

DREITEILIGE SERIE

Die Organisation nach der Koordination

Unternehmen haben jahrzehntelang zwei verschiedene Dinge unter demselben Wort bezahlt: Koordination und Urteilsvermögen. KI macht das eine fast kostenlos und das andere knapp.

Worum es geht

In wenigen Quartalen haben mehrere DAX-Konzerne dasselbe angekündigt: weniger Führungsebenen, weniger Führungsstellen. Die übliche Erklärung lautet Kostendruck, und sie stimmt. Trotzdem ist diesmal etwas anders als bei allen Sparrunden davor: Der Rückweg ist nicht eingeplant, und für einen Teil der Arbeit gibt es erstmals einen Ersatz.

Genau da beginnt das eigentliche Problem. Denn was in dieser Ebene lag, war nie eine Sache. Es waren zwei, die nie getrennt wurden, weil sie nie getrennt werden mussten.

Koordination

Berichten. Abstimmen. Übersetzen zwischen Ebenen. Kapazität zuteilen. Also Informationstransport.

wird zur Grenzkostenware

Urteilsvermögen

Erkennen, dass etwas schlüssig klingt und trotzdem nicht trägt. Entsteht aus Erfahrung mit Konsequenzen.

wird knapp

Diese Serie beschreibt, was daraus folgt — für Menschen, für Strukturen und für das Risiko, das entsteht, wenn beides gleichzeitig fehlt.

01 Koordination wird billig. Urteilsvermögen wird teuer.

Warum KI Generalisten ohne fachliche Tiefe zuerst ersetzt, und warum der Zulieferpfad für Erfahrung an zwei Stellen gleichzeitig gekappt wird.

02 Ihr mittleres Management entscheidet über Ihre KI-Strategie. Nicht Ihr CTO.

Warum 95 Prozent der KI-Piloten nichts verändern, und was Vorstände in den nächsten zwei Quartalen anders machen müssen.

03 Früher war Führung ohne fachliche Tiefe ineffizient. Jetzt ist sie gefährlich.

Warum KI nicht die Fehlerquote erhöht, sondern den Durchsatz — und aus einer sozialen Entscheidung eine Risikoentscheidung macht.

Zu den Zahlen: Jede Zahl in dieser Serie ist mit Herausgeber, Erhebungszeitraum und Stichprobe belegt; die Quellen stehen am Ende jedes Teils. Wo eine Zahl weich, vorläufig oder umstritten ist, steht es im Text dabei – auch dann, wenn sie unser Argument stützt. Zwei Zahlen sind ausdrücklich als Schätzung markiert.

Zu allem sprechfähig. Für nichts verantwortlich.

Ein Sprachmodell ist der perfekte Generalist ohne Tiefe.
Wer seine Rolle auf derselben Leistung gebaut hat,
konkurriert mit einer Kopie seiner selbst, die schneller ist
und nichts kostet.

Serie „Die Organisation nach der Koordination“, Teil 1



Koordination wird billig. Urteilsvermögen wird teuer.

Warum KI die Organisation neu sortiert und Generalisten ohne fachliche Tiefe zuerst verlieren.

Volkswagen, BMW, Bayer, Deutsche Bahn, BASF, SAP, Evonik. Innerhalb weniger Quartale haben sie alle dasselbe angekündigt: weniger Führungsebenen, weniger Führungsstellen. Personalberater berichten dieselbe Bewegung inzwischen aus dem Mittelstand.

Die übliche Erklärung lautet Kostendruck. Und die stimmt.

Bei Volkswagen geht es um Absatzkrise, China, die Kosten der Elektromobilität und eine Rendite, die bis 2030 wieder zweistellig werden soll. Von KI ist in diesen Begründungen keine Rede. Bayer hat seine Führungsebenen von zwölf auf sechs halbiert, weil ein Organisationsmodell es so vorsah, nicht weil ein Modell Berichte schreibt. Der Umbau hat nicht mit KI begonnen, und er würde auch ohne sie stattfinden.

Trotzdem ist etwas anders als bei allen Sparrunden davor.

Bisher lief es immer gleich: Ein Konzern streicht in der Krise die mittlere Ebene, und drei bis fünf Jahre später ist sie wieder da. Unter anderem Namen, in anderer Farbe im Organigramm, aber wieder da. Weil die Arbeit, die dort lag, ja nicht verschwunden war. Irgendjemand musste weiter berichten, abstimmen, übersetzen und zuteilen.

Diesmal ist der Rückweg nicht eingeplant. Und zum ersten Mal gibt es für einen Teil dieser Arbeit tatsächlich einen Ersatz.

Genau da beginnt das eigentliche Problem. Denn was in dieser Ebene lag, war nie nur eine Sache. Es waren zwei, die nie getrennt wurden, weil sie nie getrennt werden mussten.

Koordination. Und Urteilsvermögen.

Das eine lässt sich ersetzen. Das andere nicht. Und wer beides zusammen streicht, merkt den Unterschied erst, wenn es zu spät ist.

Was Agenten zuerst übernehmen

Ein großer Teil der Managerwoche besteht aus Reporting, Abstimmen, Review und Übersetzen zwischen Ebenen. Also aus Informationstransport. In Praxisberichten wird dieser Anteil auf rund 60 Prozent geschätzt; eine belastbare Primärerhebung dazu existiert nicht. Unstrittig ist die Richtung: Genau diese Tätigkeiten automatisieren sich zuerst, und zwar sauber, günstig und rund um die Uhr.

