

AI Agents: Architectures, Frameworks, Challenges and Future Directions – A Review

Mahak Garg, Kanika Garg, Neeta Verma

Department of Computer Science and Engineering - Artificial Intelligence
Inderprastha Engineering College, Ghaziabad, India

© The Author(s), under exclusive license to publication division, IPEC Journal of Science & Technology, 2026

Abstract : *The advent of AI agents is considered to be among the most important developments within the realm of Generative AI where agents are capable of independently perceiving their environments, performing tasks, reasoning, making decisions and taking actions in response to the environmental inputs with little to no human interaction. In contrast to conventional AI systems that mainly involve responses to user's queries, AI agents involve LLMs along with memory, planning, reasoning and other external tools to accomplish multi-step tasks. This review aims to provide a succinct summary on AI agents architectures, popular frameworks for developing them, application areas and research challenges. New frameworks such as LangGraph, CrewAI, AutoGen and OpenAI Agents SDK have facilitated rapid development of intelligent and autonomous systems for healthcare, educational, financial, cybersecurity and software engineering applications. This paper also sheds light on some drawbacks of AI agents such as hallucination, security issues, privacy concerns and computationally heavy nature of these systems. Finally, future directions like multi-agent collaboration, trustworthy AI agents and efficient device-level deployment are mentioned.*

Keywords : *AI Agents, Large Language Models, Agentic AI, LangGraph, CrewAI, AutoGen, Autonomous Systems, Generative AI*

I. INTRODUCTION

The fast progress of LLMs turned artificial intelligence into AI agents which are able to reason, plan, manage their memories, use different tools, etc[1,2]. An AI agent is an intelligent software which accomplishes certain predetermined goals through communication with users, applications, databases and digital environment. Contrary to chatbots which provide simple replies to users queries, AI agents have the ability to break complex tasks into simpler sub-tasks, choose appropriate tools, perform actions and then analyze results and adjust their strategies accordingly [1,11].

Due to the latest advancements in Generative AI, a few agent platforms have appeared recently, such as LangGraph, CrewAI, AutoGen and OpenAI Agents SDK. These platforms simplify the development process of autonomous applications which utilize not only LLMs but also memory and planning modules as well as external APIs[3,9,10]. Nowadays, AI agents find their application in different spheres such as healthcare, education, finance, software engineering, cybersecurity, robotics, business automation etc. Despite their outstanding abilities, these intelligent programs experience some problems concerning reliability, hallucination, security, explainability and ethical governance [2,6,12].

In this review we will try to outline the state-of-the-art of AI agents in terms of architecture, frameworks, applications, research challenges and future prospects [4,5,15].

II. LITERATURE REVIEW

The progress achieved by recent advances in LLMs has enabled researchers to develop autonomous AI agents which can not only reason, plan, manage memory but also use tools to accomplish certain tasks. As compared to traditional conversational AI, modern AI agents are able to complete multi-step processes by using third-party services, APIs and other information sources [1,2,11].

In recent years, various papers have introduced comprehensive architectures of AI agents. According to a survey conducted by Masterman et al. (2024), there is a growing interest in developing new AI agent architectures, which have reasoning, planning, reflection and tooling capabilities. The study made a comparison between single-agent and multi-agent systems and found out that planning and collaboration significantly improved performance results [1].

A systematic review of AI agents was provided by Bharti (2025). It describes the history of AI agent development from reactive agents into intelligent autonomous agents. The paper also stresses the role of large language models, reinforcement learning, memory modules and knowledge graphs in the process of autonomous decision-making [9]. Surveys have also been conducted recently regarding Agentic AI, a framework whereby autonomous agents are capable of perceiving, reasoning, planning and taking action with little human involvement. In an extensive review using the PRISMA framework in the journal Artificial Intelligence Review, dozens of studies have been classified into symbolic/classical and neural/generative frameworks. The

increasing significance of hybrid neuro-symbolic frameworks was emphasized by the researchers [5,6].

