

# Review of Multi-Agent Systems (MAS) in Education

Yujia Yang<sup>1</sup>, Jianwei Zhang<sup>1</sup>, Yifan Zhang<sup>1\*</sup>

<sup>1</sup> Smart Learning Institute of Beijing Normal University, China

[cqswuyu@163.com](mailto:cqswuyu@163.com); [2361653573@qq.com](mailto:2361653573@qq.com); [\\*zhangyifan@bnu.edu.cn](mailto:*zhangyifan@bnu.edu.cn)

**Abstract:** This systematic review synthesizes the distribution of Multi-Agent Systems (MAS) studies in education from 2015 to 2025, analyzing 523 peer-reviewed studies preliminarily. The results reveal a significant surge in publications after 2024 ( $n = 112$  in 2025), possibly driven by the integration of Large Language Models (LLMs) and accessible agentic frameworks. Higher education dominates the research landscape (61.57%), partly due to its infrastructure and flexible ethical constraints, while the K-12 and vocational sectors remain underexplored. Pedagogically, MAS applications centralize subject-specific Knowledge Mastery ( $n = 128$ ) and Learning Process Optimization ( $n = 91$ ), while placing less emphasis on broader 21st-century competencies such as collaboration and ethical development. This study highlights a critical disconnect between technical advancements and holistic educational needs, aiming to provide a landscape for bridging the gap between MAS innovation and pedagogical effectiveness in the generative AI era.

**Keywords:** Systematic Literature Review; LLM-based Agents; Multi-Agent Systems; Educational Technology

## 1. Introduction and Related Work

Multi-Agent Systems (MAS), rooted in distributed artificial intelligence (AI), integrate Agents, Environments, Interactions, and Organizations (AEIO model) and operate through decentralized cooperation (Hamal et al., 2021). While early agents relied on symbolic AI or conventional machine learning approaches (Jiang et al., 2024), the integration of Large Language Models (LLMs) has revolutionized the field. By incorporating LLMs, MAS now benefits from enhanced collaboration among specialized agents, more sophisticated reasoning capabilities, and dynamic adaptation to complex environments (Li et al., 2024). These LLM-based MAS surpass the limitations of single-agent systems, demonstrating remarkable success in addressing complex tasks.

In educational contexts, pedagogical agents (PAs) have been a subject of research since the late 1990s (Martha & Santoso, 2019; Sikström et al., 2022). Earlier review studies mainly focused on specific characteristics of these agents. Some explored their visual appearance and role design within virtual learning environments (Kim & Baylor, 2016; Martha & Santoso, 2019), while others examined their communication effectiveness and pedagogical strategies, particularly in K-12 education settings (Sikström et al., 2022; Zhang et al., 2024). These prior works primarily addressed single-agent systems developed before the emergence of LLMs and thus did not reflect the capabilities or dynamics introduced by LLM-enhanced multi-agents.

Researchers and educators need to discern emerging trends and develop corresponding strategies to align MAS innovations with pedagogical priorities by supporting the development of high-order competencies such as critical thinking, creative problem-solving, collaboration, and metacognitive regulation (Miao & Shiohira, 2024). Emerging reviews on MAS focused on diversified application scenarios in education (Jiang, 2024) and the framework for LLM-integrated intelligent multi-agent collaboration in personalized education (Zhao, 2025). However, an overview of MAS distribution in educational technology remains underexplored. To bridge the gap, this study synthesizes the distribution of MAS studies in education over the past decade. It aims to identify whether there is a potential shift in MAS publications in the context of LLM applications in education, examine relevant research across various educational levels, and assess their alignment with contemporary pedagogical objectives. This study seeks to answer the following research questions (RQs):

RQ1. How has the volume of publications on MAS in education evolved over the past decade?

RQ2. How are MAS applications distributed across different educational levels?

RQ3. What are the primary educational goals addressed in MAS research?

## 2. Methodology

In the searching and screening phase, this study employed a systematic literature review method following PRISMA guidelines (Page et al., 2021). A comprehensive search was conducted on December 3, 2025, across four major academic databases (i.e., Web of Science, Scopus, IEEE Xplore, and ACM Digital Library) restricted to publications from 2015 to 2025. Rounds of iteration resulted the search string: *("multi-agent\*" OR "multi agent\*" OR multiagent\* OR "multi-agent system\*" OR "multi agent system\*" OR "multiagent system\*" OR "multi-agent systems" OR "multi agent systems" OR "multiagent systems") AND (education OR student\* OR teacher\* OR teaching OR pedagogy OR pedagogical OR instruction)*. The initial search yielded 7,626 records. After removing 1,423 duplicates, 79 non-English, 459 non-peer-reviewed (e.g., book chapters), 45 non-retrievable, 134 short papers ( $\leq 5$  pages), and 4,963 irrelevant abstracts, the remaining 523 abstract texts. This abstract was conducted to serve as the first screening round and to inform the subsequent full-text analysis.

