Agentic AI

In the context of generative artificial intelligence, **AI agents** (also referred to as **compound AI systems** or **agentic AI**) are a class of intelligent agents distinguished by their ability to operate autonomously in complex environments. Agentic AI tools prioritize decision-making over content creation and do not require human prompts or continuous oversight [1].

Overview

AI agents possess several key attributes, including complex goal structures, natural language interfaces, the capacity to act independently of user supervision, and the integration of software tools or planning systems. Their control flow is frequently driven by large language models (LLMs) [2]. Agents also include memory systems for remembering previous user-agent interactions and orchestration software for organizing agent components [3].

Researchers and commentators have noted that AI agents do not have a standard definition [2,4,5,6]. The concept of agentic AI has been compared to the fictional character J.A.R.V.I.S.[7].

A common application of AI agents is the automation of tasks—for example, booking travel plans based on a user's prompted request [8,9]. Prominent examples include Devin AI, AutoGPT, and SIMA [10]. Further examples of agents released since 2025 include OpenAI Operator [11], ChatGPT Deep Research [12], Manus [13], Quark (based on Qwen) [14], AutoGLM Rumination [14], and Coze (by ByteDance) [14]. Frameworks for building AI agents include LangChain [15], as well as tools such as CAMEL [16,17], Microsoft AutoGen [18], and OpenAI Swarm [19].

Companies such as Google, Microsoft and Amazon Web Services have offered platforms for deploying pre-built AI agents [20].

Proposed protocols for standardizing inter-agent communication include the Agent Protocol (by LangChain), the Model Context Protocol (by Anthropic), AGNTCY [12], Gibberlink [22], the Internet of Agents [23], Agent2Agent (by Google) [24], and the Agent Network Protocol [25]. Some of these protocols are also used for connecting agents with external applications [3]. Software frameworks for addressing agent reliability include AgentSpec, ToolEmu, GuardAgent, Agentic Evaluations, and predictive models from H2O.ai [26].

In February 2025, Hugging Face released Open Deep Research, an open source version of OpenAI Deep Research [27]. Hugging Face also released a free web browser agent, similar to OpenAI Operator [28]. Galileo AI published on Hugging Face a leadership board for agents, which ranks their performance based on their underlying LLMs [29].

Memory systems for agents include Mem0 [30, 31], MemGPT [32], and MemOS [33].

History

AI agents have been traced back to research from the 1990s, with Harvard professor Milind Tambe noting that the definition of an AI agent was not clear at the time either. Researcher Andrew Ng has been credited with spreading the term "agentic" to a wider audience in 2024 [34].

Training and testing

Researchers have attempted to build world models [35,36] and reinforcement learning environments [37] to train or evaluate AI agents. For example, video games such as Minecraft [38] and No Man's Sky [39] as well as replicas of company websites [40], have also been used for training AI agents.

Architectural patterns

Common architectural design patterns for agents include:

1) Retrieval-augmented generation [41]
2) ReAct (Reason + Act) [42], an extension of chain-of-thought prompting that queries the
underlying model to explain its reasoning before taking any action [43].
3) Reflexion [41, 42, 43], which uses an LLM to create feedback on the agent's plan of action and stores that feedback in a memory cache.
4) A tool/agent registry [41], for organizing software functions or other agents that the agent can use.
5) One-shot model querying [41], which queries the model once to create the plan of action.

