

KAPITEL 7

70/30 als Architekturprinzip — wie man die zwei Schichten praktisch baut

Warum 70 Prozent Automatisierung und 30 Prozent menschliche Prüfung eine Kostenwirkung erzielen, die weder die Vollautomatisierung noch der menschliche Prozess mit aufgesetzter KI erreicht. Und wie beide Schichten — die maschinelle und die soziale — so entworfen werden, dass jede in ihrem eigenen Tempo arbeitet und beide sich in einer gemeinsamen Verantwortungsschleife schließen.

Bain & Company zeigte in einer Untersuchung von 2025: Unternehmen, die generative KI mit einem echten Umbau ihrer Prozesse verbinden, senken Kosten um bis zu 25 Prozent. Wer KI auf bestehende Prozesse aufsetzt, ohne sie umzubauen, erreicht einstellige Prozentwerte oder gar nichts. Diese 25 Prozent sind das stärkste externe Argument für den 70/30-Ansatz: Die Technologie allein liefert keine Ökonomie — die Ökonomie liefert die gemeinsame Konstruktion der menschlichen und der technischen Schicht.

Schichtenmodell: technische und soziale Schicht

Die meisten Architekturdiagramme von KI-Systemen zeigen die technische Schicht: Modell, Datenpipeline, Infrastruktur, Integrationen. Manchmal noch die Oberfläche für Endanwender. Selten mehr.

Eine 70/30-Architektur verlangt ein Diagramm, auf dem beide Schichten zusammen und in ihrer Wechselwirkung stehen: Modell, Datenpipeline, Infrastruktur auf der einen Seite; Rollen, Eskalationspfade, Mandate, Schwellenwerte, Protokollierung und deren Prüfer auf der anderen. Beide Schichten sind vollwertige, gleichberechtigte Komponenten des Systems. Erst zusammen ergeben sie ein KI-System: Die technische Schicht allein bleibt ein automatisierter Prozess ohne Kontrolle, die soziale allein eine Sammlung menschlicher Entscheidungen ohne automatisierte Stütze. Schon Norbert Wiener beschrieb in *The Human Use of Human Beings* das automatisierte System als eines, das erst in Kombination mit einer menschlichen Aufsichtsschicht Sinn erhält; ohne sie schließt sich die kybernetische Schleife auf sich selbst und verliert ihren Korrekturmechanismus.

Technische Schicht. Dazu gehören: das Modell — oder mehrere Modelle, die zusammenarbeiten und ihre Ergebnisse zu einer Entscheidung bündeln (ein solcher Verbund heißt Ensemble). Die Pipeline der Datenvorverarbeitung: Bereinigung, Normalisierung, Merkmalsbildung. Die Recheninfrastruktur für den laufenden Betrieb. Die Bewertungs- und Schwellenlogik, also die Übersetzung von Modellergebnissen in Entscheidungen oder in Markierungen zur menschlichen Prüfung. Die Protokollierung: was das System automatisch festhält. Das Monitoring: Leistungsmetriken und die Erkennung von Datendrift — der schleichenden Verschiebung der Eingangsdaten über die Zeit.

Soziale Schicht. Rollendefinition: Training, Kontrolle, Haftung — mit konkreten Namen. Eskalationspfade: wer in welcher Reihenfolge unter welchen Bedingungen benachrichtigt wird. Übersteuerungs-Prozeduren: wie Menschen Systementscheidungen aufheben und was dabei protokolliert wird. Mandate: wer welche Entscheidungen treffen darf, mit klaren Grenzen. Aufsichtsprozesse: wie oft die Eskalationsschwellen — die Zahlenwerte, an denen eine Entscheidung von der Maschine an den Menschen übergeht — geprüft werden und wer daran teilnimmt. Dokumentationspflichten: was bis wann von wem festzuhalten ist.

Crew Resource Management als institutioneller Präzedenzfall.

