

Methodological Framework for Evaluating Multi-Agent Systems (MAS) Tools in Business Training

Radoslav Kostev*

Abstract: This paper examines the options for selecting software tools to develop Multi-Agent Systems (MAS) and the possibilities for training employees to use such tools. The study is based on observations of 30 High-Tech companies across various sectors that develop and use Multi-Agent Systems. This raises the issue of training and preparing personnel for MAS development. The objective is to analyze existing technological products for MAS creation and the associated training opportunities. A set of criteria for platform selection is defined, based on specific business requirements and learner characteristics. A comprehensive methodological framework for evaluating and selecting a platform for employee training is proposed.

Keywords: Multi-Agent Systems, tools, employee training, High-Tech companies.

JEL: I21, M15, O32

I. INTRODUCTION

The rapid development of Artificial Intelligence, Agentic AI and Multi-Agent Systems (MAS) creates new opportunities for modern business. MAS aims to automate business processes, support management decisions and coordinate complex organizational activities. Unlike traditional AI solutions, such as Large language models (LLM) and Generative AI, multi-agent systems are based on the interaction between autonomous or semi-autonomous agents that can perform specialized roles, use different information sources and tools, and coordinate their actions to achieve a common goal.

There are numerous tools and technological products on the market for creating Multi-Agent Systems (MAS), each based on different operational concepts. This presents managers and entrepreneurs with an initial question: which platform is suitable for my business? Answering this leads to a second question: who will develop the Multi-Agent System? This highlights the need for a strategy regarding employee training covering both the use of MAS-building tools and the actual development and management of the systems themselves. To address these issues, a methodological approach is proposed for selecting MAS development tools and training employees, tailored to specific characteristics and needs. The proposed framework aligns the target group's requirements and competencies with the functional and educational attributes of the MAS tools. Thus, the study aims to bridge the gap between the technical assessment of MAS tools and their practical applicability in educational and business contexts.

* Radoslav Kostev is an Associate Professor at the Technical University of Sofia.

II. THEORETICAL FRAMEWORK

Scientific literature contains numerous studies on Multi-Agent Systems and their business applications. In recent years, multi-agent technology has led to the development of a new type of business management system (Zhang, 2024) and the creation of a new intelligent business ecosystem by combining natural business ecosystems with Multi-Agent Systems (Marín, Stalker, Mehandjiev, 2007).

Modern business is increasingly relying on AI as a tool to augment human cognition and automate workflows. The role of AI and Agentic AI systems in business is growing, serving as a means for decision-making (Haj Qasem et al., 2023) and the comprehensive management of resources and processes (Binyamin & Ben Slama, 2022). Agentic AI plays a pivotal role in emerging business management concepts (Anguelov & Angelova, 2026), highlighting the need for skilled personnel capable of developing and operating such information systems.

Selecting a technological environment for developing Multi-Agent Systems and deploying them in a real-world business setting involves several key factors. Raghavendra and Saikia (2026) conclude that system quality depends on its architecture and capabilities such as orchestration, memory management, tool integration, and agent interaction with the primary focus placed on the technical aspects of the tools.

The application of platforms to create Multi-Agent Systems and their application must be tailored to the specifics of the business. In this regard, it cannot be said that there is no single methodology for the selection of tools, and the assessment must be multi-criteria, combining quantitative and qualitative indicators (Bandi et al., 2025; Stübinger, & Metz, 2026).

There are scientific developments that investigate the application and effects of Multi-Agent Systems in the education and training of specialists in various fields. The authors López-Goyez et al. (2026) consider an intelligent learning system based on a Multi-Agent architecture, exploring the coordination of several specialized AI agents can provide adaptive, personalized and pedagogically relevant learning in higher education.

There are significantly fewer developments related to the analysis and evaluation of MAS tools. Wrona et al. (2024) consider the evaluation of such type of platforms through their suitability for real industrial software systems. In this context, the authors propose a system of criteria for comparing similar tools.

