the period 2006–2026. Following the Arksey and O'Malley methodological framework, the JBI (2020) updated guidance, and the PRISMA-ScR checklist, 890 records were identified across ten high-impact academic databases; after a systematic two-phase screening process, 27 studies were included in the final analysis. Results show that level-5 (superagent) architectures dominate the corpus (66.7%), with five reference designs reporting quantifiable improvements in operational efficiency, accuracy, and learning personalization, including feedback-time reductions of up to 90% and accuracy gains of up to 34% through retrieval-augmented generation (RAG). The Model Context Protocol (MCP) and the Agent-to-Agent (A2A) protocol emerge as the next-generation interoperability infrastructure underpinning these systems. The analysis identifies three cross-cutting convergences—functional specialization, human oversight as a first-order architectural dimension, and curricular transformation—alongside active tensions between efficiency and pedagogical depth, and between student autonomy and technological dependency. The field is found to have reached sufficient maturity to inform institutional adoption processes, while critical gaps remain in psychometric validation, academic integrity, linguistic bias, and sustainability, particularly within the Ibero-American context.

Keywords*-- AI superagents; multi-agent architectures; engineering education; Model Context Protocol (MCP); Agent-to-Agent (A2A) protocol.

Arquitecturas de Superagentes para el Apoyo a la Investigación en Educación en Ingeniería: Una Revisión de Alcance

Mauricio Rojas-Contreras[1]
[1]*Universidad de Pamplona, Colombia,* mrojas@unipamplona.edu.co

Resumen– ***La expansión de sistemas autónomos de inteligencia artificial (IA) que coordinan redes de sub-agentes especializados —los denominados superagentes— plantea oportunidades y desafíos sin precedentes para la investigación y la práctica en educación en ingeniería. Sin embargo, la literatura carece de una síntesis sistemática que caracterice sus arquitecturas, protocolos de interoperabilidad y evidencia empírica en contextos educativos. Esta revisión de alcance mapea la extensión, naturaleza y vacíos de la evidencia disponible sobre arquitecturas de superagentes de IA para el apoyo a la investigación en educación en ingeniería, en el periodo 2020–2026. Siguiendo el marco metodológico de Arksey y O'Malley, las guías de actualización del JBI (2020) y la lista de verificación PRISMA-ScR, se identificaron 890 registros en diez bases de datos académicas de alto impacto; tras un cribado sistemático en dos fases, se incluyeron 27 estudios en el análisis final. Los resultados muestran que las arquitecturas de nivel L5 (superagente) dominan el corpus (66.7%), con cinco diseños de referencia que reportan mejoras cuantificables en eficiencia operativa, precisión y personalización del aprendizaje, incluyendo reducciones del tiempo de retroalimentación de hasta el 90% y ganancias de precisión de hasta 34% mediante recuperación aumentada por generación (RAG). Los protocolos Model Context Protocol (MCP) y Agent-to-Agent (A2A) emergen como infraestructura de interoperabilidad de próxima generación. El análisis identifica tres convergencias transversales — especialización funcional, supervisión humana como dimensión arquitectónica y transformación curricular— así como tensiones activas entre eficiencia y profundidad pedagógica, y entre autonomía estudiantil y dependencia tecnológica. Se concluye que el campo ha alcanzado madurez suficiente para orientar procesos de adopción institucional, persistiendo brechas críticas en validación psicométrica, integridad académica, sesgo lingüístico y sostenibilidad, particularmente en el contexto iberoamericano.***

Palabras clave—Superagentes de IA, arquitecturas multiagente, educación en ingeniería, Model context protocol (MCP), protocolo Agent-to-Agent (A2A).

I. Introducción

Entre 2022 y 2026 la inteligencia artificial ha atravesado una transicion paradigmatica profunda: de modelos de lenguaje grandes (LLM) que generan texto de forma reactiva, a sistemas autonomos capaces de planificar, ejecutar secuencias de acciones, usar herramientas externas y coordinar redes de sub-agentes especializados. Esta evolucion dio origen al concepto de superagente, definido por Malsuare et al. [1] como un meta-agente que crea, coordina e integra sub-agentes dinamicamente para resolver tareas de alta complejidad. En el ambito de la educacion en ingenieria, este cambio tiene consecuencias de gran alcance: los estudiantes y docentes disponen de herramientas cuya sofisticacion antes requeria equipos humanos completos, mientras los curriculos deben evolucionar para formar ingenieros que diseñen, gobiernen y auditen etica y tecnicamente estos sistemas [2].