Die Luftfahrt kam in den 1970er-Jahren zu diesem Schluss. Bis 1979 zeigten die Unfallberichte ein stabiles Muster: Technisch einwandfreie Flugzeuge verunglückten an Koordinationsfehlern im Cockpit — der Kapitän, der keinen Widerspruch duldete, die Hierarchie-Verzerrung, die verweigerte Delegation. Der NASA-Psychologe John Lauber prägte 1979 den Begriff Crew Resource Management (CRM, Steuerung der Besatzungsressourcen). Im Jahrzehnt darauf machte die FAA mit dem Advisory Circular 120-51 das CRM-Training für alle US-Fluggesellschaften verpflichtend; die aktuelle Fassung AC 120-51E dokumentiert fünf CRM-Generationen.[1] Die Lektion für KI: Die Luftfahrt hat formell anerkannt, dass auch das perfekte Flugzeug nur die Hälfte des Systems ist. Die andere Hälfte ist die trainierte soziale Architektur des Cockpits — Rollen, Eskalationswege, das Mandat des Copiloten zu sagen: *Stopp*. Genau das ist die soziale Schicht als gleichberechtigte Komponente. KI-Systeme stehen 2026 ungefähr dort, wo die Luftfahrt 1978 stand: Die technische Perfektion wächst schnell, die soziale Schicht ist unterentwickelt.

Der Acht-Schritte-Prozess: von der Prozesslandkarte zur Eskalationsübung

Es folgt die praktische Sequenz für den Aufbau eines 70/30-Systems. Jeder Schritt geht auf eine Fehlerklasse zurück, die sich quer durch Unternehmen genau dann wiederholt, wenn dieser Schritt übersprungen wird. Die Kürzung der acht Schritte auf drei oder vier wirkt am Anfang verlockend — sie senkt jedoch drastisch die Wahrscheinlichkeit, dass das System nach der Inbetriebnahme trägt. Jeder der acht Punkte schließt eine konkrete Ausfallklasse, ohne die keine 70/30-Architektur über das Pilotprojekt hinaus skaliert.

Schritt 1: Prozesslandkarte. Der erste Schritt ist die Aufnahme des heutigen Prozesses — noch ohne Modell und ohne KI. Festgehalten wird der Ablauf, wie er heute mit Menschen aussieht: jeder Schritt, jede Entscheidung, jeder Übergang zwischen Beteiligten. Diese Phase wirkt oft übertrieben gründlich und verführt zum Abkürzen; ohne sie weiß die Organisation jedoch nicht, was genau sie automatisiert — und riskiert, die KI auf ein verzerrtes Bild der eigenen Arbeit zu setzen. Das Ergebnis ist die Prozesslandkarte mit allen Entscheidungspunkten, auf die sich die folgenden Schritte stützen.

Schritt 2: Entscheidungspunkte identifizieren. Auf der Landkarte wird jeder Punkt markiert, an dem eine Entscheidung fällt. Nicht jeder eignet sich zur Automatisierung — ein Teil bleibt unabhängig von den technischen Möglichkeiten beim Menschen, aus regulatorischen, ethischen oder Kontextgründen. Für jeden Punkt beantwortet das Team vier Fragen: Was kostet ein Fehler an dieser Stelle? Welche regulatorische Last liegt auf der Entscheidung? Wie häufig fällt sie im normalen Rhythmus? Wie variabel sind die Eingangsdaten? Die Antworten entscheiden, ob ein Punkt überhaupt für die Automatisierung infrage kommt.

Schritt 3: Die 70 Prozent definieren. Aus den Entscheidungspunkten wird die Teilmenge bestimmt, die das System automatisch ausführt. Die Kriterien: hohe Wiederholfrequenz, geringe Variabilität der Eingangsdaten, tragbare Fehlerkosten, keine regulatorische Pflicht zur menschlichen Unterschrift. Punkte, die alle vier Kriterien zugleich erfüllen, bilden die 70 Prozent. Der Wert schwankt je nach Unternehmen — in disziplinierten Prozessen erreicht er manchmal 80 Prozent, in dichter regulierten Umgebungen liegt er spürbar darunter. Fällt der automatisierbare Anteil deutlich unter 70 Prozent, ist das kein Signal zur Bescheidenheit, sondern zum Umbau der Prozesse selbst: Sie müssen so umgebaut werden, dass sie die Kriterien tragen — sonst bleibt die Ökonomie des 70/30-Ansatzes unerreichbar.

