International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

AI Agents and Autonomous Decision-Making

Ms Ambadi Varshitha
Ms Gunde Rushika
Ms Maryala Varshitha
|ambadivarshitha@gmail.com , |gunderushika2006@gmail.com
|varshithamaryala@gmail.com Mobile no :9515996974,9030953681,9618753581 Students
Vishwa Vishwani Institute of Systems and Management, BSMS
Thumkunta, Hyderabad, India

ABSTRACT:

Artificial intelligence (AI) agents are systems that can do tasks and make choices with help from people. This article looks at four research papers about AI agents, Agentic AI, industrial AI systems and human-AI collaboration.AI agents can see their environment plan what to do learn from experiences and use feedback to get better at their work. Tools like reinforcement learning and large language models help them solve problems. They are used in areas, such as healthcare, finance, transport, robotics and business. Ai agents do not always make the right choices. They can have issues, with accuracy, privacy, security and bias. Sometimes it is hard to figure out how they make decisions and who is to blame when something goes wrong. To make decisions AI agents need clear goals, good rules and regular checks. Human oversight is important especially when decisions are risky. The amount of control should match the task. This article talks about the need to use AI agents in a responsible way. It also says that they should be checked based on performance, safety, transparency, cost and human oversight.

Keywords: AI agents, decision-making Agentic AI, reinforcement learning, human-AI collaboration, AI governance.

1.INTRODUCTION

Artificial intelligence has improved a lot over time. In the past systems usually followed rules and had difficulty when situations changed. New ways, like machine learning, reinforcement learning, deep learning and large language models have helped AI systems manage complicated tasks.

An AI agent is a system that watches its environment processes information makes a choice and tries to reach a goal. It can answer questions look for information use programs check equipment or control a robot. Some agents follow instructions. Others think about steps and change what they do after getting feedback.

Autonomous decision-making means that a system can make choices with little direct help, from a person. It does not mean that the system is always correct or should have no restrictions. An agent might misinterpret information. Make a dangerous decision. Because of this it needs rules, testing, watching and human control.
These papers cover different parts of the topic: Francesco Piccialli and colleagues talk about agents; Qu and colleagues look at AI-agent design; Vijay Kumar Valaboju talks about automation and human-AI collaboration; and Viswanathan offers an Agentic AI framework. These papers cover parts of the topic. Looking at them together helps show what is already known what issues are still there and what needs study.

1.1 Objectives of the Study

· To explain what AI agents are and how they work.

· To review four papers about AI agents and autonomous decision-making. ·

To compare the main ideas in the four papers.

· To describe the uses, benefits, and risks of AI agents.

· To identify research gaps and suggest future work.

2. REVIEW OF EXISTING RESEARCH

The four papers discuss AI agents from points of view. Two of these papers are review papers. The other two papers discuss agents and Agentic AI. The following sections explain the ideas of these papers in simple terms.

2.1 Francesco Piccialli et al. (2025): Agent AI in Industry 4.0

Francesco Piccialli and colleagues published the article titled "Agent AI: A Comprehensive Survey on Autonomous Agents in Distributed AI for Industry 4.0." The article reviews AI agents that work alone or that work together across systems. The article explains how AI agents use information, learning, communication and decision-making in settings.

The authors used the PRISMA method to search databases such as Scopus, Web of Science IEEE Xplore and ScienceDirect. After checking the paper, the authors included 66 studies. The authors grouped applications of AI agents into nine areas: transport, energy, healthcare, networking, defence, gaming, governance, marketing and e-learning.

The article explains the difference between AI agents that mainly follow fixed rules or that need human supervision and AI agents that can adjust to changes with less direct help. The article also discusses how industry is moving towards systems that can act independently while still working safely with people.

This article connects AI agents with industrial needs. AI agents may affect workers, machines and equipment so AI agents must be safe and controllable. More tests in workplaces are needed to see how AI agents behave when problems or unexpected events occur.

The study also explains that AI agents can share information and work together in industries. They can help workers complete tasks faster and make better decisions. These agents can also reduce human effort and improve the way industries work. However, they may face problems such as data security issues and technical errors. Therefore, more research is needed to make AI agents safer and more reliable in real working environments.