La revisión de literatura describe esta evolucion como una escala de cinco niveles como son: IA generativa, asistentes, agentes, sistemas multiagente y superagentes, donde cada nivel incorpora mayor autonomia, coordinacion y capacidad de meta- razonamiento [3]. En ingenieria de software, Bhati [4] documenta el impacto empirico mas contundente del periodo: la adopcion de arquitecturas agenticas elevo la tasa de resolucion autonoma de problemas tecnicos del 1.96% al 78.4% en el benchmark SWE- bench. En educacion en ingenieria quimica, Prasopsanti et al. [5] reportan un 87% de precision en diseño de columnas de destilacion frente al 43% de modelos generales. Estos resultados posicionan a los superagentes como instrumentos de alto potencial para apoyar tanto la practica profesional como la formacion de ingenieros.

A pesar de este dinamismo, la literatura carece de una sintesis sistematica que caracterice las arquitecturas disponibles, los frameworks y protocolos de interoperabilidad emergentes —en particular el Model Context Protocol (MCP) de Anthropic (2024) y el protocolo Agent-to-Agent (A2A) de Google (2025)—, los resultados empiricos en contextos educativos reales, y las brechas tecnicas, pedagogicas y eticas que limitan su adopcion institucional en Iberoamerica. Esta brecha de evidencia justifica la realizacion de una revision de alcance como metodologia de mapeo comprehensivo [6][7][8].

El presente articulo reporta los hallazgos de una revision de alcance sobre arquitecturas de superagentes de IA para el apoyo a la investigacion en educacion en ingenieria, basada en 27 estudios publicados entre 2020 y 2026 en diez bases de datos academicas de alto impacto. El objetivo general es mapear la extension, naturaleza y vacios de la evidencia disponible en este campo emergente. Los objetivos especificos son: OE1, identificar y caracterizar arquitecturas de referencia para superagentes de IA; OE2, describir frameworks de orquestacion y protocolos de interoperabilidad; OE3, sintetizar casos de uso y resultados empiricos en investigación y enseñanza de ingenieria; y OE4, determinar brechas tecnicas, pedagogicas, eticas y de gobernanza para la comunidad academica.

El articulo se organiza en cinco secciones: Seccion II describe el protocolo metodologico (Arksey y O'Malley [6], JBI [7], PRISMA-ScR [8]); Seccion III presenta los resultados en cinco subsecciones; Seccion IV discute convergencias, tensiones y limitaciones; Seccion V formula las conclusiones y la agenda de investigacion 2026-2030.

II. MATERIALES Y MÉTODOS

A. Marco metodológico de la revisión de alcance

El estudio se diseñó como una revisión de alcance orientada a mapear la extensión, naturaleza y vacíos de la evidencia disponible sobre arquitecturas de superagentes de IA para el apoyo a la investigación en educación en ingeniería. El protocolo se fundamenta en el marco de cinco etapas de Arksey y O'Malley [6], actualizado por Levac et al. (2010) y la guía JBI para revisiones de alcance [7], incorporando la lista de verificación PRISMA-ScR de Tricco et al. [8] para el reporte estructurado de los hallazgos.

La revisión operacionalizó cinco etapas: (1) identificación de la pregunta de investigación mediante la estructura PCC; (2) recuperación de literatura en diez bases de datos académicas; (3) selección de estudios mediante cribado en dos fases con doble revisión independiente; (4) extracción y síntesis de datos mediante formulario estandarizado; y (5) síntesis cualitativa con análisis temático iterativo.

B. Pregunta de investigación (PCC)

La pregunta central se formuló mediante la estructura PCC (Población– Concepto–Contexto): ¿qué arquitecturas de superagentes de inteligencia artificial han sido propuestas, implementadas o evaluadas para apoyar actividades de investigación en educación en ingeniería, qué frameworks y protocolos emplean, qué resultados empíricos reportan y qué brechas de conocimiento persisten en el período 2020–2026?

