

# Von der Bedienung zur Delegation: Warum Agentic AI die Arbeit grundlegend verändert – und was das für den Mittelstand bedeutet

Phronesis AI

2026-02-16

## Der Phasenwechsel

Im Februar 2026 haben 16 autonome KI-Agents in zwei Wochen einen vollständigen C-Compiler geschrieben – 100.000 Zeilen Rust-Code, der den Linux-Kernel kompiliert und 99% einer Torture-Test-Suite besteht. Kosten: 20.000 Dollar. Noch vor zwölf Monaten verloren autonome Agents nach dreißig Minuten den Faden. Vor sechs Monaten galt es als bemerkenswert, wenn ein Agent sieben Stunden durchhielt. Von dreißig Minuten zu zwei Wochen in einem Jahr – das ist keine Trendlinie. Das ist, wie der Analyst es formuliert, ein *Phasenwechsel*.

Solche Geschichten wirken wie Nachrichten aus der Softwareentwicklung. Sind sie auch. Aber der Kern dessen, was hier passiert, betrifft nicht Softwareentwicklung. Er betrifft die Frage, was geschieht, wenn KI nicht mehr minutenweise assistiert, sondern tage- und wochenlang eigenständig arbeitet. Und diese Frage betrifft jedes Unternehmen, dessen Wertschöpfung auf Wissensarbeit beruht.

## Was sich tatsächlich verändert hat

Die öffentliche Diskussion über KI dreht sich meist um Modellgrößen, Benchmarks und Kontextfenster. Das sind die falschen Metriken. Die richtige Metrik ist eine, die kaum jemand kennt: die Fähigkeit eines Modells, Informationen innerhalb seines Kontextfensters wiederzufinden und zu nutzen.

Ein Modell, das eine Million Tokens aufnehmen kann, aber nur in einem von fünf Fällen die richtige Information daraus abrufen kann, ist wie ein Aktenschrank ohne Index. Die Dokumente sind drin, aber ob man findet, was man braucht, ist Zufall. Genau das war der Stand im Januar 2026: Die besten Modelle fanden die Nadel im Heuhaufen in 18 bis 26 Prozent der Fälle.

Opus 4.6, veröffentlicht Anfang Februar, erreicht 76 Prozent bei einer Million Tokens und 93 Prozent bei einem Viertel davon. Das ist der eigentliche Durchbruch: nicht die Menge an Information, die ein Modell aufnehmen kann, sondern die Verlässlichkeit, mit der es sie versteht und nutzt. Es ist der Unterschied zwischen einem Modell, das eine Datei sieht, und einem Modell, das ein ganzes System im Kopf hält – jede Abhängigkeit, jede Wechsel-

wirkung, jede Implikation.

Das ist die Fähigkeit, die einen erfahrenen Mitarbeitenden von einem externen Berater unterscheidet, der die Unterlagen zum ersten Mal liest. Der erfahrene Mitarbeitende weiß, dass eine Änderung im Einkauf Auswirkungen auf die Kalkulation hat, dass die Reklamationsquote mit der Lieferantenauswahl zusammenhängt, dass der Garantieanspruch anders aussieht, wenn die Montage vom Kunden selbst durchgeführt wurde. Nicht weil er es nachschlägt, sondern weil er lange genug im System lebt, um Zusammenhänge intuitiv zu erfassen.

Genau diese holistische Awareness kann ein KI-Agent jetzt leisten – nicht durch jahrelange Erfahrung, sondern durch die Fähigkeit, den gesamten Kontext gleichzeitig zu halten und über ihn hinweg zu denken.

### **Vom Werkzeug zum Gegenüber: Die eigentliche Revolution**

Die meisten Unternehmen nutzen KI heute als bessere Suchmaschine oder als Textgenerator. Man stellt eine Frage, man bekommt eine Antwort. Man gibt einen Prompt, man bekommt einen Entwurf. Das ist das Paradigma der Werkzeugbedienung: Der Mensch formuliert den Prozess, die KI führt einen Schritt aus.

Was gerade entsteht, ist etwas fundamental anderes. Anthropic nennt es *Outcome-Orientiation* – die Beschreibung von Ergebnissen statt von Prozessen. Man erklärt der KI nicht, wie sie die Tabelle bauen soll. Man erklärt, was die Tabelle zeigen muss. Man beschreibt nicht die Schritte einer Reklamationsbearbeitung. Man delegiert: “Bearbeite diese Reklamation.”