A framework is an essential tool in designing AI agents. Derouiche et al. (2025) examined popular frameworks like LangGraph, CrewAI, AutoGen, Semantic Kernel and Google

ADK. This paper reviewed the design architecture, communication protocols, memory management, safety and interoperability of the mentioned frameworks. It was found out that though the framework helps in creating AI agents, there are challenges that are yet unsolved [3,10].

Applications of AI agents have grown very fast in many areas. In the field of healthcare, AI agents help doctors to analyze data and aid in clinical decision making. In software engineering, code creation, debugging and testing can be done through the use of AI agents. In the case of cybersecurity, intelligent agents ensure constant monitoring of the network, detection of cyberattacks and their defense [2,4,8].

Nevertheless some issues still persist. Some existing research papers suggest that the main problems which hinder the development and application of AI agents include such issues as hallucination, privacy, lack of explainability, expensive computations and security concerns. The key points for future research include the development of reliable AI agents with better reasoning, efficient memory techniques, explanation of decisions and safety when interacting with external tools [6,12,13,15].

Table 1. Summary of Recent Literature

Author(s)	Research Focus	Key Contribution	Limitation
Masterman et al., 2024	AI Agent Architecture	Reviewed reasoning, planning, memory, and tool use.	Limited real-world validation.
Bharti, 2025	AI Agent Evolution	Discussed evolution from rule-based to LLM-based agents.	Focused mainly on architecture.
Abou Ali et al., 2025	Agentic AI	Proposed a taxonomy of Agentic AI architectures.	Hybrid models require further validation.
Derouiche et al., 2025	AI Agent Frameworks	Compared LangGraph, CrewAI, AutoGen, and Semantic Kernel.	Limited industrial evaluation.
Xu et al., 2025	Tool-Learning Agents	Reviewed LLM-based tool learning and integration.	Focused only on tool-learning tasks.

Yehudai et al., 2025	Agent Evaluation	Surveyed evaluation methods and benchmarks for LLM agents.	Standard evaluation metrics are limited.
Ferrag et al., 2025	Autonomous AI Agents	Reviewed reasoning, communication, and agent frameworks.	Rapid evolution of frameworks.
Xie et al., 2025	Multimodal Agents	Surveyed AI agents with text, image, and vision capabilities.	Limited to multimodal systems.
Fang et al., 2025	Self-Evolving Agents	Introduced lifelong learning AI agents.	Few real-world implementations.
Nguyen et al., 2025	GUI Agents	Reviewed GUI-based autonomous agents and applications.	Mainly applicable to GUI environments.

Table 1 provides an overview of the key advances made in AI agents between 2024 and 2025. The papers examined mainly emphasize the architectural design, development methodologies, evaluation techniques, tool integrations and novel applications of AI agents. Some of the recent papers have delved into advanced issues in AI agents such as multimodal agents, self-evolving agents and graphical user interface (GUI)-based autonomous agents. However some of the challenges, including security, scalability, explainability and standardization of evaluations, are still open research issues [12,13,14,15].

III. AI Agent Architectures and Frameworks

These are autonomous software programs that employ large language models (LLMs) together with planning, memory, reasoning and other external services to execute complicated missions. In contrast to traditional chatbots an AI agent can comprehend users goals, plan for its implementation, interface with other APIs and databases, judge interim results and adjust its actions until the goal is reached. An AI agent usually includes five main modules: perception, reasoning, planning, memory and action execution [1,2,4,11].

i. AI Agent Architecture

The architecture of the AI Agent is based on several modules working in tandem to help it make intelligent decisions [1,2].

- **Perception:** Takes inputs from users, sensors, and outside entities.

- **Reasoning:** Makes use of LLM to comprehend the nature of the problem and make logical decisions.
- **Planning:** Decomposes tasks into smaller subtasks.
- **Memory:** Keeps records of past experiences and contexts to facilitate future reasoning.
- **Action Module:** Carries out the actions through interactions with APIs, databases, search engines, or applications [1,3,4].