one really knows what an AI agent is" (https://web.archive.org/web/20250512184704/https://techcrunch.com/2025/05/12/even-a16z-vcs-say-no-one-really-knows-what-an-ai-agent-is/). *TechCrunch*. Archived from the original (https://techcrunch.com/2025/05/12/even-a16z-vcs-say-no-one-really-knows-what-an-ai-agent-is/) on May 12, 2025. Retrieved May 15, 2025.
7. Field, Hayden (August 31, 2025). "AI agents are science fiction not yet ready for primetime" (https://web.archive.org/web/20250915055451/https://www.theverge.com/the-stepback-newsletter/767376/ai-agents-jarvis-what-can-they-do). *The Verge*. Archived from the original (https://www.theverge.com/the-stepback-newsletter/767376/ai-agents-jarvis-what-can-they-do) on September 15, 2025. Retrieved November 9, 2025.
8. "AI Agents: The Next Generation of Artificial Intelligence" (https://web.archive.org/web/20250111192703/https://natlawreview.com/article/next-generation-ai-here-come-agents). *The National Law Review*. December 30, 2024. Archived from the original (https://natlawreview.com/article/next-generation-ai-here-come-agents) on January 11, 2025. Retrieved January 14, 2025.
9. "What are the risks and benefits of 'AI agents'?" (https://web.archive.org/web/20241228013835/https://www.weforum.org/stories/2024/12/ai-agents-risks-artificial-intelligence/). *World Economic Forum*. December 16, 2024. Archived from the original (https://www.weforum.org/stories/2024/12/ai-agents-risks-artificial-intelligence/) on December 28, 2024. Retrieved January 14, 2025.
10. Knight, Will (March 14, 2024). "Forget Chatbots. AI Agents Are the Future" (https://web.archive.org/web/20250105095231/https://www.wired.com/story/fast-forward-forget-chatbots-ai-agents-are-the-future/). *Wired*. ISSN 1059-1028 (https://search.worldcat.org/issn/1059-1028). Archived from the original (https://www.wired.com/story/fast-forward-forget-chatbots-ai-agents-are-the-future/) on January 5, 2025. Retrieved January 14, 2025.
11. Marshall, Matt (February 22, 2025). "The rise of browser-use agents: Why Convergence's Proxy is beating OpenAI's Operator" (https://web.archive.org/web/20250222231546/https://venturebeat.com/ai/the-rise-of-browser-use-agents-why-convergences-proxy-is-beating-openais-operator/). *VentureBeat*. Archived from the original (https://venturebeat.com/ai/the-rise-of-browser-use-agents-why-convergences-proxy-is-beating-openais-operator/) on February 22, 2025. Retrieved April 2, 2025.
12. Milmo, Dan (February 3, 2025). "OpenAI launches 'deep research' tool that it says can match research analyst" (https://web.archive.org/web/20250203142402/https://www.theguardian.com/technology

/2025/feb/03/openai-deep-research-agent-chatgpt-deepseek). *The Guardian*. ISSN 0261-3077 (https://search.worldcat.org/issn/0261-3077). Archived from the Original (https://www.theguardian.com/technology/2025/feb/03/openai-deep-research-agent chatgpt-deepseek) on February 3, 2025. Retrieved April 2, 2025.
13. Chen, Caiwei (March 11, 2025). "Everyone in AI is talking about Manus. We put it to the test" (https://web.archive.org/web/20250312113852/https://www.technologyreview.com/2025/03/11/1113133/manus-ai-review/). *MIT Technology Review*. Archived from the original (https://www.technologyreview.com/2025/03/11/1113133/manus-ai-review/) on March 12, 2025. Retrieved April 2, 2025.
14. "China is gaining ground in the global race to develop AI agents" (https://web.archive.org/web/20250602111847/https://restofworld.org/2025/china-ai-agent-openai/). *Rest of World*. June 2, 2025. Archived from the original (https://restofworld.org/2025/china-ai-agent-openai/) on June 2, 2025. Retrieved June 12, 2025.
15. David, Emilia (December 30, 2024). "Why 2025 will be the year of AI orchestration" (https://web.archive.org/web/20241230175615/https://venturebeat.com/ai/three-ways-2025-will-be-the-year-of-agentic-productivity/). *VentureBeat*. Archived from the original (https://venturebeat.com/ai/three-ways-2025-will-be-the-year-of-agentic-productivity/) on December 30, 2024. Retrieved January 14, 2025.
16. "CAMEL: Finding the Scaling Law of Agents. The first and the best multi-agent framework" (https://github.com/camel-ai/camel/). *GitHub*.
17. Li, Guohao (2023). "Camel: Communicative agents for "mind" exploration of large language model society" (https://proceedings.neurips.cc/paper_files/paper/2023/file/a3621ee907def47c1b952ade25c67698-Paper-Conference.pdf) (PDF). *Advances in Neural Information Processing Systems*. **36**: 51991–52008. arXiv:2303.17760 (https://arxiv.org/abs/2303.17760). S2CID 257900712 (https://api.semanticscholar.org/CorpusID:257900712).
18. Dickson, Ben (October 3, 2023). "Microsoft's AutoGen framework allows multiple AI agents to talk to each other and complete your tasks" (https://web.archive.org/web/20241227061127/https://venturebeat.com/ai/microsofts-autogen-framework-allows-multiple-ai-agents-to-talk-to-each-other-and-complete-your-tasks/). *VentureBeat*. Archived from the original (https://venturebeat.com/ai/microsofts-autogen-framework-allows-multiple-ai-agents-to-talk-to-each-other-and-complete-your-tasks/) on December 27, 2024. Retrieved January 14, 2025.
19. "The next AI wave — agents — should come with warning labels" (https://web.archive.org/web/20250114023632/https://www.computerworld.com/article/3