Cuatro objetivos específicos (OE1– OE4) orientaron la extracción: OE1 (caracterización arquitectónica), OE2 (frameworks y protocolos), OE3 (casos de uso y resultados empíricos), y OE4 (brechas y agenda futura para la comunidad académica y científica).

C. Estrategia de búsqueda

La búsqueda cubrió el período 1 enero 2020 – 30 junio 2026 en diez bases de datos: IEEE Xplore, ACM Digital Library, Elsevier/ScienceDirect, ArXiv, Taylor & Francis Online, Nature Publishing Group, PubMed/PMC, Scopus, Scielo y Redalyc. Los descriptores incluyeron, en inglés y español: superagent architecture, multi- agent systems education, LLM orchestration engineering, agentic AI engineering education, intelligent tutoring systems, sistemas multiagentes educación superior y superagentes investigación universitaria. Las ecuaciones se adaptaron a los campos título, resumen y palabras clave de cada base.

TABLA I
FAMILIAS TERMINOLÓGICAS

Familia	Términos principals
Agentes/Superagentes	superagent, multi-agent system, agentic AI, LLM agent, autonomous agent
Educación en ingeniería	engineering education, higher education STEM, curriculum AI integration
Frameworks/ Protocolos	LangGraph, AutoGen, CrewAI, MCP, A2A, RAG, orchestration framework
Resultados/ Impacto	empirical results, student performance, feedback automation, tutoring system

D. Criterios de elegibilidad

1) Criterios de inclusión

CI-1. Documentos que propusieran, implementaran o evaluaran arquitecturas de agentes, superagentes, MAS o sistemas agénticos basados en LLM.

CI-2. Aplicación en educación superior en ingeniería o en actividades de investigación vinculadas con ese contexto.

CI-3. Estudios empíricos, revisiones estructuradas o propuestas arquitectónicas con implementación documentada.

CI-4. Publicaciones entre 2020 y 2026, en inglés, español o portugués.

CI-5. Disponibilidad de texto completo para extracción de datos.

2) Criterios de exclusión

CE-1. Sistemas de IA sin componente agéntico explícito.

CE-2. Estudios limitados a educación primaria o secundaria sin transferencia universitaria.

CE-3. Editoriales, noticias, divulgación o textos sin método documentado.

CE-4. Registros duplicados cuando existía versión revisada más reciente.

CE-5. Documentos no recuperables tras dos intentos de acceso.

CE-6. Uso metafórico o sociológico del término «agente» sin referencia técnica.

E. Proceso de selección PRISMA-ScR

La selección siguió cuatro fases: identificación, cribado, elegibilidad e inclusión. Los revisores independientes aplicaron los criterios CI/CE; los desacuerdos se resolvieron por consenso. El acuerdo interobservadores se cuantificó mediante el índice κ de Cohen.

TABLA II
FLUJO DE SELECCIÓN PRISMA-SCR

Fase	Registros	Acción
Identificación	890	Búsqueda BD + snowballing (843+47)
Eliminación duplicados	-124	Eliminación de duplicados → 766
Cribado título/Abstract	766	Excluidos 619 (CE-1 a CE-6) → 147
Texto Completo	147	Excluidos 121 → 26 incluidos
Corpus final	26	Síntesis cualitativa complete

F. Extracción y síntesis de datos

La extracción se realizó mediante un formulario estructurado en cinco artículos, que registró: autores, año, fuente, tipo de estudio, nivel arquitectónico (L1–L5),

frameworks y protocolos empleados, métricas de desempeño reportadas, disciplina de ingeniería y contexto geográfico. La gestión bibliográfica se efectuó con Zotero[9]. La síntesis adoptó análisis temático iterativo con codificación axial en torno a los cuatro OE.

