AI agents vs. AI assistants

Imagine you are a movie star or star footballer. You probably have an agent and an assistant. Your assistant does tasks for you, based on your requests. They might make dinner reservations, pick up the dry cleaning, organize fan mail and help maintain your calendar.

Your agent is different. They are using their expertise day and night to maximize your opportunities and income. They can act based on your prompts—maybe a product you'd love to endorse—but they don't need prompts to continue to do their job. In fact, your Hollywood agent probably supports you in ways you wouldn't even know to ask.

The key difference between an artificial intelligence (AI) assistant and an AI agent is similar. AI assistants are reactive, performing tasks at your request. AI agents are proactive, working autonomously to achieve a specific goal by any means at their disposal.

Together, assistants and agents elevate great performers, making them or keeping them stars. In much the same way, AI assistants and AI agents can make individual workers and businesses better by performing simple and complex tasks.

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.

AI assistants: Awaiting your instructions

An AI assistant is an intelligent application that understands natural language commands and uses a conversational AI interface to complete tasks for a user. Many modern virtual assistants, such as Amazon's Alexa and Apple's Siri, rely on these capabilities to enhance user interactions.[1]

The first AI assistants relied mostly on rule-based instructions, preprogrammed responses and predefined tasks. Today, AI assistants are almost entirely machine learning (ML) or foundation model-based.

How AI assistants work

AI assistants are built by a foundation model (for example, IBM® Granite™, Meta's Llama models or OpenAI's models). Large language models (LLMs) are a subset of foundation models that specialize in text-related tasks. They enable assistants to understand queries that are submitted by humans and offer relevant information, suggestions or next step actions, which help organizations simplify access to information, automate repetitive tasks and streamline complicated workflows. In business, AI assistants also assist with data analysis, allowing users to efficiently extract insights.

Key features of AI assistants

- **Conversational AI**: LLM-based AI assistants can use natural language processing (NLP) to communicate with users through a chatbot interface. AI chatbot examples include Microsoft Copilot, ChatGPT and IBM watsonx™ Assistant. These assistants integrate with APIs to expand their capabilities.
- **Prompts**: AI assistants need a well-defined problem or a query to get started. AI assistants require continuous user input.
- **Recommendation**: An AI assistant can suggest information or actions based on data it can access. Users should review outputs for accuracy.
- **Tuning**: Users can adapt AI models to more specific tasks through tuning, which eliminates the need to retrain the model. With fine-tuning, they can give models that are labeled examples to tailor them to the target task. Through prompt-tuning, practitioners can give models a task-specific context.

AI assistant limitations

AI assistants have several limitations:

- **They require defined prompts to take action**. While AI assistants can use tools to perform tasks, their capabilities are limited to predefined functions they have been equipped and trained to handle. For example, an AI assistant can use a spreadsheet to generate a table comparing "x versus y," but it cannot independently decide to create such a comparison without a specific prompt.
- **They do not necessarily have persistent memory**. AI assistants can be tailored to fit a user's needs, but they do not inherently retain information from past user interactions. The AI models that power assistants do not continuously learn or evolve based on usage; instead, improvements occur only when the developers release updated versions. However, some AI assistants can reference prior conversations within a session by storing relevant details in their context window or by using a feature that is called "memory" to recall selected information and improve future responses.

AI agents: Taking initiative

To quote Elvis Presley, "A little less conversation, a little more action, please." Enter AI agents.

An AI agent refers to a system or program that can autonomously complete tasks on behalf of users or another system by designing its own workflow and by using available tools.

More autonomous, connected and sophisticated than AI assistants, AI agents can encompass a wide range of functions beyond NLP. These include decision-making, problem-solving, interacting with external environments and executing actions.

How AI agents work

Whereas AI assistants need users to provide prompts for every action, AI agents can operate independently after an initial kickoff prompt. They evaluate assigned goals, break tasks into subtasks and develop their own workflows to achieve specific objectives.

These agents are deployed across various enterprise applications, from software design and IT automation to code-generation tools and conversational assistants. Using advanced NLP from LLMs, AI agents comprehend user inputs step-by-step, strategize their actions and determine when to call on external tools.

Key features of AI agents

- **Greater autonomy**: After an initial prompt, AI agents can continue working without further input, reducing the need for human intervention at every stage. Unlike assistants, which suggest actions for users to approve, AI agents use multicomponent autonomy to independently reason, decide and problem-solve by using external data sets and tools. Their ability to break out of a pure chat-based framework enables proactive decision-making and learning, ultimately saving employees time by handling complex workflows on their own. Newer models are improving reasoning capabilities to support this.[2]

- **Connectivity:** AI agents unify various capabilities into a single workflow, eliminating bottlenecks that arise from disconnected systems. By integrating seamlessly with external applications, data sources and other AI models, they enhance productivity while reducing friction between different components of a process.

- **Decision-making and action**: The ability to call on tools by itself does not make an LLM an agent. AI agents can also act autonomously and decide which tools to use and when. Grounded in foundation models, AI agents go beyond chat to accomplish tasks on their own, based on a specific goal and go beyond the foundation model for additional information and capabilities. They analyze problems, break them into subtasks and plan their next steps autonomously. This makes them effective for handling complex, ambiguous problems. Some agents, such as Anthropic's Claude, even demonstrate computer use, where an LLM can click, type and operate a computer to complete tasks.[3]

- **Persistent memory and adaptive learning:** Compared to AI assistants, AI agents have a greater capacity to learn. They store previous actions, conversations and experiences, enabling them to refine their approach over time. With persistent memory, AI agents can recall past interactions to improve future responses, while adaptive learning allows them to adjust their behavior based on feedback and outcomes. Because they integrate with external applications and tools, they can act on real-time data rather than relying solely on their initial training. Over repeated interactions, they become more efficient, context-aware and better aligned with user needs.