Gartner hat bereits 2024 prognostiziert, dass bis 2026 jedes fünfte Unternehmen KI nutzen wird, um mehr als die Hälfte seiner mittleren Managementpositionen abzubauen. Ob die Quote so eintrifft, ist offen, und die Prognose beschreibt eine Absicht, keine Messung. Sie beschreibt aber präzise, warum diesmal niemand mit dem Wiederaufbau rechnet.

Damit kippt auch die Führungsspanne. Als Faustregel galten lange sieben bis zehn direkte Berichtslinien. Wie weit sich Spannen bei breitem Agenteneinsatz wirklich öffnen, ist empirisch uneinheitlich; in der Debatte werden deutlich höhere Werte genannt. Der Mechanismus ist unabhängig davon klar: Wer mit Agenten skaliert, braucht weniger Menschen, die Information durchreichen.

Weniger Ebenen. Breitere Spannen.

Und eine Führungsrolle, die inhaltlich etwas ganz anderes ist als vorher. Denn was übrig bleibt, ist nicht das Leichte. Es ist das Schwere: Ausnahmen entscheiden, Konflikte lösen, Menschen entwickeln, gemischte Teams aus Menschen und Agenten steuern.

Das ist eine seltenere Fähigkeit als das Verdichten von Statusberichten. Und eine, die kaum ein Unternehmen je systematisch trainiert hat.

Der Denkfehler über Generalisten

Über kaum etwas wird derzeit widersprüchlicher geschrieben. Die eine Fraktion sagt, KI belohne den vielseitigen Generalisten, der mit den neuen Werkzeugen fünf Rollen gleichzeitig ausfüllt. Die andere sagt, der Markt zahle nur noch für tiefe Spezialisierung.

Beide Lager reden aneinander vorbei, weil sie zwei völlig verschiedene Menschen mit demselben Wort belegen.

Da ist der **Makler**. Sein Wert entsteht daraus, dass er weiß, wen man fragt, wie man es formuliert und an welcher Stelle es üblicherweise hakt. Er kennt zu jedem Thema den Rahmen, aber zu keinem den Boden. Breite ohne Tiefe.

Und da ist der **Praktiker mit Reichweite**. Jemand, der in mindestens einem Feld einmal richtig tief war, lange genug, um dort Fehler gemacht und ausgebadet zu haben. Von dort aus hat er Breite aufgebaut.

KI frisst den ersten Typ. Nicht irgendwann. Jetzt.

Warum?

Weil die Kernleistung des Maklers darin besteht, zu jedem Thema plausible Anschlussfähigkeit herzustellen. Das ist exakt die Disziplin, in der ein Sprachmodell zu Grenzkosten nahe null unschlagbar ist. Ein Modell ist der perfekte Generalist ohne Tiefe. Es ist zu allem sprechfähig und zu nichts verantwortlich.

Wer seine Rolle auf derselben Leistung aufgebaut hat, konkurriert nicht mit einem Werkzeug. Er konkurriert mit einer Kopie seiner selbst, die schneller ist und nichts kostet.

Urteilsvermögen ist die Fähigkeit, richtig von plausibel zu unterscheiden

Hier liegt der eigentliche Punkt.

KI-Systeme produzieren selten offensichtlich falsche Antworten. Sie produzieren überzeugende. Der Unterschied zwischen einer belastbaren und einer gefährlichen Antwort ist im Output meist nicht sichtbar.

Er ist nur für jemanden sichtbar, der die Sache selbst schon einmal gemacht hat.

Übersicht sagt: Das klingt schlüssig. Urteilsvermögen sagt: Das klingt schlüssig, und trotzdem stimmt es hier nicht.

Diese zweite Fähigkeit entsteht nicht aus Belesenheit. Sie entsteht aus Erfahrung mit Konsequenzen. Wer nie eine Entscheidung getragen hat, die schiefging, hat kein Material, an dem er sich kalibrieren könnte. Er kann Ergebnisse nur beurteilen, indem er prüft, ob sie sich gut anfühlen. Das ist keine Prüfung. Das ist Geschmack.

Genau deshalb kommt der Begriff des Player-Coach gerade zurück: Wer führt, muss tief im Stoff stecken, so wie in einer Kanzlei niemand ein Team führt, der das Fallrecht nicht beherrscht. Auch die Kopplung von Beförderung und Teamgröße löst sich in ersten Unternehmen auf. Nicht wer viele führt, steigt auf, sondern wer mit wenig viel erreicht.

Führung wird wieder inhaltlich. Nicht, weil das Menschliche an Führung unwichtiger würde.

Sondern weil die inhaltsleere Variante von Führung automatisierbar geworden ist.

Und es gibt eine harte Kehrseite: Wer die neuen Werkzeuge nicht selbst beherrscht, wird zum Engpass für sein gesamtes Team. Der Kontrast zwischen zwei Erhebungen ist dabei aufschlussreich. In einer Studie von Stifterverband und McKinsey gaben 79 Prozent von rund 1.000 deutschen Führungskräften an, ihren Beschäftigten fehlten grundlegende KI-Kompetenzen. Der Finger zeigt nach unten. In der globalen Roland-Berger-Befragung halten dagegen 59 Prozent die eigene Führungsebene für nicht ausreichend vorbereitet.

Die Lücke sitzt also nicht nur unten. Sie wird nur meistens dort gesucht.

Das Problem, über das kaum jemand spricht

Tiefe hat einen Zulieferpfad. Und der wird gerade an zwei Stellen gleichzeitig gekappt.