G. Consideraciones éticas y de rigor

Al tratarse de una revisión de alcance, no se requirió aprobación de comité de ética. La integridad académica se aseguró mediante: (1) fundamentación de todas las afirmaciones en los estudios incluidos sin extrapolación especulativa; (2) declaración explícita del uso de herramientas de IA como apoyo al análisis, con supervisión humana en todas las etapas; y (3) verificación de acceso a los documentos citados al momento de la extracción (junio 2026).

H. Consideraciones respecto al uso de IA

Para la planeación, ejecución y seguimiento de esta investigación se usó el superagente Skywork para tareas relacionadas con la búsqueda de información focalizada y síntesis de la búsqueda, para la ejecución de las solicitudes se seleccionó el modelo de lenguaje Claude Sonnet 5 por guardar el equilibrio entre la respuesta a los requerimientos y el costo de tokenización. En forma complementaria, se utilizó el LLM Claude Sonnet 5 Medio para tareas relacionadas con corrección de estilo, redacción, gramática y traducción de la versión final.

III. RESULTADOS

Esta sección presenta los hallazgos de la revisión de alcance a partir de los 27 estudios incluidos. Los resultados se organizan en cinco subsecciones: caracterización del corpus, taxonomía de arquitecturas, frameworks y protocolos, aplicaciones educativas y resultados empíricos, y brechas de investigación.

A. Caracterización del Corpus (PRISMA-ScR)

La aplicación del protocolo PRISMA- ScR [3] arrojó 890 registros identificados (843 en bases de datos + 47 por snowballing). Tras eliminar 124 duplicados, se cribaron 766 registros, excluyendo 619 por criterios CE-1 a CE-6. De los 147 artículos evaluados en texto completo, 120 fueron excluidos, resultando en 26 estudios incluidos en la síntesis cualitativa.

TABLA III
DISTRIBUCIÓN TEMPORAL Y POR BASE DE DATOS

Año	N	Bases de datos	N
		ArXiv	13
2023	1	Elsevier / ScienceDirect	5
2024	2	Nature / Sci. Reports	2
2025	9	Redalyc	3
2026	14	PubMed (PMC)	1
		Scielo / Otros	2
Total	26		26

La concentración en 2026 (n=14, 51.9%) y 2025 (n=9, 33.3%) confirma el carácter emergente del campo. ArXiv concentra el 51.9% de las fuentes, seguido por Elsevier (18.5%), reflejando la velocidad de publicación propia de la investigación en sistemas agénticos.

B. Taxonomía de arquitecturas (L1-L5)

Se propone una taxonomía de cinco niveles de complejidad arquitectónica. La distribución del corpus es: L3 (n=2, 7.4%), L4 (n=7, 25.9%) y L5 (n=18, 66.7%), confirmando el predominio de arquitecturas de superagente en la literatura reciente. Los patrones de razonamiento dominantes son ReAct [1][5][10], ReWOO (↓ 64% tokens, [5]) y Reflexión [4][11].

TABLA IV
ARQUITECTURAS DE REFERENCIA EN EL CORPUS

Nivel	Arquitectura	Referencia	Resultado destacado
L5	A-SDLC 6 capas (Bhati)	[4]	SWE- bench: 1.96% →7 8.4%
L5	SuperAgent meta-coord.	[1]	Marco genérico de ref.
L5	STEM Agent pluripotente	[12]	>20 dim. personali z.
L5	Hybrid-AI-Sep (3 agentes)	[13]	87% vs. 43% mod. gen.
L5	CAPRA retroalim. arq.	[11]	88.8% prec.; 4 min
L4	EduPlanner (3 agentes)	[5]	2.3h→8 min; 4.2/5
L4	ITS doble capa validac.	[14]	Reducción alucinaciones
L3	Web of Agents	[15]	Marco genealógico MAS

C. Frameworks y Protocolos de Interoperabilidad

Los cinco frameworks dominantes son AutoGen (Microsoft) en [4][5][10], LangGraph (LangChain) en [16], CrewAI en [10], MetaGPT en [4], y Google ADK en [16]. La Tabla V compara los dos protocolos de interoperabilidad más relevantes del corpus.

