



KI-Agentensysteme in der Forschung – geliebte oder bedrohliche „Zauberlehrlinge“?

KI-Agentensysteme verändern den Wissenschaftsbetrieb in einem ungeahnten Tempo – von Doris Weßels

Die Entwicklung moderner KI-Systeme gleicht an vielen Stellen dem Spannungsfeld von Magie und Kontrollverlust in Goethes Ballade „Der Zauberlehrling“. Es geht aber nicht darum, Besen zum Wasserholen zu animieren, sondern KI-Agenten-Systeme einzusetzen, die eigenständig Hypothesen generieren, Experimente planen, ggf. sogar durchführen und wissenschaftliche Publikationen verfassen sollen. Was vor wenigen Jahren noch wie Science-Fiction anmutete, wird immer mehr zur Realität: KI-Agentensysteme verändern den Wissenschaftsbetrieb in einem ungeahnten Tempo. Dieser Dynamik und den Implikationen, bis hin zum drohenden Kontrollverlust, muss sich das Wissenschaftsmanagement nicht erst morgen, sondern bereits heute stellen.

Die weltweite KI-Entwicklung verläuft nach wie vor mit ungebremsster Geschwindigkeit. OpenAI kündigte an, bis 2028 vollautonome KI-Forscher zu entwickeln, die komplette wissenschaftliche Projekte selbstständig durchführen können. Google stellte mit dem „AI co-scientist“ ein Multi-Agentensystem vor, das Hypothesen generiert, Literatur durchsucht und Forschungspläne erstellt. Japan investiert massiv in „Autonomous Labs“,

d. h. vollautomatisierte Labore, die von der Hypothese bis zur Publikation eigenständig arbeiten. Das indische IIT Delhi entwickelte mit Beteiligung deutscher und dänischer Forschenden den KI-Agenten AILA, der Rasterkraftmikroskope autonom bedient und Aufgaben, die früher einen ganzen Tag dauerten, in zehn Minuten erledigt.

Besonders bemerkenswert: Nachdem im März 2025 der KI-Agent AI Scientist von Sakana.ai erfolgreich die erste Peer-reviewed Publikation erstellt hatte, publizierte zwei Monate später im Mai 2025 der KI-Agent „Zochi“ ein Peer-Review-Paper auf einer A*-Konferenz. Für Aufmerksamkeit sorgte auch die Stanford University mit der „Agents4Science“ im Oktober 2025 als weltweit erster Konferenz, bei der KI-Autorenschaft nicht nur erlaubt, sondern erforderlich war und KI auch als Gutachter für die eingereichten Beiträge eingesetzt wurde.

Die Implikationen erschüttern den Wissenschaftsbetrieb in seinen Grundfesten. Wenn ein nur wenige Minuten dauernder Durchlauf eines KI-Forschungsagenten die mehrmonatige Arbeit von Forschenden erledigen kann, stellen sich fundamentale Fragen: Wie

Prof. Dr. Doris Weßels ist Wissenschaftliche Leiterin des Zukunftslabors Generative KI an der FuE-Zentrum FH Kiel GmbH. | doris.wessels@haw-kiel.de | Foto: Andreas Diekötter

bewerten wir zukünftig wissenschaftliche Leistung? Wer haftet für fehlerhafte KI-generierte Ergebnisse? Wie sichern wir Forschungsintegrität, wenn KI-Agenten autonom publizieren?

Gleichzeitig entstehen neue Anforderungen an die Infrastruktur. Das „Science Context Protocol (SCP)“ aus Shanghai will ein globales Netz autonomer wissenschaftlicher Agenten ermöglichen, mit über 1.600 integrierten Tools. Mit Frameworks wie „Paper2Agent“ wird eine Zukunft sichtbar, in der eine Methodik aus einem wissenschaftlichen Paper automatisch in einen interaktiven KI-Agenten überführt und zur eigenen Nutzung bereitgestellt werden kann.

Der ehemalige OpenAI-Forschungsleiter Andrej Karpathy hat 2025 mit seinem „Reality Check“ zur Vorsicht gemahnt, weil den KI-Agenten noch Gedächtnis, Multimodalität und Robustheit fehlten. Agentische KI-Systeme sind keine „Zauberstäbe“, sondern neue Prozesse, die hoch risikoreich sind und neue Kontrollmechanismen erfordern. Besonders deutlich werden die Sicherheitsrisiken bei KI-gestützten „virtuellen Mitarbeitern“ mit eigenen Accounts und Passwörtern in Unternehmensnetzwerken. Neue Fragen der Cybersecurity, Zugangskontrollen und Verantwortlichkeit entstehen, nicht nur für die Wirtschaft, sondern auch für Hochschulen und Forschungseinrichtungen.

Viele Forschende haben bereits Erfahrungen gesammelt mit Digital Research Assistant Tools wie Elicit, Consensus oder Undermind, aber mit KI-Agentensystemen treten wir in die nächste KI-Evolutionsstufe ein. Damit ist die Bedeutung von KI-Governance sprunghaft angestiegen: Wann dürfen KI-Agenten eingesetzt werden? Wie wird ihre Nutzung dokumentiert? Welche Qualitätssicherung greift? Die Entwicklung von Richtlinien für KI-generierte Forschung wird im Wissenschaftsmanagement zur Pflichtaufgabe, abgestimmt mit Forschungsförderern wie z.B. der DFG und den Wissenschaftsverlagen. Schließlich erfordert die Entwicklung auch ausreichend finanzielle Mittel für den Kauf von Lizenzen, technischen Plattformen und der kontinuierlichen Qualifizierung der Forschenden und Akteure im Verwaltungsbereich. Aber es wird auch organisatorische Anpassungen geben müssen. Der „AI-

Handlungsempfehlungen

1. Pilotprojekte starten: *Testen Sie ausgewählte Forschungsagenten-Systeme in kontrollierten Umgebungen, um erste Erfahrungen zu sammeln.*
2. KI-Governance-Framework entwickeln: *Erarbeiten Sie mit allen relevanten Stakeholdern Richtlinien für den Einsatz von KI-Agenten in der Forschung mit dem besonderen Fokus auf Transparenzpflichten und Qualitätssicherung.*
3. KI-Kompetenzen fördern: *Schulen Sie kontinuierlich Forschende und Mitarbeitende in der Verwaltung im Umgang mit KI-Forschungsagenten.*
4. Infrastruktur prüfen: *Evaluieren Sie, welche Software-Lizenzen für KI-Systeme sinnvoll sind.*
5. Netzwerke nutzen: *Tauschen Sie sich mit anderen nationalen und internationalen Forschungseinrichtungen aus.*

<https://bit.ly/ki-forschungsagentensysteme>

Literacy-Beauftragte“ könnte zukünftig so selbstverständlich sein wie heute der Datenschutzbeauftragte. Die „Zauberlehrlinge“ sind da. Anders als in Goethes Ballade wird kein Meister kommen und die Geister wieder bannen. Es liegt an uns, die neuen Potenziale zu nutzen und zugleich die Risiken im Bann zu halten. Die neue Führungskompetenz „AI Leadership“ wird zum zentralen Zukunftsprozess im agentischen Zeitalter.

ZWM-Webinar „KI-Agenten in der Forschung“ am 02.12.2026: <https://bit.ly/4dcFJUb>

Die Kolumne „Innovationen im Wissenschaftsmanagement“ wird vom Zentrum für Wissenschaftsmanagement e.V. (ZWM) in Kooperation mit unserem Netzwerk herausgegeben.
www.zwm-speyer.de