Erstens unten. Das Stanford Digital Economy Lab wertet gemeinsam mit ADP monatliche Gehaltsabrechnungen von Millionen US-Beschäftigten aus. In der im August 2026 aktualisierten Fassung liegt die Beschäftigung der 22- bis 25-Jährigen in stark KI-exponierten Berufen 19 Prozent unter dem Niveau, das sie erreicht hätte, wäre sie mit weniger exponierten Altersgenossen mitgewachsen. Bei erfahrenen Beschäftigten gibt es keine vergleichbare Lücke. Zwei Einschränkungen gehören dazu: Die Autoren verstehen das ausdrücklich als beschreibenden Frühindikator, nicht als Kausalbeweis. Und der Effekt läuft über weniger Einstellungen, nicht über Entlassungen.

In Deutschland lag der Anteil ausgeschriebener Einstiegsstellen 2025 rund 42 Prozent unter dem Durchschnitt der fünf Vorjahre, ausgewertet über 4,6 Millionen Stellenanzeigen bei Stepstone. Das IAB führt diesen Einbruch primär auf die Konjunktur zurück, nicht auf KI. Die strukturelle Ursache liefert die Randstad-ifo-HR-Befragung aus dem vierten Quartal 2025: In 14 Prozent der Unternehmen übernimmt KI heute schon Aufgaben, die üblicherweise Berufseinsteigern zufielen. Für die kommenden drei Jahre planen das 40 Prozent, bei Großunternehmen 63 Prozent.

Zweitens in der Mitte. Die mittlere Ebene war nie nur Verwaltung. Sie war die Strecke, auf der Menschen lernen, Verantwortung zu tragen, bevor die Beträge groß werden. Wer sie 2026 streicht, sucht 2028 vergeblich nach Nachfolgern mit Führungserfahrung.

Das Ergebnis ist ein Markt, der etwas sucht, dessen Produktion er selbst eingestellt hat: Junioren mit Seniorprofil.

Sie werden nicht auftauchen. Man kann Urteilsvermögen nicht einkaufen, wenn niemand mehr die Jahre bereitstellt, in denen es entsteht.

Was Führung daraus machen muss

Vier Konsequenzen, die aus meiner Sicht über die nächsten Jahre entscheiden.

Aufgaben neu schneiden, nicht nur Stellen streichen. Wer Ebenen kürzt, ohne vorher zu klären, welche Entscheidungen dort tatsächlich getroffen wurden, verlagert diese Entscheidungen nicht. Er lässt sie ausfallen. Sichtbar wird das erst zwei Quartale später, in Form von Fehlern, die niemand mehr rechtzeitig gesehen hat.

Einstiegsrollen auf Bewertung umbauen. Die alte Junior-Rolle bestand aus Zuarbeit. Die neue muss aus Prüfung bestehen: Ergebnisse einer KI bewerten, Fehler finden, erklären können, warum etwas nicht trägt. Anspruchsvoller als früher. Aber es ist die einzige Version von Einstieg, die in fünf Jahren eine besetzte Senior-Bank hervorbringt.

Führungsspannen ehrlich rechnen. Gallup weist für 2025 einen deutlichen Rückgang des Engagements von Führungskräften aus, von 31 Prozent im Jahr 2022 auf 22 Prozent, während die Teams, die sie führen, größer werden. Bemerkenswert dabei: Der stärkste Treiber echter KI-Wirkung im Team ist die aktive Unterstützung durch den direkten Vorgesetzten. Im Gallup Engagement Index Deutschland berichten nur 21 Prozent der Beschäftigten in KI-nutzenden Organisationen, dass ihre Führungskraft den KI-Einsatz aktiv unterstützt und vorantreibt. Wer die Spanne verdoppelt und gleichzeitig KI-Wirkung erwartet, bekommt weder das eine noch das andere.

Entscheidungslogik explizit machen. Das ist der unterschätzte Teil. Organisationen dokumentieren Ergebnisse: Prozesse, Richtlinien, Handbücher. Was sie nicht festhalten, ist das Warum dahinter. Welche Option wurde verworfen? Woran hat man erkannt, dass ein Fall eine Ausnahme ist? Genau dieses Wissen sitzt in Köpfen, und genau dieses Wissen brauchen sowohl neue Führungskräfte als auch Agenten, um mehr zu sein als schnell und generisch. Wir bei Omnora arbeiten seit Jahren an dieser Stelle: Erfahrungswissen nicht archivieren, sondern in Wirkung bringen.

Die eigentliche Frage

Die Organisation der nächsten Jahre wird flacher. Das ist entschieden.

Ob sie dabei auch klüger wird, entscheidet etwas anderes. Nämlich, ob in ihr noch genug Menschen sitzen, die einen guten von einem gut klingenden Vorschlag unterscheiden können.

Breite war lange eine Währung. Man hat sie über Jahre verdient, und sie hat sich ausgezahlt.

Heute bekommt man sie im Monatsabo. Zwanzig Euro, ohne die Jahre.

Genau so wird eine Währung entwertet. Wer viel davon besitzt, ist deshalb nicht schlechter geworden. Er bekommt nur nichts mehr dafür.

Was knapp wird, ist die Fähigkeit, eine Sache so gut zu verstehen, dass man ihr widersprechen kann.

Welche Entscheidungen in Ihrem Unternehmen kann heute niemand mehr fachlich prüfen, sondern nur noch weiterreichen?

