

What are the components of AI agents?

What are the components of AI agents?

AI agents make intelligent decisions and interact seamlessly with digital systems, requiring minimal human intervention. But what makes these agents truly intelligent? At their core, AI agents rely on a set of interconnected components that enable them to perceive their environment, process information, decide, collaborate, take meaningful actions and learn from their experience.

There are many types of AI agents with different capabilities, and the behavior of agents is governed by the AI [agent architecture](#) in which they operate.

On one end, reactive agents are simple reflex agents that respond instantly to stimuli, sometimes with actuators that allow them to interact with their environment. Model-based reflex agents use an internal model of the environment to enhance their decision-making. On the other end of the spectrum, proactive cognitive agents are capable of advanced reasoning and long-term planning. Some agents specialize in specific tasks and others are designed to lead other agents as a kind of “conductor” in an [AI orchestration](#).

With that caveat, these are the primary components of [artificial intelligence](#) agents, each crucial for creating adaptive, intelligent systems.

Industry newsletter

The latest AI trends, brought to you by experts

Get curated insights on the most important—and intriguing—AI news. Subscribe to our weekly Think newsletter. See the [IBM Privacy Statement](#).

We use your email to validate you are who you say you are, to create your IBMid, and to contact you for account related matters.

Business email

Your subscription will be delivered in English. You will find an unsubscribe link in every newsletter. You can manage your subscriptions or unsubscribe [here](#). Refer to our [IBM Privacy Statement](#) for more information.

Your subscription will be delivered in English. You will find an unsubscribe link in every newsletter. You can manage your subscriptions or unsubscribe [here](#). Refer to our [IBM Privacy Statement](#) for more information.

Perception and input handling

Agentic AI must be able to ingest and interpret information from various sources. Inputs can come in different forms, including user queries, system logs, structured data from APIs or sensor readings. The agent must be able to parse and understand this information, often using AI technologies such as [natural language processing](#) (NLP) for text-based inputs or data extraction techniques for structured sources. The complexity of the perception module depends on the agent's purpose; for example, a

chatbot such as Amazon's Alexa relies on NLP to interpret human input, while a self-driving car processes camera feeds, LIDAR data and radar signals to recognize objects and navigate roads. This overlapping multisensor fusion coupled with [computer vision](#) gives autonomous vehicles a perception of their environment in real time.

After raw data is received, the perception module cleans, processes and structures it into a usable format. AI solutions such as speech-to-text conversion, object detection, sentiment analysis, entity recognition and anomaly detection are often employed. In real-time AI systems, perception must be efficient and adaptive, filtering out noise and prioritizing relevant information. The accuracy and robustness of this module directly impact the AI agent's effectiveness, as misinterpretations in perception can lead to incorrect decisions and actions.

[Prompt engineering](#) might be required to successfully guide the behavior of agents within certain workflows.

Planning and task decomposition

Unlike reactive agents that respond instinctively to immediate inputs, planning agents map out sequences of actions before execution. This module is important for AI applications such as autonomous robots, logistics optimization and AI-driven scheduling systems.

After the AI understands the input, it needs to break down complex problem into smaller, manageable tasks. Some key components are sequencing actions and determining dependencies between tasks. AI agents rely on logic, [machine learning](#) models or predefined heuristics to establish the best course of action.

In multiagent systems, planning becomes even more sophisticated as agents must coordinate or negotiate for resources. Effective planning also incorporates uncertainty, leveraging probabilistic [AI models](#) to prepare for unexpected events. Without a robust planning module, an agent might struggle with long-term tasks, fail to optimize processes or become inefficient when dealing with changing conditions.

Memory

The memory module enables the AI agent to retain and recall information, helping ensure that it can learn from past interactions and maintain context over time. This module is typically divided into short-term and long-term memory. Short-term memory stores session-based context, allowing an AI assistant to recall recent messages in a conversation and maintain coherence. This allows for in-context learning. Long-term memory, alternatively, consists of structured knowledge bases, [vector embeddings](#) and historical data that the agent can refer to when deciding.