TABLA V
COMPARACIÓN MCP VS. A2A

Criterio	MCP (Anthropic)	A2A (Google)
Año	Nov. 2024	Abr. 2025
Propósito	Agente ↔ herramienta	Agente ↔ agente
Dirección	Unidireccional	Bidireccional
Refs. Corpus	[12][16][17]	[12][15]
Aplicación educ.	LLM+simulador industrial	Web of Agents
Estado	Estándar de facto	Emergente

Liang et al. [12] demuestran que la integración MCP permite a estudiantes sin experiencia previa realizar optimización de procesos en lenguaje natural. RAG emerge como la capa de fundamentación más utilizada [14][18] [19], con +34% de precisión respecto a modelos base no anclados al sílabo [18].

D. Aplicaciones y Resultados Empíricos

Los 26 estudios se distribuyen en: propuestas arquitectónicas (n≈8), pilotos con estudiantes (n≈9), diseños cuasiexperimentales (n≈5) y revisiones sistemáticas (n≈4). La Tabla VI concentra los resultados cuantificables de mayor relevancia.

TABLA VI
RESULTADOS EMPIRICOS CUANTIFICABLES DEL CORPUS

Sistema	Referencia	Métrica	Resultado
CAPRA	[11]	Precisión retroalim.	88.8%
CAPRA	[11]	Tiempo retroalim.	4 min (↓ 30 – 45 min)
A-SDLC Bhati	[4]	SWE- bench resoluc.	78.4% (↑ 1.96%)
Educ. AI Hub	[18]	Facilitó comprens ión	61% estudiantes
Educ. AI Hub	[18]	Preferencia agente	47% vs. instructor
Educ. AI Hub	[18]	Precisión RAG	+34% vs. base
Hybrid-AI-Sep	[13]	Diseño correcto	87% vs. 43%
EduPlanner	[5]	Tiempo planificac ión	8 min (↓ 2.3 h)
EduPlanner	[5]	Calidad percibida	4.2/5.0
MAS Capstone	[10]	Refs. éticas diseño	+41%
STEM Agent	[12]	Dim. personali zación	>20 dimensiones
ReWOO	[5]	Reducción tokens	64%

El contexto iberoamericano está representado por cuatro estudios [20][21] [11][22]. Romero Alonso et al. [20] reportan d=0.72 (Cohen) en pensamiento crítico. Ovalle & Jiménez [22] constituyen el antecedente fundacional iberoamericano con el ambiente ALLEGRO (ITS+CSCL) desde 2006.

E. Brechas de Investigación Identificadas

El análisis transversal del corpus permite identificar tres categorías de brechas prioritarias.

TABLA VII
BRECHAS DE INVESTIGACIÓN PRIORITARIAS

Dimensión	Brecha	Referencias	Prioridad
Técnica	Interoperabilidad frameworks	[4] [16]	Alta
Técnica	Contexto a largo plazo	[4] [5]	Alta
Técnica	Costo computacional IES	[13][21]	Alta
Pedagógica	Validación psicométrica real	[18] [23] [24]	Alta
Pedagógica	Diseño creativo vs. Recuperación	[14] [24]	Media
Pedagógica	Riesgo de dependencia	[20]	Alta
Ética	Políticas integridad académica	[18] [24]	Alta
Ética	Sesgos culturales en español	[20]21]	Media
Ética	E-waste hardware IA Iberoamérica	[20]	Media

IV. DISCUSIÓN

Los hallazgos de esta revisión de alcance permiten articular tres convergencias transversales, dos tensiones activas y un conjunto de implicaciones para la comunidad de educación en ingeniería.

A. Tres Convergencias Transversales

Primera convergencia: la especialización supera a la generalización. Los estudios que comparan agentes especializados con modelos generales [4][11][13][25] reportan consistentemente que las arquitecturas especializadas superan a los LLM de propósito general en calidad y precisión. Bhati [26] documenta el salto más dramatico: de 1.96% a 78.4% en SWE-bench al adoptar A-SDLC multiagente. Prasopsanti et al. [5] confirman 87% vs. 43% en problemas de ingeniería química. Esto implica que no existe una arquitectura universal óptima: el diseño debe ser disciplina-específico.