2.2 Qu et al. (2025): A Comprehensive Review of AI Agents

Qu and colleagues published the article titled "A Comprehensive Review of AI Agents: Transforming Possibilities in Technology and Beyond." The article reviews how AI agents are developed how AI agents work, where AI agents can be used and what problems AI agents may face. The article discusses learning, reinforcement learning large language models and groups

of AI agents working together.

The review says that it included 64 studies. The review describes four parts of an AI agent: memory, tools, planning and actions. Memory keeps information. Tools let the AI agent use services such as search or calculators. Planning divides a goal into steps. Actions are the steps the AI agent carries out.

The authors also discuss how AI agents check their work and change their plans. The authors identify problems such as mistakes in situations, decisions that are hard to explain, safety concerns, ethical issues and high computing costs. The AI agent may fail when a task or its surroundings change.

This article gives an explanation of AI agents. The article shows that having memory, tools, planning and actions does not guarantee results. We need to work more to test AI agents in the way explain AI agents' choices and help AI agents work well in new situations.

2.3 Vijay Kumar Valaboju: Intelligent AI Agents and Human–AI Collaboration

Vijay Kumar Valaboju's article titled "Intelligent AI Agents: Enhancing Automation, Decision-Making and Human-AI Collaboration" discusses how intelligent AI agents may support automation and decisions. The article describes types of AI agents. Reactive AI agents respond to what's happening now while limited-memory AI agents can use information from the past. Theory-of-mind AI agents and self-aware AI agents are ideas used to describe types of intelligence; these ideas should not be treated as abilities that today's AI systems already have.

The article discusses uses of AI agents in healthcare, finance, cybersecurity, manufacturing and logistics. AI agents may help people find information complete repeated tasks and make decisions. The article also explains that AI should support people's knowledge and skills than automatically replace them.

The article raises concerns about bias, privacy the ability to work at a scale and people relying too much on AI. The article also gives claims about efficiency, cost and accuracy. How ever the supplied article does not explain its data and methods in detail to confirm all those figures. For that reason, this review does not treat those figures as proven results.

2.4 Viswanathan (2025): Agentic AI Framework

Panneer Selvam Viswanathan published the article titled "Agentic AI: A Comprehensive Framework for Autonomous Decision-Making Systems in Artificial Intelligence." The article describes Agentic AI as systems that can work toward goals make choices and adjust their actions with ongoing help, from people.
The paper lists four features: deciding on its own pursuing goals learning from outcomes and using knowledge about the context.

Its framework includes receiving data choosing a decision taking action and checking feedback.

In terms the agent receives data decides what to do acts and uses the outcome to guide its next

step. The paper also discusses reinforcement learning, neural networks and pattern recognition. Possible uses include robotics, manufacturing, healthcare, self-driving vehicles and finance.

The paper also discusses responsibility, bias, safety. Working with people. Clear limits, regular checks and a way to stop actions are important when the agent can take several steps without asking a person each time. This paper gives a framework for understanding Agentic AI.

However, claims about performance should not be treated as proof if they are not supported by clear experiments.

Future studies should test these systems to understand the conditions. Use clear measures for reliability, safety and responsibility.

[Table 1: Comparative Analysis of Selected Research Articles on AI Agents and Autonomous Decision-Making]

Article	Main focus	Contribution to this review	Research need
Francesco Piccialli et al. (2025)	Industrial and distributed agents	Shows where industrial agents may be used and how independently they can act	More safety tests in real settings
Qu et al. (2025)	General AI agent architecture	Explains memory, tools, planning, actions, and problems	More reliable agents, clearer decisions, and better testing
Vijay Kumar Valaboju	Automation and human–AI collaboration	Shows how agents may support decisions and people	Clear study methods and checked results
Viswanathan (2025)	Agentic AI framework	Explains goals, changes in behaviors and feedback	Better evidenceand clearer responsibility

This paper gives a general framework for understanding Agentic AI. However, claims about

performance should not be treated as final proof if they are not supported by clear experiments. Future studies should test these systems under stated conditions and use clear measures for reliability, safety, and responsibility.

2.5 Comparison of the Four Articles

The four papers focus on related areas. Francesco Piccialli and colleagues talk about agents, Qu and colleagues describe the main parts of AI agents Vijay Kumar Valaboju looks at automation and working with people and Viswanathan talks, about Agentic AI. Together the papers show that agents can help with tasks but agents must be balanced with safety, reliable results and human control.

