

Warum KI läuft, aber weiterhin nicht liefert

Die wichtigsten KI-Trends für 2026: Vom Einsatz zum echten Mehrwert



Nach den ersten Praxisjahren von KI stellt sich für viele Unternehmen eine zentrale Frage: Wo bleibt der ROI? KI-Agenten arbeiten nicht autonom genug, Modellnutzung leidet unter „Context Rot“ bzw. Qualitätsverlust und klassische Datenbankkonzepte stoßen an Grenzen.

Damit rückt die Frage in den Mittelpunkt, welche Ansätze tatsächlich helfen, KI-Anwendungen aus der Pilotphase in den produktiven Alltag zu bringen – und welche Hürden Unternehmen berücksichtigen müssen, bevor sie skalieren. Welche Rolle dabei konkrete technologische und organisatorische Entwicklungen spielen, zeigen die folgenden sechs Trends.



Autor:
Michael Hunger
VP Product Innovation
Neo4j
<https://neo4j.com>

1. KI-Realitätscheck: Skalierung im Fokus

Die Rückmeldungen aus den Unternehmen sprechen eine klare Sprache: Die meisten KI-Projekte liefern noch nicht das, was man sich von ihnen versprochen hat. Laut einer MIT-Studie [1] liefern 95 % der Pilotprojekte keine messbaren Ergebnisse. Gartner [2] rechnet damit, dass 40 % der Agentic-AI-Projekte bis 2027 scheitern – gebremst durch Kosten, unklaren ROI und ungeklärte Risiken. Hier zeigt sich das GenAI-Paradox: Allgemeine KI-Tools und -Assistenten lassen sich schnell ausrollen, bringen aber schwer messbare ROI-Effekte. Die wirklich wertschöpfenden, vertikal integrierten KI-Systeme dagegen kommen nur mühsam in die Unternehmen.

Von KI-Frust zu sprechen ist dennoch verfrüht. Der Endnutzer-Hype um ChatGPT, Copilot & Co. – und nicht zuletzt die enormen Investitionssummen der Tech-Riesen – haben schlichtweg Erwartungen geweckt, die mit der Realität in Unternehmen kollidieren. Denn hier müssen KI-Systeme tief und sicher in bestehende Prozesse, Datenstrukturen und IT-Landschaften integriert werden. Diese Integration braucht Zeit und Anpassungen – und selbst Künstliche Intelligenz kann diesen Prozess nur bedingt beschleunigen. Zudem ist KI experimentell: Viele prototypische Projekte müssen scheitern, damit sichtbar wird, welche Ansätze langfristig funktionieren. Die eigentliche Skalierung beginnt erst noch.

2. KI-Agenten: Die neuen Trainees

KI-Agenten verdeutlichen dieses Paradox besonders klar. Von autonomen „Agenten-Armeen“, die ganze Abteilungen ersetzen, ist in der Praxis wenig zu sehen. Die meisten Systeme arbeiten verborgen im Hintergrund und übernehmen vor allem zeitintensive Rechercheaufgaben – etwa in Recht, Compliance oder Medizin.

Zwar experimentiert laut McKinsey [3] die Mehrheit der Unternehmen mit AI-Agenten, doch nur 23 % bringen sie überhaupt in einen produktiven Bereich, und in keiner Funktion überschreitet der Anteil der skalierten Agenten derzeit rund 10 %. Das zeigt: Der Nutzen ist hochgradig kontextabhängig, der produktive Einsatz eng begrenzt. Unternehmen müssen zunächst den KI-Hype ausblenden und nüchtern

herausfinden, wo Agenten echte Wirkung entfalten.