Schritt 4: Die 30 Prozent definieren. Die Punkte, die beim Menschen bleiben, bilden die vollwertige zweite Hälfte des Systems und verdienen dieselbe Entwurfssorgfalt wie die automatisierte. Für jeden menschlichen Entscheidungspunkt werden drei Parameter fixiert: wer konkret entscheidet, unter welchen Bedingungen die Entscheidung zu dieser Person gelangt und in welchem Zeitfenster sie fallen muss. Entscheidend ist die namentliche Benennung, nicht die Funktionsrolle. Ein Eintrag wie *zuständige Bereichsleitung* oder *diensthabender Experte* genügt nicht: Hinter dieser Formel fühlt sich im Moment des Vorfalles niemand persönlich verantwortlich. Ein tragfähiger Eintrag nennt den Namen, die Position und die dokumentierte Vertretung für Abwesenheit, Urlaub oder veränderte Arbeitszeiten.

Schritt 5: Schwellenwerte festlegen. Für jeden automatisierten Punkt definiert das Team die Schwellen, an denen die Sache an den Menschen übergeht. Beispielformulierungen: *Liegt die Modellkonfidenz unter 85 Prozent, eskaliert die Entscheidung an den KI-Operator*, oder: *Betrifft die Entscheidung eine Charge über 500 Einheiten, ist menschliche Bestätigung erforderlich*. Schwellen müssen spezifisch, messbar und in der Betriebsdokumentation festgehalten sein — nicht nur im Code versteckt. Sonst bleiben sie für die Menschen, die täglich mit ihnen arbeiten, unsichtbar, und jede Revision der Schwellen setzt eine Tauchfahrt in den Code voraus.

Schritt 6: Protokollierungs-Design. Protokolliert wird jede Systementscheidung, jede Eskalation und jede manuelle Übersteuerung — mit Name, Zeitstempel und Begründung. Separat wird geregelt, wo die Protokolle liegen, wer Zugriff hat und wie lange sie aufbewahrt werden. Der EU AI Act verlangt für Hochrisiko-Systeme eine Aufbewahrung von mindestens fünf Jahren; die interne Richtlinie kann längere Fristen vorsehen, wenn Branchenstandards oder die Risikolage es verlangen.

Schritt 7: Betriebsdokumentation. Vor der Inbetriebnahme liegen drei unterschriebene Dokumente vor. Die *Systembeschreibung* hält fest, für wen das System gebaut ist, was es tut, welche Eskalationsschwellen gelten und wer Training, Betriebskontrolle und finale Entscheidung verantwortet. Das

Betriebsverfahren beschreibt, was der Operator im Normalbetrieb, bei Eskalation und bei Ausfall tut. Der *Wiederanlaufplan* regelt, was bei vollständiger Nichtverfügbarkeit geschieht und wie der Prozess manuell weiterläuft, bis das System in den Normalzustand zurückkehrt.

Schritt 8: Eskalationsübung. Vor der Inbetriebnahme simuliert das Team einen Vorfall: Ein Eskalationsszenario wird in Echtzeit durchgespielt — wer die Meldung erhält, wer reagiert, was diese Person tut, was protokolliert wird. Die Dauer wird sofort gemessen, noch in der Übung. Welche Dauer tragbar ist, hängt vom Entscheidungstyp ab: bei kritischen Vorfällen typischerweise wenige Minuten, bei Routinefällen einige Dutzend — und die konkrete Grenze gehört ins Betriebsverfahren. Im realen Vorfall wächst die Dauer weiter, durch Stress, Nachtzeit oder die Abwesenheit einer Schlüsselperson; die Übung unter ruhigen Bedingungen liefert nur die Untergrenze.

NASA „Lock the Doors“ als Präzedenzfall der Übung. In der NASA-Missionskontrolle existiert eine dokumentierte Prozedur namens „Lock the Doors“ — der Flugdirektor ruft sie im Moment eines kritischen Vorfalls aus. Die Türen des Kontrollraums werden physisch verschlossen, der Zugang von außen endet, jede Station wechselt in den Modus, den die Branche als FREEZE-ISOLATE-PROTECT beschreibt — einfrieren, isolieren, schützen: den aktuellen Systemzustand fixieren, von neuen Änderungen isolieren, Dokumentation und Protokolle sichern. Die Prozedur ist nicht für den Vorfall selbst da — sie ist für das Danach. Sie garantiert, dass die Daten bei Beginn der Untersuchung im Originalzustand vorliegen und nicht in der hektischen Reaktion überschrieben wurden. Für den KI-Betrieb liegt das Äquivalent auf der Hand: Wenn die Eskalation kommt, ist der erste Schritt nicht *reparieren* — der erste Schritt ist festzuhalten, welcher Systemzustand zur Eskalation führte, mit wem, zu welcher Zeit, auf welchen Eingangsdaten. Ohne das wird die Vorfallsanalyse eine Woche später zur Spekulation. Eine Übung ohne Einfrier-Phase trainiert die Reaktion, aber nicht die Diagnose.[2]