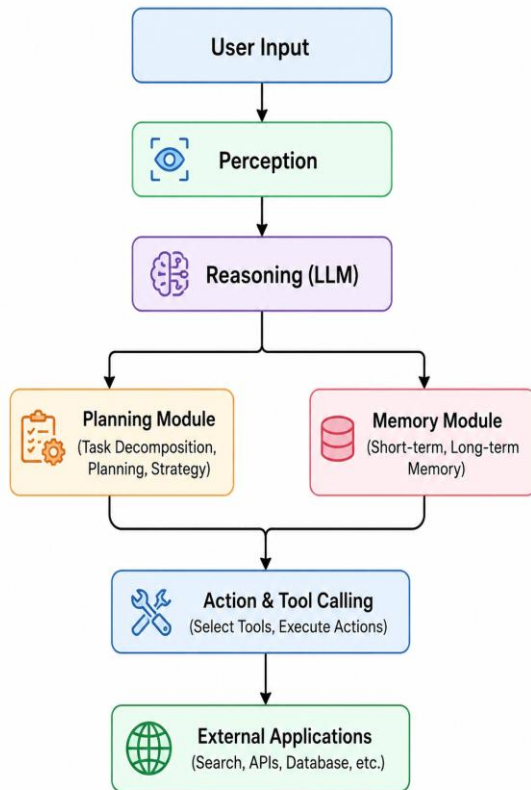


Figure 1. General Architecture of an AI Agent

An illustration of the basic structure of an AI agent is provided in Figure 1. First, the process involves receiving input from the user which is then processed using the perception and reasoning components. Planning and memory components help in making decisions through formulating task strategies and keeping relevant context. Lastly the action and tool-calling component communicates with external systems like APIs and databases to perform the required task [1,3,11].

ii. Popular AI Agent Frameworks

Different frameworks have been proposed that make the implementation of AI agents easy.

Table 2. Comparison of Popular AI Agent Frameworks

Framework	Developer	Key Features	Applications
LangGraph	LangChain	Graph-based workflow orchestration, memory support	Workflow automation
CrewAI	CrewAI	Multi-agent collaboration with role assignment	Business automation
AutoGen	Microsoft	Autonomous communication among multiple agents	Software development, research
OpenAI Agents SDK	OpenAI	Tool calling, memory, safety mechanisms	AI assistants
LlamaIndex	LlamaIndex	Retrieval-Augmented Generation (RAG) and data integration	Enterprise knowledge systems

Every framework has different abilities for workflow, memory management, collaboration between agents and integration with tools [3,9,10].

Out of all the frameworks available, LangGraph is well-suited for the design of complex workflows and CrewAI can be used in the case of collaborative multi-agent setups. The AutoGen framework allows independent conversations between different AI agents and can be used for software development and research purposes.

IV. CHALLENGES AND FUTURE DIRECTIONS

The current rapid developments in AI agents have changed the nature of interaction between humans and intelligent systems and the ways of completing complicated actions. Nevertheless there are several limitations in terms of technology and ethics that must be solved before achieving large-scale implementation of AI agents. The first one is the risk of hallucinations, when agents may give incorrect information. Moreover, there are certain security threats like prompt injection attacks, unauthorized access to tools and possible breaches of confidentiality. Another challenge consists of high costs of computations, lack of explainability of the processes and the lack of evaluation framework.

As for the future, the prospects of development of AI agents are extremely positive. First of all, currently researchers are trying to create AI agents that will provide transparency and trust. The following trends can be considered promising ones: collaboration among agents, lightweight language model, edge AI and collaboration between human and machine. Additionally continual learning, reinforcement learning and safety mechanisms will make AI agents flexible and efficient enough. With time, AI agents will become an important part of the development of the next generation of intelligent systems in healthcare, education, finance, cybersecurity and many other fields.