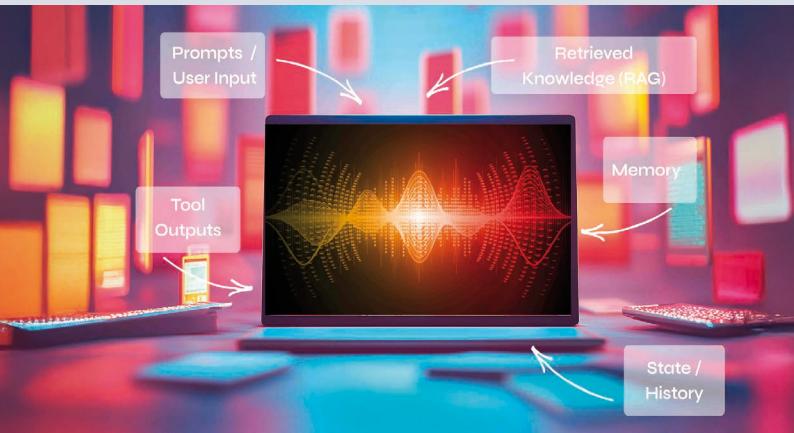
Dieses iterative Vorgehen ist berechtigt – denn KI bleibt unzuverlässig. Meist liegt es nicht an der „fehlenden Intelligenz“ der Modelle. Vielmehr wurden Kontext und Anweisungen nicht klar genug vermittelt, um relevante und zuverlässige Ergebnisse zu garantieren. Für eine funktionsfähige Integration brauchen Agenten daher – ähnlich wie neue Mitarbeitende – eine Art Onboarding: Sie müssen eingearbeitet, informiert, überwacht und regelmäßig korrigiert werden. Da sie probabilistisch arbeiten, liefern sie selbst bei identischen Eingaben nicht zwingend dieselben Ergebnisse. Die Validierung erfordert Tests, Feedback und Review-Prozesse – Aufwand, der nur begrenzt skalierbar ist.

Das wirft eine weitere Frage jenseits der technischen Umsetzung auf: Wie lassen sich KI-Agenten in bestehende Workflows, Teams und die Unternehmenskultur einbinden?

Unternehmen müssen dabei nicht nur in das Training und Setup der KI-Agenten investieren, sondern auch in das ihrer Mitarbeitenden. Sie sind es künftig, die Ergebnisse der KI-Kollegen validieren und die zugrunde liegenden Modellgrenzen verstehen müssen. Dafür gilt es, Arbeitsmodelle neu zu denken, einschließlich klarer Governance, neuer Rollen, flacherer Strukturen und besonders eindeutiger Verantwortlichkeiten.



KI-Agenten stecken oft in der Pilotphase – die Skalierung steht noch aus.
© Neo4j/Canva



Context Engineering: KI-Modelle brauchen kuratierte Infos statt Information Overload © Neo4j/Canva

3. Context Engineering: Informationsarchitektur für die KI:

KI – selbst in agentischen, iterativen Architekturen – ist nur so gut wie der Kontext, den sie erhält. Doch oft bekommt sie zu wenig, zu viel oder zu ungenauen Input. Viele denken beim Prompten an direkte Anweisungen. In realen Anwendungen aber besteht die eigentliche Aufgabe des Systems darin, den Kontext dynamisch so zu gestalten, dass das LLM genau die Informationen erhält, die es für den nächsten Schritt braucht.

LLMs funktionieren in mancher Hinsicht wie das menschliche Arbeitsgedächtnis: Sie behalten den Anfang und das Ende, doch in der Mitte verlieren sie den Faden. Langer Kontext führt zu Fehlern, Reibungsverlusten und sinkender Aufmerksamkeit (Context Rot). Modelle verwirren sich, wenn zu viele oder sehr ähnliche Tools eingesetzt werden (Context Confusion). Oder sie stolpern über widersprüchliche Arbeitsschritte (Context Clash). Obwohl die Modelle theoretisch riesige Mengen an Kontext verarbeiten könnten, zeigt die Praxis: Je mehr man ins Context Window hineinlädt, desto unzuverlässiger sind die Ergebnisse.

In Sachen KI-Kontext gilt daher, wie so oft in der Datenverarbeitung: Qualität statt Quantität. Modelle verfügen nur über eine begrenzte Aufmerksamkeitsspanne – ein Konzept, das als „Attention Budget“ bezeichnet wird. Jeder zusätzliche Kontextbaustein verbraucht davon etwas und verwässert das Wesentliche. Gut kuratiertes Kontext – Context Engineering – entwickelt sich damit zur Grundvoraussetzung für verlässliche KI.

4. Push vs. Pull: Daten auf Abruf statt auf Vorrat

Mit dem Aufkommen von Agentensystemen verändert sich die Art, wie KI auf Informationen zugreift.

Während frühere Ansätze wie Retrieval-Augmented Generation (RAG) nach dem Push-Prinzip funktionierten, setzt sich nun ein Pull-Prinzip durch. Statt Informationen im Vorfeld zu sammeln und dem Modell zuzuschieben, entscheidet die KI selbst, welche Informationen ihr fehlen, und ruft diese gezielt mittels Werkzeug ab. Statt einer Informationslawine entsteht echte Informationsauswahl.