Das klingt nach einem graduellen Unterschied. Es ist keiner. Es ist ein Paradigmenwechsel in der Mensch-Maschine-Interaktion, der so grundlegend ist wie der Übergang von der Kommandozeile zur grafischen Benutzeroberfläche in den achtziger Jahren. Damals hörte der Computer auf, eine Maschine zu sein, die man programmiert, und wurde zu einem Werkzeug, das man bedient. Jetzt hört er auf, ein Werkzeug zu sein, das man bedient, und wird zu einem Gegenüber, an das man delegiert.

Die Kompetenz, die dabei zählt, verschiebt sich: weg von der technischen Beherrschung eines Tools, hin zur Klarheit der eigenen Absicht. Wer genau weiß, was er braucht – und es so formulieren kann, wie man es einem kompetenten Mitarbeitenden sagen würde –, der kann jetzt Dinge erreichen, für die bisher ganze Abteilungen zuständig waren.

### **Die Auflösung der Grenze zwischen technisch und nicht-technisch**

Einer der bemerkenswertesten Aspekte der jüngsten Entwicklung: Bei Rakuten, dem japanischen E-Commerce-Konzern, nutzen nicht-technische Mitarbeitende dieselbe KI-Infrastruktur wie die Entwickler, um Features zu bauen und in Produktion zu bringen. Zwei CNBC-Reporterinnen – keine Ingenieurinnen – haben in unter einer Stunde ein funktionierendes Projektmanagement-Tool gebaut, das die Kernfunktionalität eines 5-Milliarden-Dollar-Produkts repliziert.

Das ist nicht die Demokratisierung von Technologie im üblichen Sinne, bei der man ein kompliziertes Werkzeug einfacher bedienbar macht. Das ist die Auflösung der Kategorie selbst. Die Unterscheidung zwischen technischen und nicht-technischen Mitarbeitenden – eine Unterscheidung, die Wissensarbeit, Gehaltsstrukturen und Organigramme seit dreißig Jahren organisiert – löst sich in Monaten auf.

Für den Mittelstand hat das eine spezifische Bedeutung. Hier gibt es selten eine IT-Abteilung mit zwanzig Entwicklern. Hier gibt es Meister, Vertriebsleiter, Sachbearbeiter, Ingenieure – Menschen mit tiefem Fachwissen in ihrem Bereich, aber ohne Programmierkenntnisse. Genau diese Menschen werden durch Agentic AI nicht ersetzt, sondern multipliziert. Ihr Fachwissen – die Fähigkeit zu beurteilen, ob ein Angebot stimmt, ob eine Reklamation berechtigt ist, ob eine Norm korrekt angewendet wurde – wird zum Hebel, der bisher fehlte.

### **Urteilsvermögen als neuer Engpass**

Die gängige Angst lautet: KI ersetzt menschliche Arbeit. Die Realität ist differenzierter und in gewisser Hinsicht anspruchsvoller.

Was KI ersetzt, ist Ausführung. Was sie nicht ersetzt – und was durch sie dramatisch an Wert gewinnt –, ist Urteilsvermögen. Domain-Expertise. Das, was man im Englischen “taste” nennt: das tiefe Verständnis dafür, was ein gutes Ergebnis ausmacht, wie ein korrektes Angebot aussieht, welche Formulierung in einer Reklamationsantwort rechtlich hält und welche nicht.

Die 16 Agents, die den C-Compiler gebaut haben, brauchten niemanden, der für sie Code schreibt. Sie brauchten jemanden, der präzise genug spezifiziert, was ein C-Compiler ist. Das Marketing-Team braucht niemanden mehr, der die Analytics-Plattform bedient – es braucht jemanden, der weiß, welche Kennzahlen relevant sind und warum.

Der Hebel hat sich verschoben: von der Ausführung zum Urteil. Und dieser Hebel multipliziert sich mit der Anzahl der Agents, die eine Person dirigieren kann. AI-native Unternehmen erzielen heute fünf bis sieben Millionen Dollar Umsatz pro Mitarbeitendem – das Fünf- bis Siebenfache dessen, was als “exzellent” in traditionellen Softwareunternehmen gilt. Nicht weil sie bessere Menschen eingestellt haben, sondern weil ihre Menschen Agents orchestrieren statt selbst auszuführen.