727412/the-next-ai-wave-agents-should-come-with-warning-labels.html). *Computerworld*. January 13, 2025. Archived from the original (https://www.computerworld.com/article/3727412/the-next-ai-wave-agents-should-come-with-warning-labels.html) on January 14, 2025. Retrieved January 14, 2025.
20. David, Emilia (April 15, 2025). "Moveworks joins AI agent library craze" (https://web.archive.org/web/20250415214729/https://venturebeat.com/ai/moveworks-joins-ai-agent-library-craze/). *VentureBeat*. Archived from the original (https://venturebeat.com/ai/moveworks-joins-ai-agent-library-craze/) on April 15, 2025. Retrieved May 14, 2025.
v>
21. David, Emilia (March 6, 2025). "A standard, open framework for building AI agents is coming
from Cisco, LangChain and Galileo" (https://web.archive.org/web/20250309045209/https://venturebeat.com/ai/a-standard-open-framework-for-building-ai-agents-is-coming-from-cisco-langchain-and-galileo/). *VentureBeat*. Archived from the original (https://venturebeat.com/ai/a-standard-open-framework-for-building-ai-agents-is-coming-from-cisco-langchain-and-galileo/) on March 9, 2025. Retrieved April 2, 2025.
22. Zeff, Maxwell (March 5, 2025). "GibberLink lets AI agents call each other in robo-language" (https://web.archive.org/web/20250305141006/https://techcrunch.com/2025/03/05/gibberlink-lets-ai-agents-call-each-other-in-robo-language/). *TechCrunch*. Archived from the original (https://techcrunch.com/2025/03/05/gibberlink-lets-ai-agents-call-each-other-in-robo-language/) on March 5, 2025. Retrieved April 2, 2025.
23. Cooney, Michael (January 30, 2025). "Cisco touts 'Internet of Agents' for secure AI agent
collaboration" (https://web.archive.org/web/20250131133538/https://www.networkworld.com/article/3812618/cisco-touts-internet-of-agents-for-secure-ai-agent-collaboration.html). *Network World*. Archived from the original (https://www.networkworld.com/article/3812618/cisco-touts-internet-of-agents-for-secure-ai-agent-collaboration.html) on January 31, 2025. Retrieved April 2, 2025.

24. Clark, Lindsay (April 10, 2025). "Did someone say AI agents, Google asks, bursting in" (https://web.archive.org/web/20250410112802/https://www.theregister.com/2025/04/10/google_agentic_ai_cloud_next/). *The Register*. Archived from the original (https://www.theregister.com/2025/04/10/google_agentic_ai_cloud_next/) on April 10, 2025. Retrieved May 14, 2025.
25. Stokel-Walker, Chris (June 11, 2025). "Can we stop big tech from controlling the internet with AI agents?" (https://archive.today/20250611131453/https://www.newscientist.com/article/2483880-can-we-stop-big-tech-from-controlling-the-internet-with-ai-agents/). *New Scientist*. Archived from the original (https://www.newscientist.com/article/2483880-can-we-stop-big-tech-from-controlling-the-internet-with-ai-agents/) on June 11, 2025. Retrieved June 12, 2025.
26. David, Emilia (March 28, 2025). "New approach to agent reliability, AgentSpec, forces agents to follow rules" (https://web.archive.org/web/20250412120324/https://venturebeat.com/ai/new-approach-to-agent-reliability-agentspec-forces-agents-to-follow-rules/). *VentureBeat*. Archived from the original (https://venturebeat.com/ai/new-approach-to-agent-reliability-agentspec-forces-agents-to-follow-rules/) on April 12, 2025. Retrieved May 14, 2025.
27. Edwards, Benj (February 5, 2025). "Hugging Face clones OpenAI's Deep Research in 24 hours" (https://web.archive.org/web/20250206125754/https://arstechnica.com/ai/2025/02/after-24-hour-hackathon-hugging-faces-ai-research-agent-nearly-matches-openais-solution/). *Ars Technica*. Archived from the original (https://arstechnica.com/ai/2025/02/after-24-hour-hackathon-hugging-faces-ai-research-agent-nearly-matches-openais-solution/) on February 6, 2025. Retrieved April 2, 2025.28. Wiggers, Kyle (May 6, 2025). "Hugging Face releases a free Operator-like agentic AI tool" (https://web.archive.org/web/20250506221518/https://techcrunch.com/2025/05/06/hugging-face-releases-a-free-operator-like-agentic-ai-tool/). *TechCrunch*. Archived from the original (https://techcrunch.com/2025/05/06/hugging-face-releases-a-free-operator-like-agentic-ai-tool/) on May 6, 2025. Retrieved May 14, 2025.
29. Ortiz, Sabrina (February 14, 2025). "Which AI agent is the best? This new leaderboard can tell you" (https://web.archive.org/web/20250330001709/https://www.zdnet.com/article/which-ai-agent-is-the-best-this-new-leaderboard-can-tell-you/). *ZDNET*. Archived from the original