The authors Barno and Phelps (2025) propose an approach to using Multi-Agent Systems in teacher education by developing a three-tiered framework. Their idea is that

individual agents can simulate real educational cases, providing immediate feedback or a summary assessment of the learner's performance. In this way, personalization, constant real-time feedback and interaction in learning are aimed (Córdova-Esparza, 2025). This determines the need for properly selected tools for creating MAS, tailored to the specifics of learners. Some developments have focused on creating conceptual frameworks for MAS learning, yet choosing a tool to create MAS remains key to effective learning. The authors Angelov and Angelova (2022, 2022) develop and offer key methods and metrics for evaluating the result and effectiveness of trainings related to information systems and technologies.

III. METHODOLOGICAL FRAMEWORK

This study focuses on the potential for employees in business organizations, including management and executive staff, to develop and utilize artificial intelligence-based Multi-Agent Systems (MAS). Emphasis is placed on analyzing various platforms and tools for building MAS, as well as on their implementation and the associated training requirements. The report presents a methodological framework for selecting tools and training business professionals in the field of Multi-Agent Systems. A methodological sequence is proposed for evaluating and selecting an MAS development platform suitable for training employees with limited or no technical and IT expertise.

The research involves observing 30 high-tech companies across various sectors that develop and use Multi-Agent Systems to manage organizational resources, activities, and processes. The aim is to analyze and evaluate the capabilities of different MAS creation tools and to propose a methodology for selecting a platform to train business professionals in their use.

The following research questions have been defined:

RQ1: Which platforms and tools are most widely used in the development of Multi-Agent Systems within high-tech companies?

RQ2: What are the key criteria for selecting a platform to train employees in the creation of Multi-Agent Systems?

Considering the defined objective and the research questions, the following research tasks are formulated:

1. To investigate and analyze various platforms and tools for developing Multi-Agent Systems.
2. To derive a set of criteria for evaluating technological products used to develop Multi-Agent Systems.
3. To develop a general methodological framework for selecting a platform suitable for training employees to build Multi-Agent Systems.

IV. RESULTS

4.1. Analysis and evaluation of platforms and tools for building Multi-Agent Systems.

The study covers companies across various sectors—such as IT (software and telecommunications), finance, insurance, banking, retail, healthcare, logistics, manufacturing, tourism, and startups that utilize Multi-

Agent Systems (MAS). Consequently, MAS finds broad application in companies with diverse organizational structures and process workflows. Furthermore, each company possesses a unique human resource profile, meaning employees vary in their levels of technical literacy, professional skills, and competencies. Therefore, the tools used to build Multi-Agent Systems must be tailored to the organization's specific characteristics and, above all, the profiles of the personnel who will use them.

In practice, a wide range of technological products and tools designed for professionals with diverse backgrounds are utilized. Some platforms cater to professionals with a technical background, requiring technical expertise such as coding or programming. Examples include LangChain, OpenAI Agents SDK, Microsoft Agent Framework, CrewAI, Google ADK, and Pydantic AI. These tools fall into the category of "Pro-code" or "Full-code" platforms. Such platforms serve as programming environments for developing the agent architecture of Multi-Agent Systems, offering a significantly higher degree of control over agent behavior. They are characterized by more complex integration and interactions, high flexibility, and superior integration capabilities with other systems.

Other tools are geared more towards business users and do not require extensive technical knowledge; they feature visual interfaces and/or allow for instructions to be given in natural language. Examples of such platforms include Microsoft Copilot Studio, n8n, Make, Zapier, Dify, and others. These fall under the category of no-code platforms. No-code and low-code platforms are software environments that enable the creation, configuration, and orchestration of multiple AI agents via a graphical interface. This allows users without significant technical or programming skills to build complex agent-based systems, enabling the visual modeling of activities and interactions between various processes and agents.

There are also tools that combine both approaches such as Langflow and Flowise to a certain extent. These can be classified as low-code platforms, as this approach offers greater flexibility. Some of the tools mentioned earlier may also fall into this category.