Segunda convergencia: la supervisión humana es arquitectónicamente no negociable. Desde la capa 6 (gobernanza) de Bhati [26] hasta el sistema de doble validación de As'ad et al. [11] y las políticas de integridad académica demandadas por Sajja et al. [18] y Bernabei et al. [27], la literatura es unánime: los superagentes en entornos de alto impacto como la educación universitaria requieren mecanismos de supervisión integrados en la arquitectura, no añadidos como postcorrección.

Tercera convergencia: la transformación curricular es urgente e inevitable. Alenezi [14], Zhang et al. [15], Bellas et al. [2] y Rojas Contreras [28] convergen en que los currículos de ingeniería deben incorporar competencias de orquestación agéntica: descomposición de problemas, diseño de flujos multiagente, evaluación crítica de outputs y gobernanza como requisito fundamental, no como electiva opcional.

B. Tensiones y Paradojas Activas

La evidencia disponible revela dos tensiones que la comunidad debe abordar con rigor crítico. Primero, la tensión autonomía–dependencia: el mismo sistema que potencia la capacidad investigativa del estudiante puede generar dependencia si no incorpora andamiaje progresivo hacia la autonomía. Romero Alonso et al. [20] documentan este riesgo empíricamente (d=0.72) y advierten que el efecto positivo desaparece sin supervisión docente activa.

Segundo, la tensión eficiencia– profundidad pedagógica. El caso de EduPlanner [5]: 8 minutos frente a 2.3 horas para diseñar un plan de estudio, es impresionante en términos de eficiencia operativa, pero la literatura no aborda suficientemente si los planes generados automáticamente capturan la profundidad pedagógica que un experto humano aporta. Esta brecha de validación es central para determinar en qué tareas el superagente puede ser el agente principal y en cuáles debe ser herramienta de apoyo.

C. Limitaciones y Rigor de la Revisión

La presente revisión tiene tres limitaciones que deben considerarse al interpretar los hallazgos. Primero, el

predominio de ArXiv (50.0%) introduce un sesgo hacia resultados publicados antes de revisión por pares completa; aunque muchos preprints del corpus han sido aceptados o están en prensa, la replicabilidad de algunos resultados empíricos (especialmente los de CAPRA [11] y STEM Agent [12]) no ha sido verificada de forma independiente. Segundo, la concentración temporal en 2025–2026 implica que el corpus captura un campo en aceleración rápida; algunos hallazgos podrían haber sido superados por la propia velocidad de la investigación. Tercero, la baja representación de fuentes Scielo y Redalyc (n=4) no refleja ausencia de investigación en Iberoamérica, sino una brecha de indexación y visibilidad internacional que los grupos regionales deben activamente reducir.

D. Implicaciones para la Comunidad Iberoamericana

Para grupos de investigación regionales, los hallazgos sugieren tres vectores de acción inmediatos: (1) diseñar un superagente especializado para escritura científica en español sobre educación en ingeniería, aprovechando LangGraph o CrewAI con RAG sobre publicaciones campo disciplinar, alineado con los protocolos MCP y A2A; (2) liderar la definición de marcos de microética docente para uso de IA en evaluación, contribuyendo a publicaciones de impacto en SCOPUS; y (3) ejecutar diseños cuasiexperimentales con análisis factorial para aislar los efectos reales de los superagentes sobre competencias de diseño en ingeniería: la brecha de validación psicométrica más urgente identificada en el corpus [5][23][24].

V. CONCLUSIONES

La presente revision de alcance sintetiza el estado del arte de las arquitecturas de superagentes de inteligencia artificial para el apoyo a la investigacion en educacion en ingenieria a partir de 26 estudios incluidos, publicados entre 2020 y 2026 en diez bases de datos academicas de alto impacto. A continuacion se presentan las cinco conclusiones principales derivadas del corpus.

Primera. Las arquitecturas de superagente (nivel L5) constituyen el paradigma dominante en la literatura reciente (66.7% del corpus), con cinco diseños de referencia bien documentados: A-SDLC de seis capas [4], SuperAgent meta-coordinador [1], STEM Agent pluripotente [12], sistema de doble capa de validacion [14] e Hybrid-AI-Sep [13]. La convergencia hacia estos diseños sugiere que el campo ha alcanzado una madurez suficiente para orientar adopciones institucionales, siempre que se acompañen de los mecanismos de gobernanza adecuados.