- **Task chaining:** AI agents don't complete tasks in isolation—they break complex workflows into smaller, manageable steps. AI agents identify dependencies between tasks, which help ensure that each step logically flows into the next. This ability enables structured execution across multi-step processes and makes automation more dynamic.

- **Team play**: AI agents often specialize in specific tasks—one may excel at fact-checking, while another is better at research. These agents can collaborate, forming teams to tackle complex challenges together. IBM currently supports AI agents that are written in LangChain, with LlamaIndex integration coming soon. Instead of being developer-heavy, IBM's framework enables users can compose and edit AI agents in a low-code or no-code environment.

Benefits of AI agents and AI assistants

AI agents and AI assistants offer numerous benefits, from optimizing workflows to enhancing user experience.

Complementary AI solutions: AI agents specialize in performing specific or complex tasks autonomously, while AI assistants excel at understanding and interacting with users naturally. Together they create more powerful and intuitive AI solutions.

Optimized workflows and increased productivity: AI tools and gen AI streamline processes, automate routine tasks, and assist humans with problem-solving, improving overall efficiency.

Enhanced user experience: AI assistants provide interactive support, adapt to user needs and learn from feedback and conversation history to offer more personalized interactions.

Autonomous operations and scalability: AI agents can operate independently, manage multiple tasks simultaneously and scale to handle complex processes without direct human intervention.

Improved task management and collaboration: AI agents can interpret user needs and assign tasks to AI assistants. Assistants can use agent-generated data to create more intuitive outputs. These abilities enhance coordination.

Improved integration potential: As AI models evolve, they can better integrate conversational and autonomous components, enabling seamless task handoffs and delivering higher-quality responses in less time.

AI assistants and AI agents use cases

Customer experience

AI assistants improve customer experience by providing real-time, real-world support across chat, voice and email. They handle common customer inquiries, guide users through self-service options, and escalate complex issues when needed. Using NLP, they personalize interactions, recommend products, and help customers complete transactions quickly. Their anytime availability improves customer satisfaction and reduces costs.

AI agents take customer experience and customer support further by adapting to user behavior in real time. Unlike AI assistants with scripted responses, AI agents learn and improve interactions, whether

it's simulating job interviews or handling complex support issues autonomously. They work across websites, apps and IoT devices to create smooth and highly personalized user experiences.

Banking and financial services

AI assistants provide secure, real-time banking support by handling balance inquiries, fraud alerts and loan applications. They also help customers manage their finances by analyzing spending habits and offering personalized budgeting advice.

AI agents proactively prevent fraud by monitoring transactions in real-time, detecting suspicious activity and blocking threats before they escalate. Unlike assistants that just send fraud alerts, AI agents adjust security protocols, refine risk models and coordinate with fraud detection systems to stay ahead of emerging threats. In trading and investment, AI agents analyze market trends, execute trades and adjust portfolios without human intervention.

Human resources

AI assistants help organizations streamline recruitment by generating job descriptions, sorting resumes and drafting personalized messages. Beyond hiring, they assist in onboarding by guiding new employees through policies, benefits and training materials.

AI agents take HR automation further by managing and optimizing talent acquisition, employee engagement and workforce planning. They screen candidates, schedule interviews and refine hiring strategies by using past data. For performance management, AI agents analyze employee feedback, detect trends and recommend training programs. They also automate onboarding, benefits administration and compliance tracking, making HR operations more data-driven and efficient.

Healthcare

AI assistants play a key role in human resources (HR) process automation by helping to improve patient experiences and streamline administrative tasks. They answer patient questions in real-time, assist with appointment scheduling, billing and prescription refills and provide self-service access to medical records. AI assistants help doctors by summarizing patient histories and flagging urgent cases. AI assistants also help organize documentation, helping to ensure that formatting remains consistent for easier accessibility.

AI agents support medical decision-making in complex environments. In emergency rooms, multi-agent systems help triage patients, adjusting priorities based on real-time data from sensors. AI agents also optimize drug supply management, predict shortages and adjust treatment plans based on patient responses.

Risks of AI agents and AI assistants

There are risks and limitations with AI-powered technologies to consider. LLMs are brittle, meaning that they are susceptible to even the smallest prompt changes that cause invalid structures, an incorrect payload or hallucinations. This means that AI agents and AI assistants might fail if, for example, the underlying foundation model hallucinates or breaks.

For AI agents, especially, it is early days. If they have trouble creating comprehensive plans or fail at reflecting on their findings, AI agents get stuck in infinite feedback loops. And because AI agents consider external environments and tools, they must deal with the changes to those tools. Over time, those changes might cause the agent set up to break. AI assistants, on the other hand, can be reliably used in most cases, as they do not use external tools.

For harder tasks, AI agents require a great deal of training, and they might still take a long time to complete them. Also, they can often be expensive.

Today's foundation models are not quite intelligent enough to reliably act as agents but advances in model reasoning will improve the situation. Therefore, we are still in the early days of understanding and seeing what AI agents can do. This future of AI might see expanded self-guided applications of AI technology. But at this stage of development, human intervention is often still necessary to offer guidance or redirection.