

Fünf Thesen zur Bewertung von
Geschäftsmodellen unter KI-Einfluss

27.04.2026
DATUM

Was Künstliche Intelligenz mit der Marge macht

Vorwort

Dieses Papier entwickelt einen Analyserahmen für Geschäftsmodelle unter KI-Einfluss. Adressat ist der Investor, der vor einer Portfolio-Entscheidung wissen muss, unter welchen Bedingungen eine heute tragfähige Marge im Investitionshorizont hält, erodiert oder verschwindet. Der Rahmen ist keine Prognose, keine Tool-Empfehlung und keine Branchenstudie. Er beschreibt Wirkmechanismen und benennt die Bedingungen, unter denen sie greifen.

Unter *Künstlicher Intelligenz* (KI) verstehen wir allgemein nutzbare, datengetriebene Modelle – Foundation Models – und die von ihnen gesteuerte Robotik. Die Modelle erbringen kognitive Arbeit in einer Breite und Tiefe, die frühere Automatisierungswellen deutlich übersteigt. Die Robotik erweitert ihre Wirkung in den physischen Raum. Technische Einzelheiten werden nur insoweit behandelt, wie sie für die ökonomische Analyse relevant sind. Das Papier bleibt bewusst auf der strategischen Ebene: Operative und taktische Instrumente der Unternehmensbewertung bleiben unberührt.

Das Papier trägt fünf Thesen, um die herum alles andere angeordnet ist:

These I. KI zerstört zuerst Friktion, nicht Arbeit. Wer Friktion für ein Gate hält, wird die Erosion seiner Marge zu spät bemerken.

These II. Die Investorenfrage ist nicht, ob KI Kosten senkt, sondern ob sie die Quelle der Marge zerstört, verlagert oder neu begründet.

These III. Die unsichtbarsten Angriffe kommen nicht aus der Substitution innerhalb der eigenen Kette, sondern aus Pfaden, die die Kette umgehen.

These IV. Zeitpunktprognosen zur KI-Entwicklung scheitern an der eigenen Methode; tragfähig ist die Beobachtung von Sperrern und ihrer Vorindikatoren.

These V. KI erzeugt eine neue Klasse verteidigbarer Gate-Quellen. Die mächtigste ist der geschlossene Lernkreislauf – das *KI-Flywheel* –, der nur dann Marge trägt, wenn der Kreislauf-Betreiber technisch und rechtlich die Kontrolle über seine Daten hält.

Jede These bekommt ein Kapitel. Vor den Thesen stehen zwei Kapitel zu den Grundbegriffen, die das Papier trägt; nach ihnen steht eine kurze Synthese. Ein Leitfaden, der die fünf Thesen in konkrete Prüfpunkte für Screening und Due Diligence überführt, ist als eigenständiges Dokument beigelegt.

Teil I – Die Grundbegriffe

1. Marge und ihre zwei Quellen

Die Marge einer Beteiligung lässt sich aus dem Jahresabschluss ablesen. Ihre Quelle nicht. Solange sich die Rahmenbedingungen nicht verändern, ist das unerheblich. Im Umbruch ist die Höhe der heutigen Marge eine schwache Information; über die Marge von morgen entscheidet ihre Quelle.

Dieses Papier reduziert die Frage nach der Quelle analytisch auf zwei Kategorien, weil sie die für KI relevanten Bedrohungsvektoren bündeln. Die erste ist strukturelle Nicht-Replizierbarkeit: eine Position, die der Wettbewerb innerhalb eines relevanten Zeithorizonts nicht angreifen kann – auch bei ausreichender finanzieller und technischer Ausstattung. Die zweite ist Marktunvollkommenheit, die der Wettbewerb angreifen könnte, bisher aber nicht getan hat, weil der Angriff sich nicht rechnet. Wir bezeichnen die erste als *Gate*, die zweite als *Friktion*. Andere Rentenquellen – Risikokompensation, temporäre Knappheit – existieren, werden aber entweder durch Gate oder Friktion dominiert oder sind für die KI-Analyse nachrangig.

Friktion kommt in drei Gestalten vor. *Intransparenz* – der Kunde kennt die Alternative nicht. *Trägheit* – er scheut den Vergleichsaufwand. *Wechselkosten* – der Wechsel wäre möglich, aber aufwendig. Jede dieser Gestalten ist eine Kostenposition auf Kundenseite, die Marge auf Anbieterseite erzeugt, solange sie bestehen bleibt.