In the coding and rating phase, to ensure data validity, the selected article abstracts were evenly assigned to two independent researchers for coding using the initial analytical frameworks of educational levels and research objectives. An inductive approach was subsequently adopted to capture publications on MAS in education that lacked a specific target learner group or learning goal. The initial inter-rater agreement was calculated at 73%. All discrepancies were subsequently resolved through a consensus discussion involving a third senior researcher, ultimately achieving agreement on the current classification.

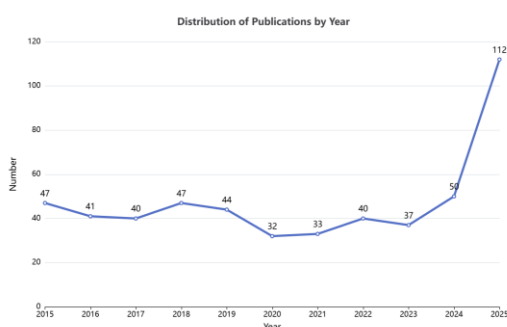
## 3. Key Findings and Discussion

### 3.1. Potential LLMs Surge in MAS Research

Research temporal trends were analyzed by publishing year (see Figure 1), revealing a sharp acceleration in educational MAS research. Publications increased from a stable baseline (2015–2023) to 50 in 2024 and reached 112 in 2025. Related research indicates that authors' motivations were influenced by the integration of LLMs (Malik et al., 2024). The increase in publications observed in 2025, rather than immediately after the release of foundational LLM models in 2023, may be partly explained by a standard academic publication lag in social science disciplines (Björk & Solomon, 2013). This surge is further interpreted through emerging agentic frameworks that increased functionality (Durante et al., 2024), accessible agentic applications that lowered technical barriers (e.g., Coze, Dify), and supportive policy and post-pandemic education digitalization that provided fertile ground (Holmes & Miao, 2023).

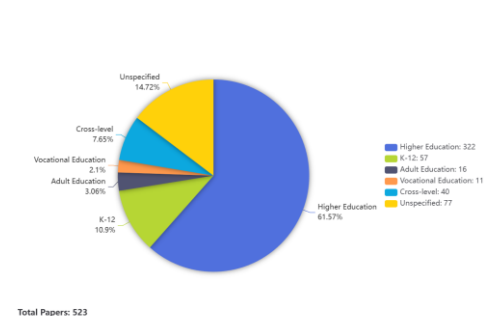
**Figure 1**

*Annual distribution of publications from 2015 to 2025*



**Figure 2**

*Distribution of MAS research by educational level*



### 3.2. The Predominance of Higher Education in MAS Research

Among the 523 reviewed publications, higher education dominates (61.57%), followed by K-12 (10.9%), cross-level (7.65%), adult (3.06%), and vocational education (2.1%), with 14.72% unspecified (see Figure 2). The predominance is likely driven by readiness-optimized infrastructure and institutional capacity suitable for research in higher education (Pedró et al., 2019). Compared to higher education, the smaller K-12 portion might be due to the restrictive, protectionist safeguards typically required in K-12 settings (Akgun & Greenhow, 2022). Adoption in the adult and vocational sectors remains limited, underscoring the adaptable MAS's span across the educational spectrum. A notable portion of studies with unspecified levels is largely simulation-based or system-design research lacking direct human involvement, and these will be further analyzed to determine whether they should be excluded from the review.

### 3.3. Educational Objectives of MAS Research

The educational objectives emerged in seven categories (see Table 1): *Knowledge Mastery* for subject acquisition, such as STEM and language learning; *Cognitive and Metacognitive Skills* for critical thinking and self-regulation; *Ethical and Affective Development* for emotional growth; and *Social and Collaboration* for teamwork. The remaining categories include *Practical Skill Training*, *Learning Process Optimization* for personalization, and *Others* for technical studies that lack explicit human evaluation or pedagogical goals. The analysis reveals that Knowledge Mastery ( $n = 128$ ) and Learning Process Optimization ( $n = 91$ ) are the predominant educational goals, indicating that MAS is utilized primarily for subject-specific content delivery and adaptive tutoring, instead of fostering broader 21st-century competencies like critical thinking ( $n = 25$ ) or collaboration ( $n = 18$ ) (Miao & Shiohira, 2024). The marginal attention to Ethical and Affective Development ( $n = 6$ ) highlights a critical gap in addressing contemporary priorities such as social-emotional learning (OECD, 2021). The high volume of technical studies ( $n = 231$ ) in the current dataset needs to be further analyzed or removed to focus on the educational impact of MAS, rather than its technical architecture. This distribution suggests an imbalance in pedagogical objectives across MAS applications, underscoring more immersive, human-centered pedagogical goals (Zawacki-Richter et al., 2019).