Based on the analysis of these various tools, the following classification can be made:

Table 1: MAS development tools category.

Category	Features	Platforms
No-Code	Visually design agents and workflows without programming	Microsoft Copilot Studio, Zapier AI Agents, Make, n8n
Low-Code	Visual build with the ability to add scripts and custom integrations	Flowise AI, Dify, n8n, CrewAI, Stack AI
Pro-Code	Development using programming languages and APIs	Microsoft AutoGen, LangChain, OpenAI Platform (Agents SDK), CrewAI, Google ADK, Pydantic AI

Since the study encompasses employees with diverse professional skill sets, training and work involving such platforms must be tailored accordingly. In this context, the study focuses on no-code and low-code platforms. Observations indicate that among high-tech companies, the preferred no-code and low-code tools for developing and managing multi-agent systems are n8n, Make, and Zapier.

n8n is an AI-powered low-code workflow automation platform that combines visual modeling with coding capabilities. It is available as a cloud service but can also be installed locally. It offers integration with other applications, information systems, and various language models (such as OpenAI, Anthropic, and others). Users describe the platform as intuitive, easy to use, and versatile, capable of handling core workflows. Compared to the other two platforms, it is characterized by greater technical complexity and requires IT proficiency, which may necessitate a longer training period (n8n, 2026).

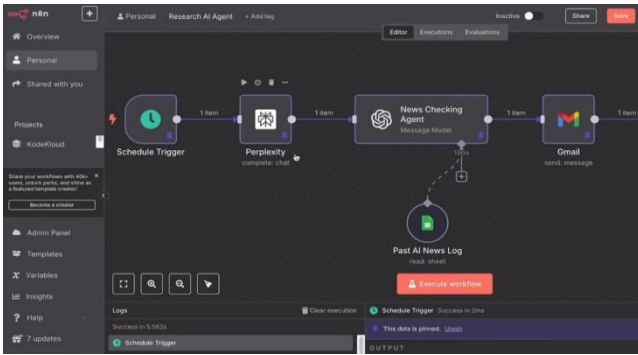


Figure 1: Screenshot of the n8n platform interface.

Like n8n, Make is a low-code process automation tool, but unlike n8n, it is available exclusively as a cloud-based service. It offers extensive integration capabilities with other platforms and tools. Users find the platform highly intuitive, noting its user-friendly interface and built-in features for processing text and datasets. Make strikes a good balance between no-code and low-code, making it well-suited for beginners with no technical background. This makes the tool highly appropriate for training purposes (Make, 2026).

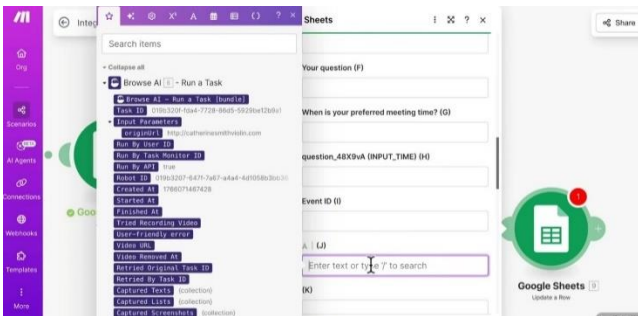


Figure 2: Screenshot of the Make platform interface.

Zapier positions itself as infrastructure for AI automation, a tool offered as a no-code, low-code, and pro-code platform that supports a library of over 9000 applications and enables the creation of custom apps. Users describe the platform as intuitive and easy to use,

highlighting its extensive integration capabilities with various tools, including LLMs and business information systems such as ERP and CRM. Key drawbacks cited include limited control over complex multi-agent architectures and rising costs associated with high automation volumes, as pricing is based on the number of automations (Zapier, 2026).