In Teil 2: Warum diese Frage nicht in der Personalentwicklung landet, sondern im Organigramm.

Quellen

- Gartner, Pressemitteilung vom 22.10.2024: Prognose, dass bis 2026 20 % der Organisationen KI zum Abflachen der Struktur nutzen und mehr als die Hälfte der bestehenden mittleren Managementpositionen streichen
- Stanford Digital Economy Lab, Brynjolfsson / Chandar / Chen, „Canaries in the Coal Mine?“, Fassung August 2026, ADP-Daten bis Juni 2026: 19 % relative Beschäftigungslücke bei 22- bis 25-Jährigen in KI-exponierten Berufen
- Stepstone, Auswertung von 4,6 Millionen Stellenanzeigen: Anteil ausgeschriebener Einstiegsstellen 2025 rund 42 % unter dem Fünfjahresdurchschnitt
- Randstad-ifo-HR-Befragung Q4/2025, veröffentlicht am 19.01.2026, 500 bis 1.000 Personalverantwortliche: 14 % heute, 40 % geplant in drei Jahren, 63 % bei Großunternehmen
- Gallup Engagement Index Deutschland, veröffentlicht März 2026, 1.700 Befragte: 21 % erleben aktive Unterstützung der KI-Nutzung durch die Führungskraft
- Gallup, State of the Global Workplace 2026: Manager-Engagement von 31 % (2022) auf 22 % (2025)
- Stifterverband / McKinsey, „KI-Kompetenzen in deutschen Unternehmen“, Januar 2025, rund 1.005 befragte Führungskräfte
- Roland Berger, „The AI-First Organization“, 472 Führungskräfte, Erhebung Dezember 2025 bis April 2026
- Kienbaum: Faustregel zur Führungsspanne von sieben bis zehn direkten Berichtslinien
- Wirtschaftspresse 2025 und 2026: Abbau von Führungsstellen bei VW, BMW, Bayer, Deutsche Bahn, BASF, SAP, Evonik

8,7×

so häufig verändert KI die Arbeitsweise einer Organisation

wenn Beschäftigte ihrer Führungskraft eine aktive Unterstützung der KI-Nutzung bescheinigen. In Deutschland erleben das nur 21 Prozent.

Gallup, Global Indicator: Artificial Intelligence, Q1/2026 ·
Gallup Engagement Index Deutschland, März 2026



Ihr mittleres Management entscheidet über Ihre KI-Strategie. Nicht Ihr CTO.

Warum der Umbau der Organisation eine Frage von Quartalen ist. Und warum Abbau nicht Umbau ist.

Im ersten Teil ging es um die Fähigkeit, richtig von plausibel zu unterscheiden, und darum, warum sie aus fachlicher Tiefe entsteht. Das klingt nach einem Thema für die Personalentwicklung.

Ist es nicht. Es ist ein Strukturthema, und es hat eine Frist.

Die meistzitierte Zahl des letzten Jahres stammt vom MIT: Rund 95 Prozent der untersuchten generativen KI-Pilotprojekte lieferten keinen messbaren Effekt auf die Gewinn- und Verlustrechnung.

Die Zahl ist umstritten, und das zu Recht. Der Bericht aus dem Project NANDA des MIT ist vorläufig, nicht peer-reviewed, beruht auf einer kleinen Interviewbasis und definiert Erfolg eng über messbare Ergebniseffekte nach etwa sechs Monaten. Die Kritik trifft die Präzision der Zahl. Sie trifft nicht deren Richtung.

Denn die interessante Erkenntnis ist nicht die Zahl. Es ist die Begründung. Die Projekte scheiterten nicht an den Modellen. Sie scheiterten daran, dass hinterher kein einziger Prozess anders lief als vorher. Das MIT nennt es die Lernlücke.

Die Roland-Berger-Studie „The AI-First Organization“, für die zwischen Dezember 2025 und April 2026 weltweit 472 Führungskräfte befragt wurden, macht dasselbe Bild an anderer Stelle sichtbar: 62 Prozent erwarten große oder radikale Veränderungen ihres Betriebsmodells. Nur 38 Prozent haben mit der Transformation überhaupt begonnen. 59 Prozent sagen offen, dass ihre Führung darauf nicht ausreichend vorbereitet ist. Und die größte genannte Hürde ist nicht Technologie, sondern fehlende KI-Fähigkeiten, genannt von 49 Prozent.

Zwischen Erwartung und Umbau liegt also eine Lücke von 24 Prozentpunkten. Diese Lücke hat eine Adresse.

Wer wirklich darüber entscheidet, ob KI ankommt

Gallup hat etwas gemessen, das in der Debatte um Rechenzentren und Modellgrößen völlig untergeht: US-Beschäftigte in KI-einführenden Organisationen, die ihrer Führungskraft eine aktive Unterstützung der KI-Nutzung bescheinigen, stimmen 8,7-mal so häufig voll zu, dass KI die Arbeitsweise in ihrer Organisation verändert hat.

8,7-mal. Nicht 20 Prozent besser. Fast das Neunfache.

Im Gallup Engagement Index Deutschland berichten nur 21 Prozent der Beschäftigten in KI-nutzenden Organisationen, dass ihre Führungskraft den Einsatz aktiv unterstützt und vorantreibt. In den USA ist es weniger als ein Drittel.

Und dazu passt die zweite Zahl: 65 Prozent der US-Beschäftigten in KI-nutzenden Organisationen berichten von persönlichen Produktivitätsgewinnen. Aber nur 12 Prozent sagen, dass KI die Arbeitsweise ihrer Organisation tatsächlich verändert hat.