Beide Quellen werden durch KI angegriffen, aber mit unterschiedlichen Mechanismen und Geschwindigkeiten. Friktion fällt unmittelbar, weil Vergleichsagenten, automatisierte Empfehlungen und KI-gestützte Migrationswerkzeuge genau die Kosten auf Kundenseite senken,

aus denen Friktion wirtschaftlich besteht. Gates fallen mittelbar: Die Investitionskosten eines Angriffs sinken, weil KI die dafür nötige kognitive Arbeit verbilligt; die Wertschöpfungskette verschiebt sich, so dass Gates umgangen werden; und neue, KI-spezifische Gates entstehen, die an die Stelle alter treten können.

Die Unterscheidung zwischen Gate und Friktion ist das zentrale diagnostische Werkzeug dieses Papiers. Eine Marge, die auf Friktion ruht, ist unter KI-Einfluss akut angreifbar. Eine Marge, die auf einem Gate ruht, ist nicht automatisch sicher – ihre Halbwertszeit hängt vom Typ des Gates und von der Stabilität der Wertschöpfungskette ab, in der es wirkt.

2. Der Zerstörbarkeitstest

Die Unterscheidung ist nur dann nützlich, wenn sie sich operativ anwenden lässt. Man denkt sich mehrere gut finanzierte Wettbewerber – Neueinsteiger oder Fast Follower mit großem Kapitaleinsatz –, die mit aktueller KI und Robotik ausgestattet drei bis fünf Jahre Zeit haben, gegen die zu prüfende Marge anzugehen. Was sie unter diesen Bedingungen nicht beseitigen können, ist Gate. Was sie beseitigen würden, sobald sich der Angriff rechnet, ist Friktion. Der Zeithorizont orientiert sich am typischen PE-Investitionshorizont; die Ausstattung mit aktueller Technik ist keine Verschärfung, sondern Notwendigkeit, wenn der Test die relevanten Mechaniken abbilden soll.

Eine Reihe von Wettbewerbsvorteilen, die im klassischen Raster als Gate geführt werden, bestehen den Test nicht. *Unsere Analysten sind die besten* verlangt, dass zehn neue Anbieter mit aktueller KI keine vergleichbare Qualität liefern – eine Bedingung, die in den meisten kognitiven Domänen nicht standhält. Die eigentliche Barriere war nicht die Kompetenz, sondern die Schwierigkeit des Kunden, Qualität zu vergleichen: Intransparenz, nicht strukturelle Nicht-Replizierbarkeit. *Wir haben jahrzehntelange Kundenbeziehungen* hält, solange der Kunde keinen Anlass hat, sich umzusehen; sobald ein KI-Agent die Alternativen kontinuierlich und kostenfrei vergleicht, entfällt dieser Schutz.

Gates, die den Test bestehen, lassen sich nach Angriffspunkten sortieren. Physische Kapazitäts-Gates (Häfen, Pipelines, Spezialinfrastruktur) halten, solange das System, in dem sie wirken, bestehen bleibt – dagegen kann weder KI noch Robotik direkt etwas ausrichten, weil die Physik des Standorts die Replikation verhindert. Regulatorische Gates (Lizenzen, Zulassungen, gesetzliche Monopole) sind stabil gegen Technologie, aber politisch angreifbar und unter KI-Einfluss zunehmend unter Rechtfertigungsdruck. Netzwerkeffekte erzeugen superlinear wachsenden Wert, den Arbeit allein nicht einholt. Technologische Gates und Patente sind durch KI-beschleunigte Umgehungsforschung angreifbarer geworden. Akzeptanz-Gates – Barrieren auf der Nachfrageseite – folgen eigenen, generationellen Dynamiken, die in Kapitel 6 behandelt werden.

Gates aus akkumulierter Substanz – Datenbestände, Softwareplattformen, Wissensarchive – sind die verletzlichste Kategorie unter KI-Einfluss, und sie sind scharf von Netzwerkeffekten zu unterscheiden. Ein Netzwerkeffekt-Gate schützt seinen Inhaber, weil Arbeit das Netzwerk nicht nachbauen kann. Ein Substanz-Gate schützt nur so lange, wie die Replikation der Substanz eine wirtschaftlich prohibitive Menge an Arbeit voraussetzt. Viele Wettbewerbspositionen – gerade in der Software-Branche – beruhen genau auf dieser Prohibitivität: Der Wert lag nicht in struktureller Unangreifbarkeit, sondern in der schiereren Menge kognitiver Arbeit, die zur Nachbildung nötig

gewesen wäre. Wenn KI diese Arbeitsmenge drastisch reduziert, wird aus dem vermeintlichen Gate Friktion.

Der Test erfasst einen Angriffsvektor nicht: die Umgehung. Ein Gate kann den Test bestehen und gleichzeitig ökonomisch entwertet werden, weil eine Verschiebung der Wertschöpfungskette es irrelevant macht. Das Gate bleibt intakt; die Marge, die es trug, verschwindet. Die systematische Behandlung folgt in These III.