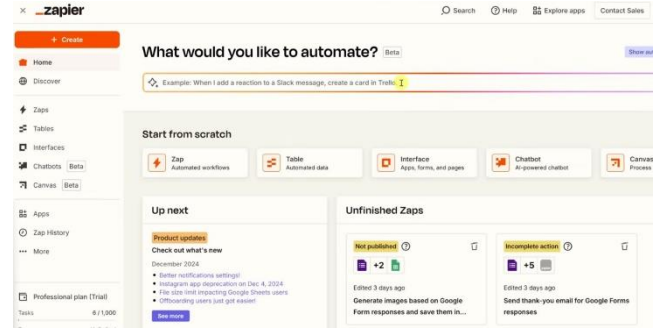


Figure 3: Screenshot of the Zapier platform interface.

An examination of the three Multi-Agent System (MAS) tools leads to the conclusion that all three are suitable solutions for training business professionals with limited IT and technical skills.

Training employees with diverse professional backgrounds in the field of Multi-Agent Systems presents specific challenges. This necessitates a careful selection of the training platform and consideration of future application in a real-world work environment. Consequently, it is necessary to establish a set of criteria for evaluating and selecting MAS tools; these criteria are presented in Table 2.

Table 2: Evaluation criteria for MAS development tool.

No	Criteria	Measures:
C1	No-Code Accessibility	Ability to work with a graphical interface, without coding or programming.
C2	Business Process Modeling	Business process modeling capability of the platform.
C3	Ease of use	Intuitive and easy for users to understand and operate.
C4	Integration capabilities	Capability for integration with various platforms, applications, and information systems.
C5	AI integration	Level of AI development and/or integration with LLMs.
C6	Multi-Agent capabilities	Capabilities for creating and coordinating multiple specialized agents.
C7	Technical extensibility and scalability	System expandability and scalability.
C8	Price	Price for service provision and technical support.
C9	Security	Level of security and reliability of the platform.

The first criterion concerns the platform's operational modes, specifically, whether it relies on a graphical interface or coding. In our case, the choice of platform considers the profile of employees who lack technical or programming expertise. This criterion serves as the starting point for selecting a suitable tool for both training and daily operations.

The second criterion is crucial for both no-code and pro-code platforms, as it determines the platform's ability to support the organization's business processes and activities. It dictates how effectively the resulting MAS (Multi-Agent System) will execute the required business logic.

The third criterion relates to the system's intuitiveness and ease of use. This is vital for training employees who are encountering this type of system for the first time.

The fourth criterion assesses the platform's integration capabilities with other internal or external information systems, such as email, ERP, CRM, websites, mobile apps, and many others.

The fifth criterion is linked to the third and concerns the integration of artificial intelligence, either directly within the platform or agent itself, or through integration with external AI models.

The sixth criterion relates to the platform's product specifically, the capabilities for creating, orchestrating, and managing agents within the system, as well as facilitating collaboration among them.

The seventh criterion concerns system scalability and the ability to handle increasing workloads while maintaining required performance levels, accommodating the growing number and expanding functionalities of the agents within the system. Technical support can also be included under this criterion.

The eighth criterion is cost, which is a significant factor, particularly for smaller companies and startups. This includes the cost of additional services associated with the overall use of the platform.

The final but most important criterion evaluates the platform's security and its ability to safeguard sensitive data, thereby ensuring the security of the agents within the system.

Collectively, these criteria enable a comprehensive assessment of tools used to train employees in building Multi-Agent Systems. Selecting a platform based on these criteria can be achieved by calculating a platform suitability index. This assessment can be performed using the following formula:

$$PSI_j = \sum_{w=1}^n w_i s_{ji}$$

where:

PSI_j – platform suitability index for platform j ;

s_{ji} – evaluation of platform j according to criterion i ;

w_i – weight of criterion i .

Each of the analyzed platforms is evaluated against every criterion using a Likert-type scale (e.g., from 1 to 5). Each criterion is assigned a weight coefficient. The weight

of each criterion is determined based on the organization's specific characteristics (processes, personnel, technical infrastructure, etc.). The sum of the weights for all criteria equals 1. Specific methodologies and models for evaluating training outcomes and effectiveness have been developed by Angelov and Angelova (2022).