Dokumentationsstandard: was schriftlich existieren muss, bevor das Modell in Produktion geht

Dies ist die Mindestliste ohne Anspruch auf Vollständigkeit. Für Hochrisiko-Systeme nach dem EU AI Act gilt deutlich mehr.

Systemkarte. Name und Version des Systems. Komponenten: Modell, Version, Infrastruktur, Integrationen. Klassifikation nach dem EU AI Act. Datum der letzten Aktualisierung und Unterschrift.

Rollen und Kontakte. Training: Name, Position, Kontakt. Kontrolle (operativ): Name, Position, Kontakt, Vertretungskontakt. Haftung: Name, Position, Kontakt. Datum der Freigabe, Unterschrift jedes Beteiligten.

Operative Schwellen. Jede automatische Eskalationsschwelle als Zahl. Jeder kontextuelle Eskalationsauslöser. Datum der letzten Prüfung und die Person, die sie freigegeben hat.

Eskalationspfad. Szenario → wer wird benachrichtigt → Antwortzeit → was diese Person tut. Pro Szenario: Standard-Eskalation, Nacht- und Wochenend-Eskalation, kritische Eskalation.

Übersteuerungs-Prozedur. Wie wird die Übersteuerung technisch ausgeführt? Welche Felder gehören ins Protokoll — in welchem Format? Wer prüft die Übersteuerungs-Protokolle, und wann?

Wiederanlaufplan. Was geschieht bei vollständigem Systemausfall? Die manuelle Ersatzprozedur. Das Kriterium für die Rückkehr zum System nach der Wiederherstellung.

Typische Fehler

Fehler 1: Mandats-Konzentration. Ein erfahrener Techniker bekommt alle Rollen — Training, Kontrolle, Haftung, dazu die *allgemeine KI-Verantwortung*. Auf dem Papier verständlich. In der Betriebsrealität ist diese Person überlastet, kann nicht überall zugleich sein und wird im Moment des Vorfalls zum einzigen Ausfallpunkt. Die Rollenverteilung muss die reale Verfügbarkeit abbilden, nicht die organisatorische Bequemlichkeit.

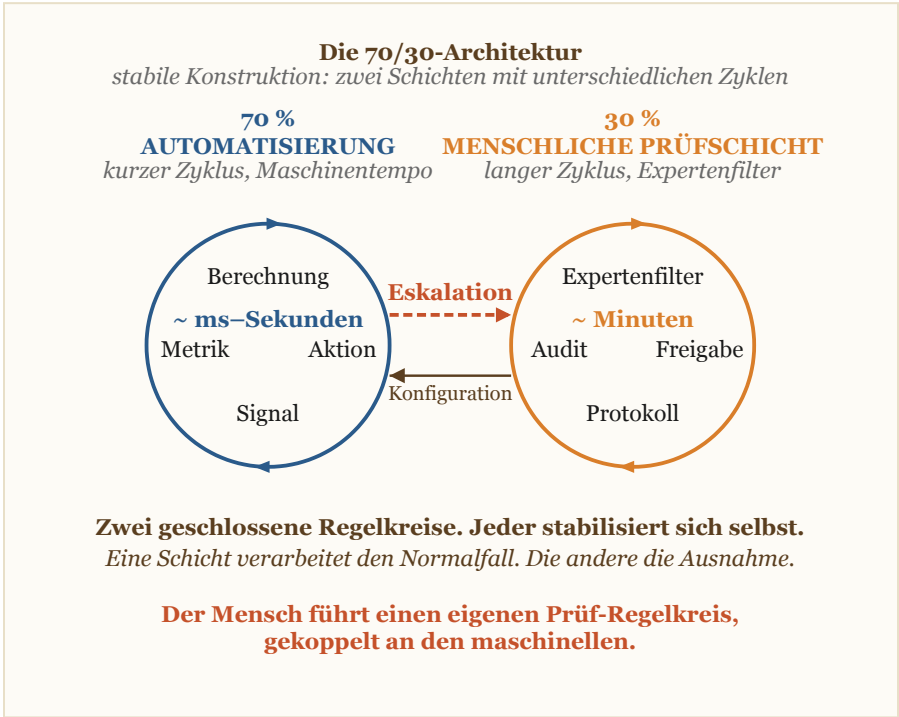
Fehler 2: Technische Schwellen ohne menschlichen Kontext. *Eskalation bei Modellkonfidenz unter 85 Prozent* klingt präzise. Aber wenn der Operator, bei dem die Eskalation landet, nicht weiß, was *Konfidenz 84 Prozent bei diesem Entscheidungstyp* praktisch bedeutet, kann er keine sinnvolle Entscheidung treffen. Technische Schwellen brauchen eine menschliche Erklärung: *Das bedeutet, das System ist sich seiner Entscheidung nicht sicher — zum Beispiel wegen eines unbekannten Dokumentformats oder eines neuen Lieferanten. In diesem Fall prüft der Operator selbst den Dokumenttyp und die Schlüsselfelder und bestätigt dann die Entscheidung oder gibt das Dokument in die erneute Verarbeitung.*