Die vollständige Zerlegung der Preisstruktur in sieben Kategorien – physische Arbeit (P), kognitive Arbeit (K), Substanzerhalt (S), Wagnis (W), Lenkung (L), Gate (G), Friktion (F) – ist im beigelegten Leitfaden als kompakte Referenz ausgeführt. Die Zerlegung ist KI-spezifisch geschnitten, nicht nach buchhalterischer oder volkswirtschaftlicher Herkunft. Für die Thesen genügen K (weil KI hier ansetzt), G und F (weil sie die Marge tragen). Die anderen Kategorien treten punktuell auf, wo sie die Argumentation tragen.

Teil II – Die fünf Thesen

3. These I – KI zerstört zuerst Friktion, nicht Arbeit

KI zerstört zuerst Friktion, nicht Arbeit. Wer Friktion für ein Gate hält, wird die Erosion seiner Marge zu spät bemerken.

Die Debatte über KI dreht sich um ersetzte Berufe, automatisierte Prozesse, eingesparte Kosten. Das ist empirisch korrekt und strategisch unvollständig. Substitution von Arbeit trifft die Kostenseite; dort löst sie den Red-Queen-Effekt aus, den Kapitel 4 entfaltet: In wettbewerbsintensiven Segmenten landet der Effizienzgewinn bei den Kunden, in konzentrierten Segmenten teilweise beim Anbieter. In keinem Fall entsteht aus Kostensenkung allein zusätzliche Marge. Für den Investor ist Effizienz unter KI deshalb kein Vorteil, sondern ein Eintrittsticket – wer nicht mitzieht, liegt in kurzer Zeit über den Grenzkosten des Markts.

Der erste Effekt, der die Marge tatsächlich bewegt, ist die Zerstörung von Friktion. Friktion besteht wirtschaftlich aus drei Substanzen: aus den Kosten, Information zu beschaffen; aus dem Aufwand, Angebote zu vergleichen; aus der Reibung, von einem Anbieter zum nächsten zu wechseln. Genau diese drei Kostenarten sind das, was KI am effizientesten reduziert – Recherche, Vergleich, Migration. Das Internet senkte einen schmalen Ausschnitt davon, vor allem Preistransparenz in definierten Kategorien. Was das Internet für Hotelpreise war, wird KI für jede Beurteilung, die bisher einen Experten verlangte.

Im Mittelstand wird regelmäßig verwechselt, was Kompetenz ist und was Intransparenz. *Unsere Analysten sind die besten* wirkt als Gate, weil der Kunde nicht vergleichen kann. Die Marge entsteht nicht aus der Kompetenz selbst, sondern aus ihrer Unvergleichbarkeit. Sobald die Unvergleichbarkeit fällt, wird Qualität messbar, und die Marge, die aus ihrer Messbarkeitslücke bestand, verschwindet. Das heißt nicht, dass Kompetenz nichts mehr wert wäre: Wo Kompetenz knapp und *nicht-replizierbar* ist – der Spitzenchirurg, der Vorstandsvorsitzende in einer personenbezogenen Beziehung –, trägt sie auch in voller Transparenz Marge. Was unter KI fällt, ist die *replizierbare* Kompetenz, die als Gate verkauft wurde.

Das Muster wiederholt sich jenseits kognitiver Leistungen. *Wir haben die schnellste Lieferung* war Friktion, solange kein Wettbewerber dieselbe Geschwindigkeit erreichte oder dem Kunden sichtbar machen konnte, dass er sie erreicht. *Unser Service ist überlegen* war Friktion, solange Servicequalität nicht messbar war. *Unsere Zulassungen sichern uns ab* kann Gate sein, wenn die Zulassung einen Marktzugang blockiert, oder Friktion, wenn sie nur einen administrativen Aufwand darstellt, den ein gut finanzierter Wettbewerber innerhalb der Halteperiode nehmen würde. Der Zerstörbarkeitstest aus Kapitel 2 ist die verlässliche Methode, diese Unterscheidung zu treffen.

Für den Investor folgt daraus: Ein Jahresabschluss mit hoher Marge ist kein Signal für die Zukunft. Die nächste Frage muss immer sein, aus welcher Quelle die Marge stammt. Fällt der größte Teil auf Friktion, ist sie unter KI-Einfluss akut gefährdet; fällt er auf Gate, hängt die Langlebigkeit vom Gate-Typ und von der Pfadstabilität ab (Kapitel 5). Unternehmen, die sich in der Selbstbeschreibung nahezu ausschließlich mit Kompetenz- und Erfahrungsargumenten beschreiben, gehören in die erste Kategorie – unabhängig davon, ob Führung und Belegschaft das wissen oder nicht.