Damit übernimmt KI zunehmend eine organisatorische Rolle: Sie analysiert die Aufgaben, identifiziert Arbeitsschritte und notwendige Informationen und wählt Tools oder Datenquellen aus, die diese Lücken schließen. Sie wird zum Koordinator der Informationsbeschaffung, was den Fähigkeiten des Sprachmodells entgegenkommt. Für Unternehmen heißt das: Denken wie ein Informations-Architekt. Entscheidend ist nicht die Menge, sondern die korrekte Dosis: das Prinzip des „Minimum Viable Context“ (MVC). Die KI soll genau die Informationen erhalten, die sie für den nächsten Schritt braucht – nicht mehr und nicht weniger.

5. Graphen: Navigationsystem für KI-Agenten

Welche Informationen die KI im nächsten Schritt benötigt, hängt stark vom jeweiligen Anwendungsfall ab: mal tiefe, lineare Kontextketten, mal breite, verzweigte Wissensstrukturen, Cluster von relevanten Informationen, oder nur ein einzelner präziser Ausschnitt. Genau an dieser Stelle beginnen klassische Datenstrukturen zu schwächeln. Graphendatenbanken bieten hier einen strukturell anderen Ansatz – von Forrester [4] als Rückgrat für LLMs bezeichnet – um Kontext abzubilden, einzufangen und wieder bereitzustellen.

Gerade im Zusammenspiel mit KI-Agenten rücken Graphen 2026 stärker in den Fokus. Da KI-Systeme

zunehmend selbstständig Entscheidungen, Tools und Prozesse koordinieren, benötigen sie robuste und nachvollziehbare Kontextmodelle. Graphen verknüpfen Wissen, Aktionen und Interaktionen in Echtzeit und machen Agenten dadurch navigierbar, überprüfbar und skalierbar. So entsteht eine semantische Informationsschicht (Knowledge Layer), die nicht nur präzisere Antworten ermöglicht, sondern vor allem einen schafft: Agenten, die verstehen, wo sie stehen, was sie tun, warum sie es tun und welche Folgen der nächste Schritt hat.

6. Datenbank der Zukunft: Adaptiv

Datenbanken und Dateninfrastrukturen entwickeln sich damit zu einem Dreh- und Angelpunkt für den KI-Erfolg. Nach vier Jahren KI-Hype stellt sich dabei immer deutlicher heraus: Während Hardware und Modelle in neue Dimensionen vorstoßen, stecken die Datenbanken darunter noch immer im Denken der 70er fest. KI-Systeme sollen Spitzenleistungen liefern, arbeiten aber auf Architekturen, die nie für sie gebaut wurden. Die zentrale Frage lautet nicht länger, wie Datenbanken verbessert werden, sondern wie eine Datenbank aussieht, die für KI gebaut ist.

Die KI-Datenbank der nächsten Generation könnte zum Beispiel ähnlich wie „Live Code“ funktionieren. Abfragen werden während der Ausführung iterativ neu geschrieben und optimiert – angelehnt an moderne Compiler-Designs wie Just-in-Time-(JIT)-Techniken. Der Ausführungsplan passt sich laufend an Datenverteilungen, Lastmuster und

die jeweils verfügbare Hardware. So entsteht eine permanente Feedbackschleife, in der die Datenbank mit jeder Iteration effizienter wird, selbst wenn Komplexität und Datenmengen wachsen. Genau diese dynamische Architektur bildet das Fundament für die Knowledge Layer, die KI-Agenten künftig brauchen.

Wer schreibt:

Michael Hunger beschäftigt sich seit mehr als 35 Jahren leidenschaftlich mit Softwareentwicklung. In den letzten 13 Jahren hat er an der Open-Source-Graphdatenbank Neo4j in verschiedenen Funktionen gearbeitet, zuletzt als Leiter der Produktinnovation und Entwicklerstrategie. Sein aktueller Fokus liegt auf generativer KI, Cloud-Integrationen und Developer Experience. ▶

Quellen

[1] MIT: The GenAI Divide State of AI in Business 2025; www.artificial-intelligence-news.com/wp-content/uploads/2025/08/ai_report_2025.pdf

[2] Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027; www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027

[3] McKinsey: The state of AI in 2025: Agents, innovation, and transformation; www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

[4] Forrester: AI Wakes The Sleeping Giant: Continuous Improvement Will Finally Fulfill Its Promise, www.forrester.com/blogs/ai-wakes-the-sleeping-giant-continuous-improvement-will-finally-fulfill-its-promise/



Robustes Kontextmodell: Knowledge Graphen bilden den Knowledge Layer moderner KI-Systeme. © Neo4j