Memory persistence and organization are crucial for improving personalization in applications such as customer support bots, recommendation engines and virtual assistants. Without an efficient memory module, an agent functions statelessly, forcing users to repeat information and diminishing the user experience. Memory also plays a role in multiagent systems, where agents share and update a collective knowledge base to improve collaboration.

Reasoning and decision-making

The simple chatbots of the previous decade used predefined rules to choose from a narrow set of decisions. More advanced AI agents work to evaluate different solution paths, assess performance and refine their approach over time. At the core of an agent is the reasoning module. This module determines how an agent reacts to its environment by weighing different factors, evaluating probabilities and applying logical rules or learned behaviors. Depending on the complexity of the AI, reasoning can be rule-based, probabilistic, heuristic-driven or powered by [deep learning](#) models. Two popular reasoning paradigms are ReAct (Reasoning and Action) and ReWOO (Reasoning WithOut Observation).

Various agent types approach reasoning differently. For example, goal-based agents decide by considering a predefined goal and selecting actions that lead to achieving that specific goal. These agents focus on whether an outcome is achieved, rather than optimizing for the best possible outcome. Whereas utility-based agents take decision-making one step further by evaluating not just whether a goal is met, but how optimal the outcome is, based on a utility function.

Simple, rule-based AI systems follow predefined logic, such as "if X happens, do Y." More advanced systems use Bayesian inference, [reinforcement learning](#) or [neural networks](#) to adapt dynamically to new situations. This module can also implement [chain-of-thought](#) reasoning and multistep problem-solving techniques, which are essential for AI applications such as automated financial analysis or legal contract review. The agent's ability to effectively reason and make informed decisions determines an agent's overall intelligence and reliability in handling complex tasks.

Action and tool calling

The action module implements the agent's decisions in the real world, allowing it to interact with users, digital systems or even physical environments. After the reasoning and planning modules determine an appropriate response, the action module executes the necessary steps, whether it's calling a tool such as an API or interacting with the external environment by moving a robotic arm.

Agentic workflows may require access to external tools, datasets, [APIs](#) and [automation](#) systems to complete tasks. Tool calling is the mechanism used in [agentic AI](#) systems where an agent invokes external tools, APIs or functions to extend its capabilities beyond its native reasoning and knowledge. This allows the AI to perform actions, retrieve real-time data, execute computations and interact with external systems dynamically.

In short, tool calling enables a [large language model](#) (LLM) to interface with structured tools, thus granting the model access to information beyond the data used in training.

Communication

The communication module enables an agent to interact with humans, other agents or external software systems, helping ensure seamless integration and collaboration. This module handles natural language generation (NLG) and protocol-based messaging. The sophistication of communication varies; simple agents can follow predefined scripts, while advanced agents use [generative AI](#) models trained on vast amounts of data to generate dynamic, context-aware responses.

The communication component is important for [multiagent systems](#) (MASs) for sharing knowledge, negotiating actions or coordinating tasks. For instance, in finance, multiple agents can analyze market trends and exchange insights to optimize trading strategies. Similarly, AI-powered supply chain networks rely on software agents to sync inventory data, predict shortages and optimize logistics. In human-facing use cases, such as virtual assistants or chatbots, this module helps ensure that responses

feel natural, informative and engaging. The ability to communicate effectively with human agents enhances an agent's usability, making it more valuable across different domains.

Learning and adaptation

A key feature of intelligent agents is their ability to learn from past experiences and improve over time. Learning algorithms enable an agent to recognize patterns, refine predictions and adjust its decision-making processes based on feedback. This is achieved through various learning paradigms, including **supervised learning**, **unsupervised learning** and **reinforcement learning**.

For example, a customer service **chatbot** with a learning module can analyze past interactions to improve its tone, accuracy and response efficiency. Similarly, a recommendation system can continuously refine its suggestions based on user preferences. Reinforcement learning agents, such as those used in robotics and gaming, optimize their actions by maximizing rewards and minimizing penalties. Without a learning module, an AI system would remain static and unable to adapt to new trends, user expectations or unforeseen challenges such as dependency failures.

Across various industries, from healthcare to supply chain to transportation, we can expect the deployment of many more agents, enabled by their impressive scalability. Leaders will need to keep abreast of the current state of agentic technology in order to take full advantage of these tools while also taking **ethical considerations** into account.