Genau in dieser Lücke sitzt das mittlere Management.

Der Effekt ist multiplikativ, nicht additiv. Eine Führungskraft ohne KI-Kompetenz spart nicht einfach ihren eigenen Produktivitätsgewinn ein. Sie neutralisiert den ihres gesamten Teams. Was sie nicht versteht, kann sie nicht priorisieren, nicht verteidigen und im Zweifel nicht genehmigen.

Anders gesagt: Das mittlere Management ist der Verstärker jeder KI-Einführung. In deutschen Unternehmen ist dieser Verstärker bei etwa einem Fünftel der Teams eingeschaltet.

Warum das kein Charakterproblem ist

Es wäre bequem, hier von Beharrungskräften zu sprechen. Das trifft die Sache nicht.

Der Wert des mittleren Managements bestand jahrzehntelang aus Koordination: Informationen verdichten, zwischen Ebenen übersetzen, Kapazität zuteilen, Status berichten. Das war eine reale, knappe und gut bezahlte Leistung.

Jetzt bittet man dieselben Menschen, die Einführung genau der Technologie zu beschleunigen, die diese Leistung überflüssig macht. Ohne ihnen vorher zu sagen, worin ihre neue Leistung bestehen soll.

Wer das tut und dann über mangelnde Veränderungsbereitschaft klagt, hat ein Designproblem und nennt es ein Haltungsproblem.

Die Frage ist also nicht, wie man diese Ebene motiviert. Die Frage ist, wofür man sie ab morgen bezahlt.

Die Antwort heißt Builder

Nicht Programmierer. Builder.

Damit meine ich die Fähigkeit, den Abstand zwischen einem Problem und einer funktionierenden Sache selbst zu schließen. Ohne Projekt. Ohne Lastenheft. Ohne den Umweg über drei Abteilungen und ein Quartal.

Warum das strukturell alles verändert:

Die klassische Organisation ist eine Rationierungsmaschine. Briefing, Spezifikation, Priorisierung, Freigabe, Backlog, Roadmap: Jeder dieser Schritte existiert, weil Umsetzungskapazität knapp und teuer war. Man musste sie zuteilen, also brauchte man eine Ebene, die zuteilt.

Diese Ressource ist nicht mehr knapp.

Eine Rationierungsmaschine für ein Gut, das im Überfluss vorhanden ist, erzeugt keinen Wert. Sie erzeugt nur noch Reibung. Und Reibung fühlt sich innen wie Sorgfalt an.

Deshalb greift die Rechnung „weniger Ebenen gleich mehr Tempo“ zu kurz. Tempo entsteht nicht dadurch, dass man Genehmiger streicht. Es entsteht dadurch, dass die Menschen, die eine Idee haben, sie auch bauen können.

Eine Führungskraft, die die neuen Werkzeuge nicht selbst beherrscht, ist in dieser Logik nicht langsam. Sie ist ein Engpass für alle, die schneller wären.

Schnell umbauen ist nicht schnell abbauen

Hier liegt der teuerste Denkfehler des Jahres.

Viele Unternehmen streichen gerade Ebenen und nennen das Transformation. Es ist eine Kostenmaßnahme. Sie wirkt sofort in der Bilanz und verändert an der Arbeitsweise nichts, weil die verbliebenen Führungskräfte exakt dieselbe Rolle ausüben wie vorher, nur für doppelt so viele Menschen.

Das Ergebnis ist die schlechteste aller Varianten: weniger Koordinationskapazität und trotzdem keine neue Kompetenz.

Umbau bedeutet etwas anderes. Er bedeutet, den Zuschnitt von Rollen zu ändern, bevor man Stellen antastet. Welche Entscheidungen werden hier eigentlich getroffen? Welche davon kann ein Agent vorbereiten? Welche darf er nie treffen? Was bleibt, und wer muss dafür was können?

Diese Arbeit ist unbequemer als eine Abbauliste. Aber sie ist die einzige, die nach zwei Jahren noch trägt.

Warum Tempo hier wirklich zählt

Ich bin generell vorsichtig mit Dringlichkeitsrhetorik. In diesem Punkt halte ich sie für berechtigt, aus zwei Gründen.

Strukturschulden verzinsen sich. Jedes Quartal, in dem eine Organisation in der alten Struktur weiterarbeitet, trainiert sie ihre Leute im alten Verhalten. Prozesse zementieren sich, Erwartungen verfestigen sich, Karrierewege bilden ab, was gestern belohnt wurde. Der Umbau wird nicht linear teurer. Er wird schwerer.

Die eigentliche Lernkurve kann man nicht kaufen. Berater und Anbieter liefern Modelle, Werkzeuge und Referenzarchitekturen. Was sie nicht liefern können, ist das Wissen darüber, wo KI in Ihrem Geschäft, mit Ihren Fällen und Ihren Ausnahmen zuverlässig scheitert. Dieses Wissen entsteht nur durch eigenes Tun, und es entsteht langsam. Wer heute damit anfängt, hat in zwei Jahren etwas, das nicht nachkaufbar ist.

Alles andere an KI kann man kaufen. Das nicht.

Vier Konsequenzen für die nächsten zwei Quartale

Nachweis statt Schulung. KI-Schulungen für Führungskräfte erzeugen Teilnahmequoten, keine Kompetenz. Die härtere Anforderung: Jede Führungskraft hat im letzten Quartal mindestens einen realen Prozess in ihrem eigenen Verantwortungsbereich selbst automatisiert oder umgebaut. Selbst. Nicht beauftragt.