### **Management als emergente Eigenschaft**

Ein faszinierendes Ergebnis der jüngsten Entwicklungen: Wenn man mehrere KI-Agents auf eine komplexe Aufgabe ansetzt, organisieren sie sich selbstständig in hierarchische Strukturen. Ein Lead-Agent zerlegt das Projekt in Teilaufgaben, weist sie Spezialisten zu, verfolgt Abhängigkeiten, löst Blockaden. Die Spezialisten kommunizieren nicht nur über den Lead, sondern auch direkt miteinander – Peer-to-Peer-Koordination.

Das ist keine aufgezwungene Struktur. Es ist konvergente Evolution. Hierarchie ist keine kulturelle Konvention, die Menschen KI-Systemen auferlegen. Sie ist eine emergente Ei-

genschaft der Koordination mehrerer intelligenter Akteure bei komplexen Aufgaben. Menschen haben Management erfunden, weil Management das ist, was Intelligenz tut, wenn sie skalieren muss. KI-Agents haben dasselbe entdeckt – aus denselben strukturellen Gründen.

Für die Argumentation zugunsten einer Plattform wie Phronesis ist das zentral: Die Plattform bildet nicht einfach bestehende Arbeitsabläufe digital ab. Sie stellt die Infrastruktur bereit, aus der sich Agents selbst organisieren – mit Skills als definierten Arbeitsabläufen, Tools als Einzelfähigkeiten und Kontexten als abteilungsspezifischem Wissen. Die Plattform ist das, was ein gutes Unternehmen seinen Mitarbeitenden bietet: klare Strukturen, verfügbares Wissen, definierte Prozesse. Der Agent nutzt das alles – aber er entscheidet selbst, was er für die jeweilige Aufgabe braucht.

### **Das Tempo und seine Konsequenzen**

Der Phasenwechsel, der hier stattfindet, ist nicht nur in seiner Richtung bemerkenswert, sondern vor allem in seiner Geschwindigkeit. Die Werkzeuge, die im Januar Stand der Technik waren, sind im Februar eine andere Generation. Der Forscher bei Anthropic, der am C-Compiler-Projekt beteiligt war, formulierte es so: “I did not expect this to be anywhere near possible so early in 2026.”

Diese Geschwindigkeit hat eine paradoxe Konsequenz: Wer sich heute auf ein spezifisches KI-Tool einlässt und es meistert, muss damit rechnen, dass sein Wissen in wenigen Monaten veraltet ist. Das gilt für ChatGPT ebenso wie für Copilot. Jeder, der seinen Workflow auf ein bestimmtes Prompt-Pattern oder eine bestimmte Modellversion optimiert hat, erlebt mit jedem Update eine Entwertung seiner Expertise.

Die Antwort darauf ist nicht, schneller einzelne Tools zu lernen. Die Antwort ist eine Abstraktionsebene: eine Plattform, die das Fachwissen des Unternehmens von der konkreten KI-Technologie entkoppelt. Skills, die definieren, was getan werden soll, bleiben stabil, auch wenn das Modell darunter sich alle drei Monate ändert. Kontexte, die festlegen, welches Wissen in welcher Abteilung relevant ist, überdauern jeden Modellwechsel. Das Unternehmenswissen – Produktdaten, Preislisten, Normen, Richtlinien – bleibt unabhängig davon, ob unter der Haube Opus 4.6, Opus 5 oder etwas ganz anderes arbeitet.

Das ist die architektonische Kernidee von Phronesis: das Unternehmenswissen und die Arbeitsabläufe von der sich rapide verändernden KI-Technologie zu entkoppeln. Die Plattform absorbiert den technologischen Wandel, damit das Unternehmen sich auf das konzentrieren kann, was stabil bleibt: sein Fachwissen, seine Prozesse, sein Urteilsvermögen.

### **Warum der Mittelstand nicht wartet, sondern handelt**

Die Zahlen aus dem Silicon Valley – Cursor mit 5 Millionen Dollar Umsatz pro Mitarbeitendem, McKinsey mit dem Ziel der Agent-Mensch-Parität bis Ende 2026, Amazon-Teams, die sich zu “zwei Menschen plus Agent-Fleet” reorganisieren – das klingt nach einer anderen Welt als das Küchenstudio in Niederbayern oder der Maschinenbauer im Bergischen Land.

