IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, October 2026
Impact Factor: 8.2

A Systematic Literature Review of Agentic AI: Definitions, Architectures, and Challenges

Sandhya Pravin Naral[1], Prof. K. M. Borude[2], Prof. L. S. Bhadake[3], Dr. H. B. Jadhav[4]
[1,2,3,4]Department of Computer Engineering,
Adsul's Technical Campus, Chas, India

Abstract: *Agentic Artificial Intelligence (Agentic AI) refers to systems that can plan, make decisions, use tools, and execute multi-step actions toward goals, with varying levels of human supervision. This condensed review summarizes research on Agentic AI published between 2021 and 2025. The reviewed literature is organized around five areas: memory and cognition, networking systems, trust and safety, evaluation limits, and applications. The source review reports that practical applications and autonomy receive substantial attention, while adaptability and new architectural models remain underexplored. Major challenges include architectural complexity, unreliable behavior in unfamiliar settings, limited evaluation methods, computational cost, and questions of accountability. Future research should prioritize resilient architectures, measurable recovery from errors, adaptive planning, and stronger safety controls.*

Keywords: Agentic AI, artificial intelligence, systematic literature review, autonomous agents, architectures, adaptability, safety

I. INTRODUCTION

Artificial Intelligence has progressed from systems that mainly respond to predefined inputs toward systems that can pursue goals through planning, feedback, memory, and tool use. Agentic AI describes this shift from passive assistance to goal-directed action. Modern agentic systems may combine large language models (LLMs), external tools, memory, planning modules, and feedback loops to complete tasks that require several steps.

However, the term Agentic AI does not have one universally accepted definition. Some studies treat it as a form of autonomous software agent or multi-agent system, while others describe agency as behavior emerging when an LLM is connected to tools and iterative planning. This review focuses on the common features, architectures, applications, limitations, and research gaps identified in the provided literature.

II. SEARCH METHODOLOGY

The source review follows a systematic literature review approach informed by Kitchenham's methodology and the PRISMA 2020 reporting principles. It searched major computing and engineering research databases for studies published from 2021 to mid-2025. Search terms included "agentic AI," "AI agent," "autonomous agent," "LLM," and "large language model."

The reported selection process began with 125 studies. After removing unsuitable and duplicate records and screening titles, abstracts, and full texts, 48 studies remained for qualitative assessment. The review considered research purpose, context, related work, methodology, results, conclusions, and future directions. These figures are reported by the source paper and are not a new independent search.

III. THEORETICAL FOUNDATIONS

Agency and autonomy are related but distinct. Autonomy refers to operating without continuous external control; agency additionally emphasizes goal-directed behavior, decision-making, and the ability to influence outcomes. A system can act automatically without necessarily planning or adapting toward broader objectives.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, October 2026
Impact Factor: 8.2

Traditional rule-based systems operate within predefined rules and knowledge. Belief–Desire–Intention (BDI) architectures model agents through beliefs about the environment, desired outcomes, and intentions or plans. Reinforcement learning and multi-agent systems expanded the range of adaptive behavior. More recently, LLM-based agents have combined language-based reasoning with tools, memory, and iterative action. Yet, tool use alone does not guarantee robust agency or reliable decision-making.

IV. ARCHITECTURES AND APPLICATION AREAS

Agentic systems commonly integrate several components: input or perception, memory, planning and reasoning, learning, tool or action interfaces, and feedback. These components must coordinate as a coherent system. Modular designs can support flexibility, but interactions between modules may also create unexpected failures. LLM-based agents additionally face challenges in context tracking, prompt management, tool errors, and verification of intermediate steps.

The literature reviewed in the source paper covers applications across digital services, software development, education, agriculture, healthcare, networking, and other domains. Agents can coordinate workflows, retrieve information, operate tools, assist users, or collaborate with other agents. Multi-agent systems may distribute tasks among specialized agents and coordinate their outputs. The usefulness of these applications depends on the task, the quality of the surrounding architecture, and the safeguards applied.

V. KEY CHALLENGES

Architectural complexity: perception, memory, planning, learning, and execution must work together. Small errors in one component can affect the entire workflow.

Evaluation limits: conventional accuracy or reward metrics do not fully measure long-term goal pursuit, proactivity, adaptation, or recovery from failure. There is no single standard benchmark that captures all aspects of agency.

Generalization and robustness: an agent that performs well in a controlled task may behave unpredictably when inputs, tools, or environmental conditions change.

Trust, safety, and accountability: autonomous actions raise questions about oversight, explainability, privacy, and responsibility. High-impact actions should have clear permission boundaries, logging, and human review.

Computational demands: repeated model calls, memory retrieval, tool execution, and multi-agent coordination can increase latency and operating cost.

VI. FINDINGS AND DISCUSSION

The source review identifies a research landscape strongly oriented toward practical applications and the implementation of autonomy. Its intersection analysis highlights a relative lack of work on adaptability and new models. In other words, many studies demonstrate that agents can complete tasks, but fewer establish how agents can remain reliable when conditions change or unexpected failures occur.

This imbalance matters because real-world environments are dynamic. Systems optimized only for a predefined workflow may be brittle outside expected conditions. A stronger research direction is to evaluate not only whether an agent completes a task, but also whether it detects mistakes, recovers safely, revises plans, and explains or records consequential actions.

VII. FUTURE SCOPE

Future work should develop architectures that support self-reflection and real-time plan revision while maintaining clear limits on autonomous action. Evaluation should include resilience measures, such as recovery from injected errors, performance under changing conditions, and safe escalation to a human. Researchers should also explore adaptive learning after deployment, reproducible benchmarks, observability, and methods that reduce unnecessary tool calls and computational cost.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, October 2026
Impact Factor: 8.2

VIII. CONCLUSION

This review summarizes the development, architectures, applications, and challenges of Agentic AI. The field has progressed toward systems that use planning, memory, tools, and feedback to carry out multi-step goals. The reviewed research demonstrates growing practical interest, but also reveals an important gap: adaptability and resilient behavior remain less developed than application-focused autonomy. Reliable agentic systems will require stronger architectures, better evaluation methods, clear safety boundaries, and continuous monitoring. The central next step is to move beyond demonstrating task completion and build agents that can respond safely and effectively to unfamiliar conditions.

REFERENCES

[1] D. B. Acharya, K. Kuppan, and B. Divya, "Agentic AI: Autonomous intelligence for complex goals—A comprehensive survey," IEEE Access, vol. 13, pp. 18912–18936, 2025.
[2] B. Kitchenham, "Procedures for Performing Systematic Reviews," Keele University, Technical Report, 2004, pp. 1–26.
[3] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," BMJ, vol. 372, p. n71, 2021.
[4] A. S. Rao, "BDI agents: From theory to practice," in Proc. International Conference on Multi-Agent Systems (ICMAS), 1995, pp. 312–319.
[5] M. Wooldridge and N. R. Jennings, "Intelligent agents: Theory and practice," Knowledge Engineering Review, vol. 10, no. 2, pp. 115–152, 1995.
[6] S. Yao et al., "ReAct: Synergizing reasoning and acting in language models," in International Conference on Learning Representations (ICLR), 2023.
[7] C. Moralles et al., "A Systematic Literature Review of Agentic AI: Definitions, Architectures, and Challenges," source paper provided for this project, 2026.