Führung an gebauten Ergebnissen messen, nicht an Teamgröße. Solange Aufstieg an der Anzahl der Berichtslinien hängt, optimiert jeder rational auf mehr Menschen statt auf mehr Wirkung. Erste Unternehmen haben diese Kopplung inzwischen gelöst. Das ist keine Symbolpolitik, sondern die Stellschraube, die das Verhalten am direktesten dreht.

Umbaufähigkeit statt Umbau. Die eine große Reorganisation ist das falsche Format. Was Sie brauchen, ist eine Organisation, die ihren Zuschnitt regelmäßig überprüft, weil sich die Grenze zwischen dem, was Menschen und dem, was Systeme tun, alle paar Monate verschiebt. Wer alle fünf Jahre umbaut, baut immer zum falschen Zeitpunkt.

Kontext bereitstellen, bevor Sie Tempo verlangen. Ein Builder ohne Zugriff auf die Entscheidungslogik seines Unternehmens baut schnell in die falsche Richtung. Geschwindigkeit ohne Urteilsvermögen ist nur schnelleres Danebenliegen. Das MIT benennt genau das als Kern der Lernlücke: Werkzeuge, die den Kontext der Organisation nicht behalten und sich nicht an ihn anpassen. Wir bei Omnora arbeiten an dieser Stelle, weil das Erfahrungswissen, das Agenten handlungsfähig macht, in Köpfen sitzt und nicht in Ordnern.

Die Frage, die ich Vorständen stellen würde

Nicht: Wie hoch ist Ihr KI-Budget?

Nicht: Wie viele Piloten laufen?

Sondern:

Wie viele Ihrer Führungskräfte haben in den letzten 90 Tagen selbst etwas gebaut, das heute jemand benutzt?

Diese Zahl ist Ihre eigentliche KI-Strategie. Alles andere ist Absichtserklärung.

Quellen

- MIT Project NANDA, „The GenAI Divide: State of AI in Business 2025“: 95 % der untersuchten GenAI-Piloten ohne messbaren Ergebniseffekt, Ursache ist laut Bericht die Lernlücke bei Werkzeugen und Organisationen. Vorläufiger Bericht, nicht peer-reviewed, Angaben zur Stichprobe schwanken je nach Darstellung
- Roland Berger, „The AI-First Organization – Turning AI power into enterprise performance“, veröffentlicht am 06.07.2026. 472 Führungskräfte aus großen und mittelständischen Unternehmen, Erhebung Dezember 2025 bis April 2026, DACH-Anteil 35 %: 62 % erwarten große oder radikale Veränderungen des Betriebsmodells, 38 % haben begonnen, 59 % halten die Führung für nicht ausreichend vorbereitet, 42 % zweifeln an den Governance-Strukturen, 49 % nennen fehlende KI-Fähigkeiten als größte Hürde
- Gallup, „Global Indicator: Artificial Intelligence“ sowie State of the Global Workplace 2026, Basis US-Workforce-Befragung Q1/2026: Faktor 8,7 bei voller Zustimmung zur Aussage, dass KI die Arbeitsweise in der Organisation verändert hat; 65 % persönliche Produktivitätsgewinne gegenüber 12 % organisationaler Veränderung
- Gallup Engagement Index Deutschland, veröffentlicht März 2026, 1.700 repräsentativ Befragte: 21 % erleben aktive Unterstützung der KI-Nutzung durch die Führungskraft
- Stifterverband / McKinsey, „KI-Kompetenzen in deutschen Unternehmen“, Januar 2025, rund 1.005 Führungskräfte: 79 % sagen, ihren Beschäftigten fehlten grundlegende KI-Kompetenzen

Die Fehlerquote ändert sich nicht. Der Durchsatz schon.

Früher drei substanzielle Entscheidungen pro Woche, jede mit Vorlauf. Heute Dutzende fertig formulierter Ergebnisse pro Tag. Und die einzige verbleibende Aufgabe ist Prüfung.

Serie „Die Organisation nach der Koordination“, Teil 3



**Früher war Führung ohne fachliche Tiefe ineffizient.
Jetzt ist sie gefährlich.**

Warum KI die Kosten fehlenden Urteilsvermögens nicht erhöht, sondern vervielfacht.

In jeder Organisation, die ich kenne, gab es Führungskräfte, die inhaltlich nicht getragen haben.

Nicht faul. Nicht böswillig. Meist umgänglich, oft loyal, manchmal seit zwanzig Jahren im Haus. Nur ohne fachliches Fundament in dem Bereich, für den sie verantwortlich waren.

Und es ging trotzdem. Es gab einen starken Stellvertreter. Es gab ein Team, das die Lücke stillschweigend füllte. Es gab Prozesse, die Fehler abfingen, bevor sie Geld kosteten.

Das war teuer. Aber es war verkraftbar.

Dieser Puffer ist gerade verschwunden. Und zwar nicht ein bisschen.

Was sich rechnerisch geändert hat

Die Fehlerquote eines Menschen ändert sich durch KI nicht. Sein Durchsatz schon.

Früher traf eine Führungskraft in der Woche vielleicht drei oder vier substanzielle Entscheidungen. Jede hatte Vorlauf. Es gab eine Vorlage, Rückfragen, ein Gremium, Zeit. Zeit, in der auffallen konnte, dass die Sache nicht durchdacht war.

