

Die Evolution Agentischer KI

VON REGELBASIRTER PLANUNG ZU AUTONOMEN LLM-AGENTEN

Chronologischer Report •
Überarbeitete Auflage

Zusammenfassung: Agentische KI-Systeme basieren auf jahrzehntelanger akademischer und industrieller Forschung. Dieser Bericht zeichnet den Weg von frühen Planungsarchitekturen (BDI/SOAR) über den Transformer-Durchbruch bis hin zur Symbiose aus kommerziellen Frontier-Modellen und der Open-Source-Demokratisierung nach.

1. HISTORISCHE GRUNDLAGEN & ENTWICKLUNGSPFADE

Die Entwicklung autonomer Agenten verlief über verschiedene Paradigmenwechsel – von symbolischer Logik und regelbasierten Systemen bis hin zu LLM-gestützten geschlossenen Feedback-Schleifen.

1950er–1980er

Symbolische KI, Expertensysteme & frühe Planung

Fokus auf regelbasierte Logik und deterministische Suchalgorithmen. Entwicklung von Kognitionsarchitekturen wie SOAR und ACT-R.

Late
1980er/90s

BDI-Architekturen (Belief-Desire-Intention)

Formalisierung des BDI-Modells (Rao & Georgeff, 1987). Agenten planen Handlungen basierend auf Überzeugungen, Zielen und Absichten.

2016–2017

Hybrid RL (AlphaGo) & Transformer-Durchbruch

DeepMind demonstriert mit AlphaGo komplexe Strategieplanung (Kombination aus RL, Supervised Learning und MCTS). 2017 revolutioniert die Transformer-Architektur die Textverarbeitung.

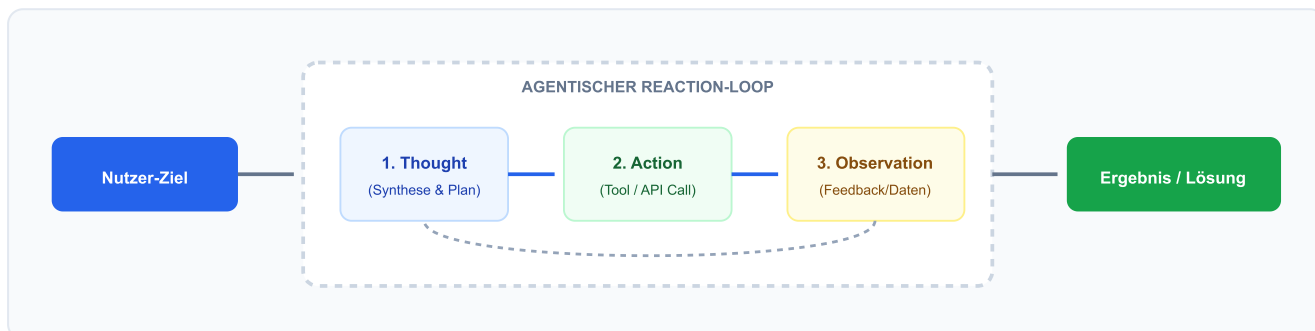
Okt 2022

Das ReAct-Paradigma (Princeton / Google)

Paper "ReAct: Synergizing Reasoning and Acting in Language Models" etabliert ein Kernkonzept moderner LLM-Agenten: Das Verknüpfen von Nachdenken (Thought) und Handeln (Action).

2. FUNKTIONSPRINZIP DES REACT-LOOPS

Im Gegensatz zu klassischen Single-Turn-Prompting agiert ein agentisches System in einer Schleife aus Planung, Werkzeugnutzung und Umwelt-Feedback:



Klassisches LLM (Single-Turn)

Nimmt Kontext auf und erzeugt unmittelbar eine einmalige Textantwort. Keine direkte Interaktion mit der Außenwelt.

LLM-Agent (Multi-Step Loop)

Nutzt Modelle als Steuerungseinheit zur schrittweisen Ausführung von Aufgaben, API-Aufrufen und Selbstkorrektur.