V. CONCLUSION

AI agents mark a considerable improvement in artificial intelligence through the use of Large Language Models in conjunction with planning, reasoning, memory and tool utilization for accomplishing tasks autonomously. In this review, architectures of AI agents, their frameworks and existing challenges faced in relation to AI agents were presented. The development of recent frameworks like LangGraph, CrewAI, AutoGen and the Agents SDK developed by OpenAI has increased the usage of AI agents in the fields of healthcare, education, finance, cybersecurity and software engineering. There is still much more that needs to be done to ensure secure, scalable and efficient AI agents can operate in real-world environments. In summary, the use of AI agents is anticipated to be of great importance for future intelligent automation and decision support systems. More efforts for research in reliability, safety and efficiency will lead to further advancement in their adoption.

REFERENCES

- [1] L. Wang, C. Ma, X. Feng, *et al.*, “A Survey on Large Language Model Based Autonomous Agents,” *Frontiers of Computer Science*, vol. 18, no. 6, 2024.
- [2] M. A. Ferrag, N. Tihanyi, and M. Debbah, “From LLM Reasoning to Autonomous AI Agents: A Comprehensive Review,” *arXiv preprint arXiv:2504.19678*, 2025.
- [3] H. Derouiche, Z. Brahmi, and H. Mazeni, “Agentic AI Frameworks: Architectures, Protocols, and Design Challenges,” *arXiv preprint arXiv:2508.10146*, 2025.
- [4] T. Liang, Z. He, W. Jiao, *et al.*, “Understanding Agentic AI: Algorithms and Infrastructure,” *IEEE/CAA Journal of Automatica Sinica*, 2026.
- [5] “The Role of Agentic AI in Shaping a Smart Future: A Systematic Review,” *ScienceDirect*, 2025.
- [6] “Agentic AI Systems: A Systematic Survey of Multi-Agent Architectures, Cognitive Foundations, Interaction, Explainability, Security, and Performance Evaluation,” *Neurocomputing*, 2026.
- [7] L. Stauffer, K. Feng, K. Wei, *et al.*, “The 2025 AI Agent Index: Documenting Technical and Safety Features of Deployed Agentic AI Systems,” *arXiv preprint arXiv:2602.17753*, 2026.
- [8] H. Wang, J. Gong, H. Zhang, and Z. Wang, “AI Agentic Programming: A Survey of Techniques, Challenges, and Opportunities,” *arXiv preprint arXiv:2508.11126*, 2025.
- [9] S. Joshi, “Review of Autonomous Systems and Collaborative AI Agent Frameworks,” *International Journal of Science and Research Archive*, vol. 14, no. 2, pp. 961–972, 2025.
- [10] V. Patel, “Systematic Comparison of Agentic AI Frameworks for Scholarly Literature Processing,” *International Journal of Scientific Research and Technology*, vol. 2, no. 9, pp. 169–183, 2025.
- [11] X. Li, “A Review of Prominent Paradigms for LLM-Based Agents: Tool Use, Planning (Including RAG), and Feedback Learning,” *Proceedings of the 31st International Conference on Computational Linguistics (COLING)*, 2025.
- [12] A. Yehudai, L. Eden, A. Li, *et al.*, “Survey on Evaluation of LLM-Based Agents,” *arXiv preprint arXiv:2503.16416*, 2025.
- [13] M. Mohammadi, Y. Li, J. Lo, and W. Yip, “Evaluation and Benchmarking of LLM Agents: A Survey,” *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2025.
- [14] J. Xie, Z. Chen, R. Zhang, *et al.*, “Large Multimodal Agents: A Survey,” *Visual Intelligence*, vol. 3, 2025.
- [15] J. Fang, Y. Peng, X. Zhang, *et al.*, “A Comprehensive Survey of Self-Evolving AI Agents: A New Paradigm Bridging Foundation Models and Lifelong Agentic Systems,” *arXiv preprint arXiv:2508.07407*, 2025.