Heute liegen an derselben Stelle Dutzende fertig formulierter Ergebnisse pro Tag. Argumentiert, mit Zahlen, in gutem Deutsch, in der richtigen Gliederung.

Und die einzige verbleibende Aufgabe an dieser Schnittstelle ist Prüfung.

Also genau die Fähigkeit, die dieser Person fehlt.

Der Punkt ist nicht, dass sie falsch entscheidet. Der Punkt ist, dass sie nicht mehr nicht entscheiden kann. Jedes Durchwinken ist eine Entscheidung. Wer fünfzig Vorlagen am Tag freigibt, ohne eine davon beurteilen zu können, trifft fünfzig Entscheidungen, die er nicht getroffen hat.

Der Befund, der das ungemütlich konkret macht

Das Forschungslabor METR hat 2025 eine randomisierte kontrollierte Studie durchgeführt. 16 erfahrene Open-Source-Entwickler, 246 echte Aufgaben, in Codebasen, die sie im Schnitt seit fünf Jahren kannten.

Vorher schätzten sie, mit KI rund 24 Prozent schneller zu sein. Hinterher glaubten sie, etwa 20 Prozent schneller gewesen zu sein.

Gemessen brauchten sie 19 Prozent länger.

Diese Zahl ist berühmt geworden. Der eigentlich wichtige Befund steht daneben: Die Entwickler akzeptierten weniger als 44 Prozent des vorgeschlagenen Codes. Sie haben also mehr als die Hälfte weggeworfen. Und sie brauchten rund 9 Prozent ihrer Arbeitszeit dafür, den Output zu prüfen und zu korrigieren.

Das ist Urteilsvermögen im Betrieb. Sichtbar als Kostenstelle.

Jetzt die Frage, um die es in diesem Text geht.

Was passiert, wenn an derselben Stelle jemand sitzt, der die weggeworfenen 56 Prozent nicht erkennt?

Er ist schnell. Er fühlt sich produktiv. Er liefert pünktlich. Und niemand merkt etwas, bis es teuer wird.

Beachten Sie dabei den zweiten Teil des METR-Befundes: Selbst Experten haben ihre eigene Geschwindigkeit um fast 40 Prozentpunkte falsch eingeschätzt. Wer über Tiefe verfügt, irrt sich über sich selbst. Wer keine hat, hat nicht einmal das Material für den Irrtum.

Warum das in Führungsrollen eskaliert

Ein Sachbearbeiter, der schlechten KI-Output übernimmt, produziert einen fehlerhaften Vorgang.

Eine Führungskraft, die dasselbe tut, produziert einen Standard.

Sie gibt frei. Sie priorisiert. Sie eskaliert nicht. Ihr Urteil ist die letzte Instanz, an der ein plausibler Fehler noch auffallen könnte, bevor er nach außen geht oder in ein System wandert. Fällt diese Instanz aus, wird aus dem Fehler eine Entscheidung, und aus der Entscheidung ein Präzedenzfall.

Dazu kommt ein zweiter Effekt, der langsamer wirkt und länger nachhallt: Man kann nicht beibringen, was man nicht beurteilen kann. Ein Team lernt von seiner Führungskraft, was als gut genug durchgeht. Wenn die Antwort lautet „alles, was plausibel aussieht“, dann lernt es das.

So entsteht in zwei Jahren eine Abteilung, in der niemand mehr Nein sagen kann.

Langsamkeit war eine Sicherung

In Teil 2 habe ich beschrieben, dass die klassische Organisation eine Rationierungsmaschine ist: Briefing, Spezifikation, Freigabe, Backlog. Alles Mechanismen, die eine knappe Umsetzungskapazität zuteilen. Und alles Reibung, sobald die Kapazität nicht mehr knapp ist.

Das bleibt richtig. Es ist aber nur die halbe Beschreibung.

Diese Schleifen waren nicht nur Rationierung. Sie waren auch Redundanz. Vier Augenpaare an vier Stellen. Wenn drei nichts merkten, merkte das vierte etwas.

Wer Ebenen streicht und gleichzeitig den Durchsatz vervielfacht, entfernt beides auf einmal: die Reibung und die Absicherung.

Deshalb ist Umbau ohne Kompetenzaufbau nicht neutral. Er ist riskant. Und deshalb ist auch das Gegenteil kein Ausweg: Wer langsam bleibt, um sicher zu sein, verliert die Sicherheit trotzdem, weil die Menschen in der Kette dieselben bleiben.

Der unbequeme Teil

Eine Führungskraft ohne fachliche Tiefe zu halten war lange eine soziale Entscheidung. Sie hatte einen Preis, und der wurde in Effizienz bezahlt.

Heute ist es eine Risikoentscheidung. Der Preis wird in Qualität, Haftung und Reputation bezahlt, und er fällt nicht in dem Quartal an, in dem die Entscheidung getroffen wurde.

Das ist ausdrücklich kein Aufruf zum Aussortieren. Die meisten dieser Menschen sind nicht ungeeignet. Sie sind falsch eingesetzt oder nie ausgebildet worden, oft von genau der Organisation, die jetzt enttäuscht ist. Wer zwanzig Jahre für Koordination befördert wurde, hat keine fachliche Tiefe versäumt. Er hat getan, was belohnt wurde.

Aber eine Entscheidung muss fallen, und sie muss ausgesprochen werden. Entweder Tiefe aufbauen, mit Zeit, Budget und einer realistischen Frist. Oder die Rolle ändern, ohne die Person zu beschädigen.