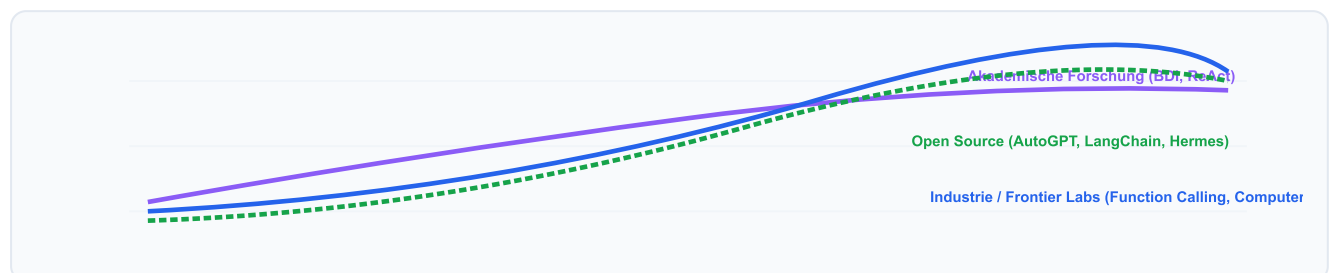
3. SKALIERUNG & FOKUSSIERUNG (2023 – 2025)

Ab 2023 entwickelten sich Agentensysteme schrittweise zu einem der primären Schwerpunkte von akademischen Teams, Open-Source-Communities und Industrie-Laboratorien.

2023 (Frühjahr)	Pionierphase Open Source (AutoGPT, LangChain, Function Calling) AutoGPT und BabyAGI zeigen das Potenzial autonomer Schleifen. OpenAI führt strukturiertes <i>Function Calling</i> ein, was Schnittstellen-Aufrufe verlässlicher macht.
2023/2024	Multi-Agenten-Frameworks (CrewAI, AutoGen) Systeme spezialisieren sich auf Rollenverteilungen: Mehrere spezialisierte Agenten arbeiten kollaborativ an komplexen Workflows.
22. Okt 2024	Anthropic "Computer Use" & Model Context Protocol (MCP) Claude 3.5 Sonnet lernt die optische Bedienung von Benutzeroberflächen. Mit MCP wird ein verbreitetes offenes Protokoll zur Schnittstellen-Anbindung ins Leben gerufen.
2025	Systemische Agenten-Integration (OpenAI, Google) Rollout von Browser-Agenten (OpenAI Operator) sowie tief integrierten Workspace-Agenten (Google Gemini Computer Use).

4. DAS ZUSAMMENSPIEL VON FORSCHUNG, INDUSTRIE & OPEN SOURCE

Die Evolution agentischer KI ist keine Einbahnstraße: Forschungseinrichtungen, globale Tech-Konzerne und die Open-Source-Community befruchten sich kontinuierlich gegenseitig.



Lokale Barrierefreiheit & Spezialisierung

Open-Source-Projekte wie der *Hermes Agent* (Nous Research) machen agentische Architekturen und Agentic Fine-Tuning lokal und ohne Vendor-Lock-in nutzbar.

Fazit zur Entwicklung

Kein Akteur hat Agenten "alleine erfunden". Das Paradigma reifte durch akademische Grundlagen, Open-Source-Experimente und industrielle Rechenkapazität.

5. LITERATUR- & QUELLENVERZEICHNIS

- Rao, A. S., & Georgeff, M. P. (1991): *Modeling Rational Agents within a BDI-Architecture*. Proceedings of KR'91.
- Silver, D. et al. (DeepMind, 2016): *Mastering the game of Go with deep neural networks and tree search*. Nature 529.
- Vaswani, A. et al. (Google Brain/Research, 2017): *Attention Is All You Need*. NeurIPS 2017.
- Yao, S. et al. (Princeton & Google Research, 2022): *ReAct: Synergizing Reasoning and Acting in Language Models*. arXiv:2210.03629.
- OpenAI (2023): *Function Calling and Other API Updates*. Official OpenAI Blog (June 2023).
- Anthropic (2024): *Developing Computer Use capabilities and Model Context Protocol (MCP)*. Research & Product Announcement (October 2024).
- Nous Research (2026): *Hermes Agent & Advanced Function Calling Framework*. Released under MIT License (February 2026).