Aber der Kern des Arguments trifft den Mittelstand sogar stärker als Großkonzerne. Denn: **Der Mittelstand hat, was KI nicht hat: tiefes, spezifisches Fachwissen.** Die Fähigkeit zu beurteilen, ob ein Küchenangebot korrekt kalkuliert ist. Das Wissen, welche DIN-Norm bei einer bestimmten Bauart greift. Die Erfahrung, wie man eine Reklamation so beantwortet, dass der Kunde zufrieden und das Unternehmen rechtlich abgesichert ist. Dieses Wissen sitzt in den Köpfen von Mitarbeitenden, die oft seit Jahrzehnten im Unternehmen sind – und die zunehmend schwer zu ersetzen sind.

**Was der Mittelstand nicht hat: beliebig skalierbare Arbeitskraft.** Die Fachkräfte fehlen, und sie werden weiter fehlen. Jeder Meister, jede Sachbearbeiterin, jeder Vertriebsleiter verbringt einen erheblichen Teil der Arbeitszeit mit Aufgaben, die zwar Fachwissen voraussetzen, aber im Kern repetitiv sind: Angebote schreiben, Berichte erstellen, Normen nachschlagen, Reklamationen bearbeiten. Nicht weil diese Aufgaben trivial wären – sie sind es nicht –, sondern weil sie einem Muster folgen, das ein Agent erlernen kann.

**Agentic AI multipliziert genau diese Kombination.** Das Fachwissen des Mitarbeitenden wird zum Hebel, die Agent-Infrastruktur zur Multiplikation. Der Meister schreibt nicht mehr jedes Angebot selbst – er delegiert es und prüft das Ergebnis. Die Sachbearbeiterin bearbeitet nicht mehr jede Reklamation von Grund auf – sie delegiert die Standardfälle und konzentriert sich auf die, die echtes Urteilsvermögen erfordern. Der Ingenieur durchsucht nicht mehr stundenlang Normen – er delegiert die Recherche und bewertet das Ergebnis.

Das ist keine Automatisierung im industriellen Sinne, bei der ein Roboter den Menschen ersetzt. Es ist Delegation im eigentlichen Sinne: Ein kompetenter Mitarbeitender übergibt eine Aufgabe an ein kompetentes Gegenüber, das die Abläufe kennt, das Wissen hat und das Ergebnis in der richtigen Form liefert.

### **Die Frage, die sich stellt**

McKinsey empfiehlt seinen eigenen Partnern, bis Ende 2026 die Anzahl der KI-Agents auf die Anzahl der menschlichen Mitarbeitenden zu bringen. Die Frage für den Mittelstand lautet nicht, ob diese Entwicklung kommt. Sie lautet, ob man sie mit generischen Tools wie ChatGPT begleitet – Tools, die das Unternehmenswissen nicht kennen, die keine Skills haben, die nicht wissen, wie ein Angebot in diesem konkreten Unternehmen aussehen muss – oder mit einer Infrastruktur, die auf das eigene Fachwissen, die eigenen Abläufe und die eigenen Qualitätsansprüche zugeschnitten ist.

Die Frage lautet nicht: “Sollen wir KI einsetzen?” Die Frage lautet: “Wie hoch ist unser Verhältnis von Agents zu Mitarbeitenden – und was muss jeder Mitarbeitende exzellent können, damit dieses Verhältnis funktioniert?”

Phronesis ist die Infrastruktur, die diese Frage beantwortbar macht. Nicht als Versprechen, sondern als produktives System: 39 Skills im Einsatz, über 40 Tools verfügbar, Unternehmenswissen vollständig integriert, DSGVO-konform auf eigener Infrastruktur. Nicht irgendwann. Jetzt.

*Basierend auf einer Analyse der jüngsten Entwicklungen in Agentic AI, insbesondere der Ergebnisse*

*von Anthropic's Opus 4.6 (Februar 2026), Rakutens Produktiveinsatz von Agent-Teams und der sich abzeichnenden Reorganisation von Wissensarbeit hin zu Human-Agent-Teams.*