(https://www.zdnet.com/article/which-ai-agent-is-the-best-this-new-leaderboard-can-tell-you/) on March 30, 2025. Retrieved April 2, 2025.
30. Dickson, Ben (May 8, 2025). "Mem0's scalable memory promises more reliable AI agents that remembers context across lengthy conversations" (http://web.archive.org/web/20250827005641/https://venturebeat.com/ai/mem0s-scalable-memory-promises-more-reliable-ai-agents-that-remembers-context-across-lengthy-conversations/). *VentureBeat*. Archived from the original (https://venturebeat.com/ai/mem0s-scalable-memory-promises-more-reliable-ai agents-that-remembers-context-across-lengthy-conversations/) on August 27, 2025. Retrieved November 28, 2025.
31. Kene-Okafor, Tage (October 28, 2025). "Mem0 raises $24M from YC, Peak XV and Basis Set to build the memory layer for AI apps" (https://archive.ph/0G5cH). *TechCrunch*. Archived from the original (https://techcrunch.com/2025/10/28/mem0-raises-24m-from-yc-peak-xv-and-basis-set-to-build-the-memory-layer-for-ai-apps/) on October 28, 2025. Retrieved November 28, 2025.
32. Bort, Julie (September 23, 2024). "Letta, one of UC Berkeley's most anticipated AI startups, has just come out of stealth" (https://web.archive.org/web/20251006164502/https://techcrunch.com/2024/09/23/letta-one-of-uc-berkeleys-most-anticipated-ai-startups-has-just-come-out-of-stealth/). *TechCrunch*. Archived from the original (https://techcrunch.com/2024/09/23/letta-one-of-uc-berkeleys-most-anticipated-ai-startups-has-just-come-out-of-stealth/) on October 6, 2025. Retrieved November 28, 2025.
33. Nuñez, Michael (July 8, 2025). "Chinese researchers unveil MemOS, the first 'memory operating system' that gives AI human-like recall" (http://web.archive.org/web/20250901004623/https://venturebeat.com/ai/chinese-researchers-unveil-memos-the-first-memory-operating-system-that-gives-ai-human-like-recall/). *VentureBeat*. Archived from the original (https://venturebeat.com/ai/chinese-researchers-unveil-memos-the-first-memory-operating-system-that-gives-ai-human-like-recall/) on September 1, 2025. Retrieved November 28, 2025.
34. O'Brien, Matt (November 18, 2025). "What does 'agentic' AI mean? Tech's newest buzzword is a mix of marketing fluff and real promise" (https://web.archive.org/web/20251118171507/https://apnews.com/article/agentic-ai-agents-microsoft-amazon-518d6ae159d1f4d3343e98a456cb5221). *Associated Press*. Archived from the original (https://apnews.com/article/agentic-ai-agents-microsoft-amazon-518d6ae159d1f4d3343e98a456cb5221) on November 18, 2025. Retrieved November 28, 2025.
35. Knight, Will (May 22, 2025). "A United Arab Emirates Lab Announces Frontier AI Projects—and a New Outpost in Silicon Valley"