Following the analysis and evaluation of a suitable tool, the process can proceed to the final stages: conducting the training using the selected platform and evaluating the results.

4.2. Methodological framework for selecting a tool for creating Multi-Agent Systems (MAS) for employee training.

The process of selecting a suitable employee training tool follows this methodological sequence (Figure 4):

I. Training Needs Analysis – This is the initial stage, aimed at creating a profile of the employees to be trained. The objective is to analyze the trainees, their characteristics, and the competencies they need to acquire. This involves, first, analyzing the trainees' skills, competencies, and qualities, as well as identifying knowledge and skill gaps. Next, a training plan is developed, and measurable training outcomes are formulated.

II. Platform Requirements Definition – This stage aims to translate training needs into specific requirements for the software tool. Requirements are defined based on business specifics and the characteristics of the trainees/employees, alongside technical requirements for the platform. Platform selection criteria are established, and the relative importance (weighting factor) of each criterion is determined through expert assessment. A preliminary analysis of software solutions available on the market is also conducted at this stage.

III. Platform Analysis and Evaluation. This stage involves the actual process of selecting a technological tool for training. A detailed analysis of the platforms is conducted, covering features, capabilities, design, and more. Subsequently, each platform is evaluated based on established criteria and pre-defined weighting factors. The resulting expert assessment provides the basis for selecting a specific platform tailored to the needs and specific characteristics of the business and the trainees.

IV. Training Design and Implementation. Once a specific tool has been evaluated and selected, the actual training phase begins. At this stage, the training curriculum is developed, considering the specifics of the business, the employees, and the platform itself. Training is conducted on using the platform, solving a specific business case, and building a functional AI agent. The objective here is for employees (trainees) to acquire competencies in areas such as MAS platforms, Agentic AI, Multi-Agent Systems, and business processes.

V. Evaluation of Results. The final stage concludes with an evaluation of the training outcomes. This assessment may be carried out by experts based on specific criteria, such as business relevance, technical accuracy, feasibility of real-world deployment, integration capabilities, and others. This evaluation enables the refinement of the training curriculum to align with the evolving needs and specific characteristics of the business and its employees.

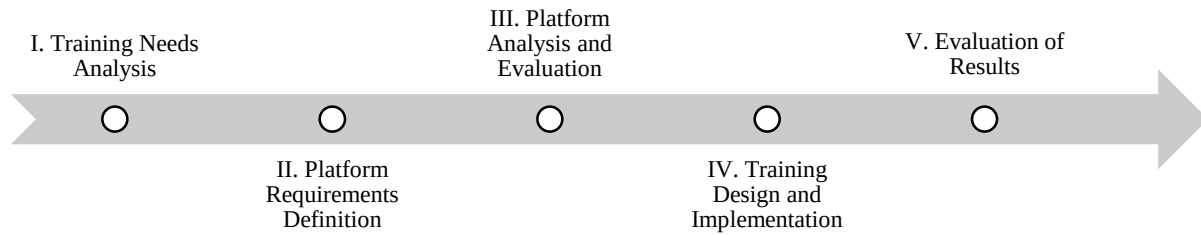


Figure 4: Methodological framework for selecting a tool for creating Multi-Agent Systems for employee training.

CONCLUSION

The implementation of Agentic AI and Multi-Agent Systems lies at the core of the future development of intelligent enterprises. Analyses reveal the capabilities of existing tools for developing Multi-Agent Systems, as well as their potential applications in business operations and employee training. The proposed methodological framework enables a comprehensive assessment for selecting a platform to create Multi-Agent Systems intended for training employees and managers. Naturally, any training initiative and tool selection depend on specific business requirements and learner characteristics. Consequently, the framework can be adapted and further refined based on the specific context of implementation, thereby opening avenues for future research in this field.

ACKNOWLEDGEMENTS

This research was funded by the Bulgarian National Science Fund under Grant No. KII-06-M95/5.