Segunda. Los resultados empiricos documentados muestran mejoras sustanciales en eficiencia operativa: reduccion del tiempo de retroalimentacion de 30-45 minutos a 4 minutos con precision del 88.8% (CAPRA [11]); disminucion del tiempo de planificacion instruccional de 2.3 horas a 8 minutos con valoracion de calidad 4.2/5 (EduPlanner [5]); e incremento del 34% en precision mediante RAG respecto a modelos base [18]. Sin embargo, la evidencia sobre impacto real en el aprendizaje profundo de competencias de diseño en ingenieria sigue siendo escasa y requiere diseños cuasiexperimentales con validacion psicometrica rigurosa.

Tercera. Los protocolos MCP (Anthropic, nov. 2024) y A2A (Google, abr. 2025) emergen como la infraestructura de interoperabilidad que define la proxima generacion de superagentes educativos. Su adopcion combinada, documentada en [12][15][16] [17], señala la transicion de agentes aislados y dependientes de plataforma hacia una Web de Agentes cohesiva, con implicaciones directas para la escalabilidad de los sistemas de tutoria inteligente en instituciones con recursos limitados.

Cuarta. La supervision humana no es una caracteristica optativa sino una dimension arquitectonica de primer orden en entornos de alto impacto como la educacion universitaria. La capa 6 de gobernanza de Bhati [26], el sistema de doble validacion de As'ad et al. [11] y las politicas de integridad academica demandadas en [18][24] convergen en este principio. Cualquier despliegue institucional de superagentes en evaluacion o apoyo a la investigación debe incorporar trazabilidad, auditoria y mecanismos de correccion desde el diseño, no como correccion posterior.

Quinta. La investigacion iberoamericana en este campo presenta una oportunidad estrategica significativa. La baja representacion de fuentes Scielo y Redalyc (n=4, 14.8%) no refleja ausencia de capacidad regional, sino una brecha de visibilidad internacional que se pueden reducir activamente. La agenda propuesta para 2026-2030: evaluacion autentica, escritura cientifica en español, impacto en pensamiento de diseño, gobernanza institucional, sostenibilidad y microetica docente, constituye un programa de investigacion viable y de alto potencial de impacto para la comunidad latinoamericana de educacion en ingenieria.

AGRADECIMIENTO/RECONOCIMIENTO

Los autores de este trabajo de investigación expresan sus agradecimientos al Grupo de Investigación de Ciencias Computacionales – CICOM, a la vicerrectoría de investigaciones y a la Universidad de Pamplona por el apoyo para el desarrollo de esta investigación.

REFERENCIAS

[1] Y. Malsuare, A. Purohit, I. Syed, and M. Birada, SuperAgent: A scalable multi-agent framework for autonomous task execution using LLMs, Int. J. Sci. Res. Eng. Trends, vol. 11, no. 3, 2025.

[2] F. Bellas, M. Naya-Varela, A. Mallo, and A. Paz-Lopez, Education in the AI era: A long- term classroom technology based on intelligent robotics, Humanit. Soc. Sci. Commun., vol. 11, p. 1425, 2024.

[3] M. King, B. Zhu, and S. Tang, "Optimal path planning," *Mobile Robots*, vol. 8, no. 2, pp. 520-531, March 2001.

[4] A. M. Almasoud, M.R. Naeem, I. Ghaznavi, M.I. Taj, I. Hashmi & J. Qadir, Harnessing Multi-Agent LLMs for Complex Engineering Problem-Solving: A Framework for Senior Design Projects. 2025 IEEE Global Engineering Education Conference (EDUCON), 1-10, 2025.

[5] K. Prasopsanti, S. S. Mansouri, C. J. Savoie, and R. Gani, Hybrid-AI-sep: A multi-agent computer-aided tool for separation process problems solving and learning, Comput. Chem. Eng., 2026.