Fehler 3: Dokumentation, die niemand liest. Eine hervorragende Dokumentation in einem Ablageordner des Firmenportals, den seit der Inbetriebnahme niemand geöffnet hat, ist keine Dokumentation. Betriebsdokumentation gehört dorthin, wo Menschen sie brauchen: an den Arbeitsplatz des Operators, in die Systemoberfläche, ins erste Fenster nach Eingang einer Eskalation.

Fehler 4: Aufsicht ohne Termin. *Die Schwellen werden bei Bedarf geprüft* heißt: Sie werden nicht geprüft. Aufsicht gehört in den Kalender — einmal pro Quartal, oder nach jedem signifikanten Vorfall: ein konkreter Termin mit konkreten Personen, der konkrete Metriken prüft und entscheidet, ob Schwellen oder Rollenverteilung anzupassen sind.

Fehler 5: Kein Übergabeplan beim Personalwechsel. Der KI-Operator nimmt Urlaub. Oder kündigt. Wer übernimmt seine Funktion? Ist diese Person eingearbeitet? Weiß sie, wo die Betriebsdokumentation liegt? Die Wissensübergabe darf nicht situativ passieren — sie ist eine geplante Komponente des Systems.

Bain & Company (2025): Unternehmen, die generative KI mit dem Umbau ihrer Prozesse kombinieren, senken Kosten um 25 Prozent. Die Ökonomie entsteht erst zusammen mit dem Umbau — und der 70/30-Ansatz ist genau die Methode dieses Umbaus.



Ankerdiagramm K7/1. Zwei geschlossene Regelkreise: der maschinelle (Normalfall, Sekunden) und der menschliche (Ausnahme, Minuten). Jeder stabilisiert sich im eigenen Tempo.

Drei Fragen für den eigenen Kontext

1. **Schichten-Diagramm.** Zeichnen Sie Ihr KI-System in zwei Schichten — technisch und sozial. Lässt sich die soziale Schicht nicht zeichnen, existiert sie konstruktiv nicht.
2. **Schwellen-Prüfung.** Wann wurden die operativen Schwellen Ihres Systems zuletzt geprüft? Wer war dabei? Was wurde geändert? Wenn die Antwort *weiß ich nicht mehr* oder *nie* lautet — handeln Sie.
3. **Eskalationsübung jetzt.** Wählen Sie ein Szenario und stoppen Sie die Zeit. Wer erhielt die Meldung? Wer reagierte? Was dauerte am längsten? Das Ergebnis der Übung ist die Tagesordnung des nächsten Aufsichtstermins.

Quellen

1. John K. Lauber, „Resource Management on the Flightdeck“, NASA Conference Publication 2120, Juni 1979; FAA Advisory Circular 120-51E „Crew Resource Management Training“, 22. Januar 2004 (aktuelle Fassung).
2. NASA Johnson Space Center, Flight Director Operations Handbook (FDOH); Beschreibung der Prozedur „Lock the Doors“ in: Gene Kranz, *Failure Is Not an Option*, Simon & Schuster, 2000.