3. HOW AI AGENTS MAKE DECISIONS

AI agents usually follow a cycle when making decisions. They start by looking at what's happening around them and collect data. Then they use the data they have or what they remember to think about steps. After that they choose one action. Do it. They see what happens after the action. If the goal is still not achieved, they might rethink their plan. Try a way. The steps they follow can change depending on the agent and the job it needs to do.

3.1 Main Components

• Perception: This is how the agent collects data from things like text, pictures, sensors, databases or other sources.

• Memory: This stores information that important for the job or what the agent learned

earlier. • Reasoning: The agent breaks the goal into parts and thinks about what to do • Tools:

These let the agent use services, like search engines, calculators or software programs. •

Action: This is when the agent actually does the task, either in a space or, in the world. •

Feedback: This step checks the outcome. Helps the agent decide what to do. **3.2 Decision-Making Flowchart**

Start

↓

Observe the surroundings

↓

Collect and understand information

↓

Use memory and plan possible actions
↓

Choose and carry out an action

↓

Check the result

↓

Has the goal been reached?

```
├─── Yes → Finish the task
└─── No → Change the plan and observe again
```

[Figure 1. Simple flowchart of how an AI agent makes decisions.]

Feedback matters because an agent should check whether an action worked. It should not assume that every result is correct. For risky tasks, a person may need to approve an action before it happens.

3.3 Learning and Planning Methods

Reinforcement learning helps an agent learn by taking actions and getting feedback. Reinforcement learning gives rewards when the agent performs correct actions and gives penalties when the agent does something wrong. Large language models help an agent understand language and use it effectively. Large language models also help with planning when they are used with tools and memory. Rule-based methods follow instructions. Other methods can deal with situations where there is uncertainty. The best choice depends on the task the risks involved and the data available. The level of clarity needed in explaining the decision matters a lot.

4. APPLICATIONS OF AI AGENTS

4.1 Healthcare

The AI agents, in healthcare can help organise information, schedule appointments remind patients and monitor health data.AI agents may help healthcare workers manage information. Wrong results can put patients at risk. Personal information must be safe. Important medical decisions need review by professional**s.**

4.2Finance and Business

In finance agents can help find unusual transactions and look at the risks involved. In business they can answer questions, from customers keep track of stock and help with planning. Errors can hurt people and companies so the results need to be checked for accuracy fairness. If they follow the rules.

4.3 Transportation and logistics

In transport and logistics agents can help plan routes look at traffic patterns manage warehouse operations and support self-driving systems. Things can change fast on the road or in shipping so these systems must be tested carefully. Any system that controls vehicles or machines also

needs emergency plans in case something goes wrong.

4.4 Manufacturing and Robotics

In factories agents can watch machines, spot defects, in products, schedule maintenance and help robots work together smoothly. They can cut down on repeat tasks. If they make mistake, it could break equipment or put workers at risk. That's why safety rules and human oversight are essential.

4.5 Education and cybersecurity

In education agents can support students in their learning. Help teachers create lesson materials. In cybersecurity they can detect activity and assist in responding to threats. Both of these uses need to be checked for answers, privacy concerns and false alerts. Because mistakes here can cause real problems.

5.MAIN CHALLENGES

5.1 Accuracy and Reliability

AI agents might give wrong answers, use the wrong tool or may fail when faced with something new. Just because an agent does well in a test doesn't mean it will work reliably in life. Testing should include tricky situations and it should keep going even after the system is into use

5.2 Privacy and Security

AI agents can handle financial or business data. If security isn't strong that data could be exposed. Also, bad or harmful inputs might trick the agent into acting in ways. Systems must have limits on who can access what, handle data safely and get security checks.

5.3 Bias and Fairness

AI agents can copy patterns from the data they were trained on or from how they were built. This can be a problem in areas like hiring giving loans or medical care. Developers should look at results across groups and check for unfair differences.

5.4 Transparency and Accountability

People often don't know why an AI made a decision. It can also be unclear who is to blame if that decision causes harm. Keeping records giving clear reasons, for decisions assigning clear responsibility and letting users report issues can make things better.