REFERENCES

- Anguelov, Kiril, Angelova, Miglena (2022) Indicators for evaluation the effectiveness of trainings through educational computer games. AIP Conf. Proc. 2505 (1): 040008. <https://doi.org/10.1063/5.0103044>.
- Anguelov, Kiril, Angelova, Miglena (2022) Methodology for evaluation the effectiveness of corporate distance training. AIP Conf. Proc. 2505 (1): 040009. <https://doi.org/10.1063/5.0103048>.
- Anguelov, Kiril, Angelova, Miglena (2026). Algorithmic Management as a Driver of Accelerated Growth in High-Growth Firms. Secure Systems Design and Technology Development. Proceedings of the International Scientific Conference. Veliko Tarnovo, Bulgaria. Volume I, 31-36.
- Bandi, Ajay, Kongari, Bhavani, Naguru, Roshini, Pasnoor, Sahitya, & Vilipala, Sri Vidya (2025). The Rise of Agentic AI: A Review of Definitions, Frameworks, Architectures, Applications, Evaluation Metrics, and Challenges. Future Internet, 17(9), 404. <https://doi.org/10.3390/fi17090404>.
- Barno, Erin, & Phelps, Geoffrey (2025). Using a Multi-Agent System and Evidence-Centered Design to Integrate Educator Expertise Within Generated Feedback. Education Sciences, 15(10), 1273. <https://doi.org/10.3390/educsci15101273>

Binyamin, Sami Saeed, & Ben Slama, Sami (2022). Multi-Agent Systems for Resource Allocation and Scheduling in a Smart Grid. Sensors, 22(21), 8099. <https://doi.org/10.3390/s22218099>

Córdova-Esparza, Diana-Margarita (2025). AI-Powered Educational Agents: Opportunities, Innovations, and Ethical Challenges. Information, 16(6), 469. <https://doi.org/10.3390/info16060469>

Haj Qasem, Mais, Aljaidi, Mohammad, Samara, Ghassan, Alazaidah, Raed, Alsarhan, Ayoub, & Alshammari, Mohammed (2023). An Intelligent Decision Support System Based on Multi Agent Systems for Business Classification Problem. Sustainability, 15(14), 10977. <https://doi.org/10.3390/su151410977>.

López-Goyez, Juan P., González-Briones, Alfonso, Demazeau, Yves, & Guaytarilla G., Jhonatan M. (2026). An Exploratory Study of a Generative AI-Based Intelligent Tutoring System Using a Multi-Agent Architecture in Higher Education. Applied Sciences, 16(11), 5453. <https://doi.org/10.3390/app16115453>.

Marín, César A., Stalker, Iain, Mehandjiev, Nikolay (2007). Business Ecosystem Modelling: Combining Natural Ecosystems and Multi-Agent Systems. In: Klusch, M., Hindriks, K.V., Papazoglou, M.P., Sterling, L. (eds) Cooperative Information Agents XI. CIA 2007. Lecture Notes in Computer Science(), vol 4676. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-75119-9_13.

Raghavendra, Priyadarshini & Saikia, Manob Jyoti (2026). Agentic AI: A Perspective on Architecture, Frameworks and Applications. AI, 7(6), 219. <https://doi.org/10.3390/ai7060219>.

Stübinger, Johannes, & Metz, Fabio (2026). Understanding AI Agents—A Data-Driven Literature Review. Mathematics, 14(9), 1478. <https://doi.org/10.3390/math14091478>

Zhang, Hongbo (2024). Construction of a New Type of Business Management System in the Supply Chain Using Multi-agent Technology. In: Pei, Y., Ma, H.S., Chan, YW., Jeong, HY. (eds) Proceedings of Innovative Computing 2024, Vol. 2. IC 2024. Lecture Notes in Electrical Engineering, vol 1215. Springer, Singapore. https://doi.org/10.1007/978-981-97-4125-0_3.

Websites:

- Make. What is Make?. Official Website. Available at: <https://www.make.com/en>.
- n8n. Official Website. Available at: <https://n8n.io/>.
- Zapier. Official Website. Available at: <https://zapier.com/>.