(https://web.archive.org/web/20250522202354/https://www.wired.com/story/the-united-arab-emirates-announces-frontier-ai-projects-and-a-new-lab-in-silicon-valley/). *Wired*. ISSN 1059-1028 (https://search.worldcat.org/issn/1059-1028). Archived from the original https://www.wired.com/story/the-united-arab-emirates-announces-frontier-ai-projects-and-a-new-lab-in-silicon-valley/) on May 22, 2025. Retrieved November 9, 2025.
36. Orland, Kyle (December 6, 2024). "Google's Genie 2 "world model" reveal leaves more questions than answers" (https://web.archive.org/web/20241207000413/https://arstechnica.com/ai/2024/12/googles-genie-2-world-model-reveal-leaves-more-questions-than-answers/). *Ars Technica*. Archived from the original (https://arstechnica.com/ai/2024/12/googles-genie-2-world-model-reveal-leaves-more-questions-than-answers/) on December 7, 2024. Retrieved November 9, 2025.
37. Zeff, Maxwell (September 21, 2025). "Silicon Valley bets big on 'environments' to train AI agents" (https://web.archive.org/web/20250916191353/https://techcrunch.com/2025/09/16/silicon-valley-bets-big-on-environments-to-train-ai-agents/). *TechCrunch*. Archived from the original (https://techcrunch.com/2025/09/21/silicon-valley-bets-big-on-environments-to-train-ai-agents/) on September 16, 2025. Retrieved November 9, 2025.
38. Shazhaev, Ilman (November 24, 2025). "Why Game Engines Are Becoming A.I.'s Most Important Testbeds" (https://web.archive.org/web/20251203023639/https://observer.com/2025/11/gaming-training-next-gen-ai/). *Observer*. Archived from the original (https://observer.com/2025/11/gaming-training-next-gen-ai/) on December 3, 2025. Retrieved December 3, 2025.
39. David, Emilia (March 13, 2024). "Google's new AI will play video games with you — but not to win" (https://web.archive.org/web/20250602020104/https://www.theverge.com/2024/3/13/24099024/google-deepmind-ai-agent-sima-video-games). *The Verge*. Archived from the original (https://www.theverge.com/2024/3/13/24099024/google-deepmind-ai-agent-sima-video-games) on June 2, 2025. Retrieved December 3, 2025.
40. Metz, Cade (December 2, 2025). "Silicon Valley Builds Amazon and Gmail Copycats to Train A.I. Agents" (https://archive.ph/QsDWu). *The New York Times*. Archived from the original (https://www.nytimes.com/2025/12/02/technology/artificial-intelligence-amazon-gmail.html) on December 2, 2025. Retrieved December 2, 2025.
41. Liu, Yue; Lo, Sin Kit; Lu, Qinghua; Zhu, Liming; Zhao, Dehai; Xu, Xiwei; Harrer, Stefan; Whittle, Jon (February 1, 2025). "Agent design pattern catalogue: A collection of

architectural patterns for foundation model based agents" (https://www.sciencedirect.com/science/article/pii/S0164121224003224). *Journal of Systems and Software*. 220 112278. doi:10.1016/j.jss.2024.112278 (https://doi.org/10.1016%2Fj.jss.2024.112278). ISSN 0164-1212 (https://search.worldcat.org/issn/0164-1212).
42. Masterman, Tula; Besen, Sandi; Sawtell, Mason; Chao, Alex (April 17, 2024), *The Landscape of Emerging AI Agent Architectures for Reasoning, Planning, and Tool Calling: A Survey* (http://arxiv.org/abs/2404.11584), arXiv:2404.11584 (https://arxiv.org/abs/2404.11584), retrieved December 3, 2025.
43. Wray, Robert E.; Kirk, James R.; Laird, John E. (August 10, 2025). "Applying Cognitive Design Patterns to General LLM Agents" (https://doi.org/10.1007/978-3-032-00800-8_28). *Artificial General Intelligence*. Lecture Notes in Computer Science. Vol. 16058. Berlin, Heidelberg: Springer-Verlag. pp. 312–325. doi:10.1007/978-3-032-00800-8_28 (https://doi.org/10.1007%2F978-3-032-00800-8_28). ISBN 978-3-032-00799-5.