**Table 1**

*Pedagogical landscapes of MAS*

| Pedagogical Goals                  | Number | Pedagogical Goals             | Number |
|------------------------------------|--------|-------------------------------|--------|
| Knowledge Mastery                  | 128    | Practical Skill Training      | 24     |
| Cognitive and Metacognitive Skills | 25     | Learning Process Optimization | 91     |
| Ethical and Affective Development  | 6      | Others                        | 231    |
| Social and Collaboration           | 18     |                               |        |

## 4. Limitation and Future Plan

The limitation in the current study lies mainly in the methodology. First, a preliminary, abstract-level coding may overlook complex pedagogical objectives or the specific technical nuances of MAS architectures, and may therefore not be precise. Second, a Cohen's Kappa coefficient or a percentage agreement of 80% should be established as the threshold for acceptable consistency, requiring a refined coding framework to ensure inter-rater reliability. Third, regarding RQ3, some studies exhibited potential pedagogical overlap and necessitated a multi-label coding approach or theoretical framework for MAS implementation that serves varied instructional goals. Future phases of this review will include an updated criterion and a full-text synthesis to elucidate MAS innovations and pedagogical functions in education, focusing on the evolution enabled by LLMs across system architecture, pedagogical configuration, and empirical impacts on learning outcomes.

## References

- Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and ethics*, 2(3), 431-440.
- Björk, B. C., & Solomon, D. (2013). The publishing delay in scholarly peer-reviewed journals. *Journal of informetrics*, 7(4), 914-923.
- Durante, Z., Huang, Q., Wake, N., Gong, R., Park, J. S., Sarkar, B., ... & Gao, J. (2024). Agent AI: Surveying the horizons of multimodal interaction. *arXiv preprint arXiv:2401.03568*.
- Hamal, O., Faddouli, E., & Harouni, M. H. A. (2021). Design and implementation of the multi-agent system in education. *World journal on educational technology: Current issues*, 13(4), 775-793.
- Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.
- Jiang, Y. H., Liu, T. Y., Zhuang, X., Hu, H., Li, R., & Jia, R. (2024). Enhancing educational practices with multi-agent systems: A review. *Enhancing educational practices: Strategies for assessing and improving learning outcomes*, 47-65.
- Kim, Y., & Baylor, A. L. (2016). Research-based design of pedagogical agent roles: A review, progress, and recommendations. *International journal of artificial intelligence*, 26(1), 160-169.
- Li, X., Wang, S., Zeng, S., Wu, Y., & Yang, Y. (2024). A survey on LLM-based multi-agent systems: Workflow, infrastructure, and challenges. *Vicinagearth*, 1(1), 9.
- Malik, M. A., Amjad, A. I., Aslam, S., & Fakhrou, A. (2024). Global insights: ChatGPT's influence on academic and research writing, creativity, and plagiarism policies. *Frontiers in research metrics and analytics*, 9, 1486832.
- Martha, A. S. D., & Santoso, H. B. (2019). The design and impact of the pedagogical agent: A systematic literature review. *Journal of educators online*, 16(1), n1.
- Miao, F., & Shiohira, K. (2024). *AI competency framework for students*. UNESCO Publishing.
- OECD (2021), *Beyond academic learning: First results from the survey of social and emotional skills*. OECD Publishing.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic reviews*, 10(1), 89.
- Pedró, F., Subosa, M., Rivas, A., & Valverde, P. (2019). *Artificial intelligence in education: Challenges and opportunities for sustainable development*, UNESCO Publishing.
- Sikström, P., Valentini, C., Sivunen, A., & Kärkkäinen, T. (2022). How pedagogical agents communicate with students: A two-phase systematic review. *Computers & Education*, 188, 104564.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators?. *International journal of educational technology in higher education*, 16(1), 1-27.
- Zhang, S., Jaldi, C. D., Schroeder, N. L., López, A. A., Gladstone, J. R., & Heidig, S. (2024). Pedagogical agent design for K-12 education: A systematic review. *Computers & Education*, 223, 105165.
- Zhao, Z. (2025). A new paradigm of personalized education driven by multi-agent collaboration. *Academic journal of sociology and management*, 3(3), 9-17.