5.5 Cost and Overdependence

Some AI agents need computers, constant upkeep and experts to run them. People might also trust AI results without thinking.
Companies should think about all the costs involved. Help users learn when to question an ai result

6. HUMAN OVERSIGHT AND RESPONSIBLE USE

Human oversight means that real people can look at what an AI agent does guide it or stop it if

needed. In a human-in-the-loop system a person must say yes before the AI takes an action. In a human-on-the-loop system someone watches the AI while it works and can jump in if something goes wrong. In a human-out-of-the-loop system people don't get involved with every decision. The system must have strong safety rules built in.

The level of control should match the possible risk. For example, sorting emails doesn't need oversight but helping with medical decisions or running factory machines does. Responsible use means setting goals limiting what the AI can do testing it often protecting privacy checking for unfair bias keeping records of its choices and having a safe way to shut it down if needed.

7. EVALUATING AI AGENTS

AI agents must be examined for safety, clear explanations, cost and human control as well as whether they complete a task. The table, below shows how to judge AI agents. It does not contain test results from the four papers.

[Table 2: Evaluation Criteria for AI Agents And Decision Making]

Area	**Question**	**Possible measures**
Performance	Does the agent complete tasks successfully?	Success rate, accuracy, response time
Safety	Does it avoid harmful actions?	Error rate, safety tests, shutdown tests
Transparency	Can decisions be checked?	Decision logs, explanations, audit results
Cost	Is it practical to operate?	Running cost, computing resources, energy use
Human control	Can people review or correct it?	Approval points, intervention process, assigned responsibility

Testing should happen before and after an agent is put into use. researchers should report mistakes and limits well as successful results. using shared ways to test agents would make it easier to compare systems and choose the one for a task.

8. Research Gaps and Future Scope The four papers highlight areas that need more research. Agents need tests in real settings not only in controlled tests. Researchers should find better ways to explain decisions protect systems from harmful inputs and stop unsafe actions.

More research is also needed on how people and agents work. When should a person approve

a decision? How should an agent show that it is unsure? How can users correct mistakes? Researchers should also improve fairness checks make responsibility clearer and find ways to lower computing costs and energy use. When several agents work together, they need ways to share information and avoid actions that conflict. I am hopeful that research gaps will be addressed.

These are areas, for future research not results proved by this review. Future studies should explain their methods clearly. Provide evidence that other researchers can check or repeat.

9.Conclusion

Artificial intelligence agents are changing from systems that follow rules to systems that can plan observe act and learn from feedback. These agents could be useful in healthcare, finance, transport, manufacturing, education and cybersecurity. This review looked at four papers about agents, design of AI agents, automation working with humans and AI together and Agentic AI.

The papers show that AI agents can help with tasks that happen over and again and with making decisions. Just being independent is not enough. Agents can make errors show unfairness, share information or make choices that are hard to understand. Safety, correctness, privacy, openness, cost and who is responsible must also be taken into account.

Future studies should test agents, in life situations explain their decisions better make security and fairness better use the same ways to check how well they do and decide when humans should look at what they do. AI agents should have limits and proper human supervision. People should be able to check or stop actions when needed.

References

[1] Piccialli, F., Chiaro, D., Sarwar, S., Cerciello, D., Qi, P., & Mele, V. (2025). Agent AI: A comprehensive survey on autonomous agents in distributed AI for Industry 4.0. Expert Systems with Applications, 291, 128404. https://doi.org/10.1016/j.eswa.2025.128404

[2] Qu, X., Damoah, A., Sherwood, J., Liu, P., Jin, C. S., Chen, L., Shen, M., Aleisa, N., Hou, Z., Zhang, C., Gao, L., Li, Y., Yang, Q., Wang, Q., & De Souza, C. (2025). A comprehensive review of AI agents: Transforming possibilities in technology and beyond. arXiv:2508.11957. https://arxiv.org/abs/2508.11957

[3] Vijay Kumar Valaboju, V. K. (2025). Intelligent AI agents: Enhancing automation, decision making, and human-AI collaboration. Zenodo. https://doi.org/10.5281/zenodo.14947942

[4] Viswanathan, P. S. (2025). Agentic AI: A comprehensive framework for autonomous decision-making systems in artificial intelligence. International Journal of Computer Engineering and Technology, 16(1), 862–880. https://doi.org/10.34218/IJCET_16_01_069