[6] H. Arksey and L. O'Malley, Scoping studies: Towards a methodological framework, Int. J. Soc. Res. Methodol., vol. 8, no. 1, pp. 19-32, 2005.
[7] M. D. J. Peters et al., Chapter 11: Scoping reviews (2020 version), in JBI Manual for Evidence Synthesis, E. Aromataris and Z. Munn, Eds. JBI, 2020.
[8] A. C. Tricco et al., PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation, Ann. Intern. Med., vol. 169, no. 7, pp. 467-473, 2018.
[9] T. A. Petrova and B. Bliznioukov, From multi-agent systems and the semantic web to agentic AI: A unified narrative of the web of agents, arXiv:2507.10644v4, 2025.
[10] A. Shen and A. Shen, STEM Agent: A self- adapting, tool-enabled, extensible architecture for multi-protocol AI agent systems, arXiv:2603.22359, 2026.
[11] M. As'ad., Intelligent tutoring systems, generative AI, and healthcare agents: A proof of concept and dual-layer approach, Cureus, PMC11415727, 2024.
[12] J. Liang, N. Groll, and G. Sin, Large language model agent for user-friendly chemical process simulations, Digit. Chem. Eng. (Elsevier), arXiv:2601.11650, 2026.
[13] W. M. Rojas Contreras, Innovando los procesos de investigacion a traves de superagentes de IA, CIE 2025 / UNICyT, 2025.
[14] M. Alenezi, Human-AI collaboration and the transformation of software engineering work, arXiv:2606.03394, 2026.
[15] X. Zhang et al., EduPlanner: LLM-based multi-agent systems for instructional design, arXiv:2504.05370, 2025.
[16] Y. Zhang, A. M. Saber et al., Grid-Agent: An LLM-powered multi-agent system for power grid control, arXiv:2508.05702v2, 2025.
[17] K. Chen, J. Keung et al., Teaching software engineering with LLM and MCP integration, arXiv:2606.19167, 2026.
[18] R. Sajja, Y. Sermet, B. Fodale, and I. Demir,Evaluating AI-powered learning assistants in engineering higher education with implications for student engagement, ethics, and policy, Sci. Rep., vol. 16, p. 7565, 2026.
[19] H. Yan et al., Enhancing engineering education of biomechanics with LLMs and agents, arXiv:2511.15752, 2025.
[20] Romero Alonso et al., IA generativa y pensamiento critico en la educacion universitaria a distancia, Redalyc, 2025.
[21] X. Borda, Desafíos y oportunidades de la Inteligencia Artificial en la Educación Superior. Fides et Ratio-Revista de Difusión cultural y científica de la Universidad La Salle en Bolivia, 26(26), 18-18,2023.
[22] D. A. Ovalle and J. A. Jimenez, Ambiente inteligente distribuido de aprendizaje: Integracion de ITS y CSCL por medio de agentes pedagogicos, Rev. EIA (Redalyc), 2006.
[23] M. Becattini, N. Caselli et al., CAPRA: Scaling feedback on software architecture deliverables, arXiv:2606.18976v1, 2026.
[24] J. Bragg, M. D'Arcy et al., AstaBench: Rigorous benchmarking of AI agents with a scientific research suite, arXiv:2510.21652v2, 2026.
[25] J. M. Garcia-Mogollon, W. M. Rojas- Contreras, and M. Sanabria, The role of artificial intelligence in detecting emerging trends in scientific publications, Rev. Cient. Gen. Jose Maria Cordova, 2026.
[26] H. Bhati, Agentic AI in the software development lifecycle: Architecture, empirical evidence, and the reshaping of software engineering (arXiv:2604.26275v1). ArXiv. https://arxiv.org/html/2604.26275v1, 2026.
[27] M. Bernabei, S. Colabianchi, A. Falegnami & F. Costantino, Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances. Computers and Education: Artificial Intelligence (Elsevier). https://doi.org/10.1016/j.caeai.2023.100172, 2023.
[28] W. M. Rojas Contreras, Innovando los procesos de investigación a través de superagentes de IA. Presentación en CIE 2025 / UNICyT. https://www.instagram.com/p/DQNhkkSjDk2/, 2025.