Was nicht mehr funktioniert, ist die dritte Variante, die in den meisten Unternehmen gerade läuft: nichts tun und hoffen, dass die Organisation es wieder auffängt.

Sie fängt es nicht mehr auf. Genau das war die Leistung der Ebene, die gerade gestrichen wird.

Drei Dinge, die Sie in den nächsten Wochen prüfen können

Erstellen Sie die Liste, die es nirgends gibt. An welchen Stellen im Unternehmen liegt heute eine Freigabe bei jemandem, der den Inhalt fachlich nicht beurteilen kann? Diese Liste existiert in keinem Organigramm und in keinem Risikoregister. Sie zu erstellen dauert zwei Wochen und verändert jede weitere Diskussion über KI.

Messen Sie die Ablehnungsquote, nicht die Nutzungsquote. Fast alle Unternehmen berichten heute, wie viele Mitarbeitende KI nutzen. Fast keines berichtet, wie oft KI-Output zurückgewiesen wird. Bei METR lag diese Quote über 56 Prozent, unter Experten, in vertrautem Terrain. Wenn sie in Ihrem Unternehmen bei nahe null liegt, ist das kein Qualitätssignal. Es ist ein Warnsignal.

Machen Sie die Sicherung explizit. Welche Entscheidungen brauchen zwingend ein zweites, fachlich kompetentes Augenpaar, und welche nicht? Die Antwort darauf steht in keinem Prozesshandbuch, weil sie nie ausgesprochen werden musste. Sie sitzt in den Köpfen der Erfahrenen. Genau dieses Wissen aus Köpfen in etwas Nutzbares zu übersetzen, ist die Arbeit, an der wir bei Omnora seit Jahren sitzen. Ohne sie prüft niemand, weder Mensch noch Agent, weil niemand weiß, woran man erkennt, dass etwas nicht trägt.

Zum Schluss

Es hat immer Menschen in Führungsrollen gegeben, die inhaltlich nicht getragen haben. Die Organisation hat sie mitgetragen, weil sie langsam genug war, um es sich zu leisten.

Diese Nachsicht war nie eine Tugend. Sie war ein Nebenprodukt von Trägheit.

Die Trägheit ist weg.

Wo in Ihrem Unternehmen hängt heute eine Freigabe an jemandem, der den Inhalt nicht beurteilen kann?

Quellen

- METR, randomisierte kontrollierte Studie, erstes Halbjahr 2025: 16 erfahrene Open-Source-Entwickler, 246 reale Aufgaben in eigenen Repositories. Erwartung vorab plus 24 % Geschwindigkeit, Selbsteinschätzung danach plus 20 %, gemessen 19 % längere Bearbeitungszeit. Weniger als 44 % des vorgeschlagenen Codes akzeptiert, rund 9 % der Arbeitszeit für Prüfung und Korrektur des Outputs. METR selbst bezeichnet das Ergebnis als Momentaufnahme des Stands Anfang 2025 und warnt vor Verallgemeinerung.
- MIT Project NANDA, „The GenAI Divide: State of AI in Business 2025“: Lernlücke bei Werkzeugen und Organisationen als Hauptursache ausbleibender Ergebniseffekte.
- Gallup, State of the Global Workplace 2026 und „Global Indicator: Artificial Intelligence“: Diskrepanz zwischen individuellen Produktivitätsgewinnen und organisationaler Wirkung.
- Roland Berger, „The AI-First Organization“, Erhebung Dezember 2025 bis April 2026, 472 Führungskräfte: 59 % halten die eigene Führung für nicht ausreichend vorbereitet, 49 % nennen fehlende KI-Fähigkeiten als größte Hürde.

Die Liste, die es in keinem Organigramm gibt

An welchen Stellen in Ihrem Unternehmen liegt heute eine Freigabe bei jemandem, der den Inhalt fachlich nicht beurteilen kann?

Diese Liste existiert in keinem Organigramm und in keinem Risikoregister. Sie zu erstellen dauert etwa zwei Wochen — und verändert danach jede weitere Diskussion über KI.

Drei Fragen, mit denen Sie anfangen können

- 1 Welche Entscheidungen kann bei Ihnen heute niemand mehr fachlich prüfen, sondern nur noch weiterreichen?
- 2 Wie hoch ist Ihre Ablehnungsquote bei KI-Output? Nicht die Nutzungsquote.
- 3 Wo steht bei Ihnen geschrieben, woran man erkennt, dass ein Fall eine Ausnahme ist?

Warum wir daran arbeiten

Die Antwort auf die dritte Frage steht in keinem Prozesshandbuch, weil sie nie ausgesprochen werden musste. Sie sitzt in den Köpfen der Erfahrenen. Genau dieses Wissen aus Köpfen in etwas Nutzbares zu übersetzen, ist die Arbeit, an der wir bei Omnora seit Jahren sitzen — nicht archivieren, sondern in Wirkung bringen.

Omnora ist eine KI-gestützte SaaS-Plattform für Corporate Learning und Intelligence Management. Mehr als 650 Unternehmen nutzen Omnora, um Wissen zu sichern, auffindbar zu machen und zu teilen. Die Plattform wird in der EU gehostet und ist DSGVO-konform.

Wenn Sie diese Liste nicht allein bauen wollen: 30 Minuten, kein Pitch, nur die Fragen. Schreiben Sie an **info@omnora.com**.