Friktion verschwindet nicht vollständig, sie verlagert sich. Wo alte Friktion fällt, entsteht neue: *Plattformabhängigkeit*, weil Anwendungen auf Foundation Models aufsetzen, über die der Anwender keine Kontrolle hat; *Komplexitätsasymmetrie*, weil KI-durchdrungene Prozesse eine Verständnislücke zwischen Betreibern und Nutzern erzeugen; *Vertrauensprämie*, weil maschinell erzeugte Leistung häufig einen menschlichen Validator braucht, um am Markt akzeptiert zu werden. Die Nettomenge an Friktion im System kann steigen oder fallen – beobachtbar ist, dass sie zu dem wandert, der sich schneller bewegt. Wer alte Friktion aktiv abräumt, zerstört die Marge des Langsameren. Wer neue Friktion aufbaut, bevor der Markt sie versteht, fängt Marge ein, die vorher nicht existierte.

→ **Exkurs 1 – F ohne G darunter: Reisebüros**

Reisebüros lebten von reiner Friktions-Marge. Der Kunde wusste nicht, was Flüge und Hotels anderswo kosteten (Intransparenz), scheute den Aufwand eigener Recherche (Trägheit) und vertraute dem vertrauten Büro (Wechselkosten). Gates gab es keine, die den Zerstörbarkeitstest bestanden hätten; zehn Wettbewerber hätten das Geschäft in der damaligen Technologielage ohne Weiteres nachbauen können. Das Internet zerstörte die drei Friktionen binnen weniger Jahre durch Vergleichsplattformen, einfache Buchungsschnittstellen und einheitliche Oberflächen. Weil kein Gate darunter lag, kollabierte die Marge vollständig. KI ist die nächste, breitere Welle desselben Mechanismus – sie trifft jedes Geschäftsmodell, dessen Marge auf Unwissenheit, Bequemlichkeit oder Aufwand beim Kunden beruht. Versicherungsmakler, Finanzberater, Immobilienmakler, intransparente Dienstleister aller Art stehen vor derselben Erosion, schneller und breiter als beim Internet-Schub.

4. These II – Die Margenquelle ist die Investorenfrage, nicht die Effizienz

Die Investorenfrage ist nicht, ob KI Kosten senkt, sondern ob sie die Quelle der Marge zerstört, verlagert oder neu begründet.

Vorstände reagieren auf KI mit Effizienzprogrammen. Sie müssen es tun: Wer die eigenen Prozesse nicht mit KI-Werkzeugen beschleunigt, liegt binnen zwei Jahren messbar über den Grenzkosten der Konkurrenz. Für den Investor verdeckt diese Bewegung jedoch die eigentliche Frage – ob die Margenquelle hält. Ein Unternehmen, das sichtbar in KI investiert und Effizienzgewinne ausweist, kann gleichzeitig die Quelle seiner Marge verlieren, ohne dass es in der Berichterstattung auftaucht.

Der erste Mechanismus ist die *Doppelbewegung*. Wenn alle Wettbewerber dieselben KI-Werkzeuge einsetzen, wird der Effizienzgewinn zum Branchenstandard; der Wettbewerb treibt die Preise auf die neuen, niedrigeren Grenzkosten. Der Effizienzgewinn landet beim Kunden, nicht beim Anbieter, der ihn erzielt – zumindest in wettbewerbsintensiven Segmenten. Wer nicht mitrennt, fällt zurück. Das ist der *Red-Queen-Effekt*: Man rennt, um nicht zurückzufallen, und kommt dabei nicht voran. Eine Besonderheit der Bewegung ist ökonomisch präzise: Wenn der Gesamtpreis unter KI-Einfluss fällt, sinkt die absolute Marge pro Einheit stärker als der relative Margenanteil – weil die Grundgesamtheit, auf die der Anteil bezogen ist, selbst kleiner wird. Wer aus Margenprozenten auf Ergebnisqualität schließt, unterschätzt den Verlust.

Die zweite Bewegung ist unsichtbarer und schärfer. Parallel zur Effizienzsteigerung kommoditisiert KI die kognitiven Fähigkeiten selbst. Der Kostenabstand zu potenziellen Wettbewerbern, die diese Fähigkeiten bisher nicht hatten, sinkt – unabhängig davon, ob der Incumbent selbst investiert. Wer nicht investiert, verliert früher; wer investiert, verlangsamt, verhindert aber nicht. Ein Geschäftsmodell, dessen Marge auf der Unerreichbarkeit kognitiver Kompetenz ruhte, verliert sein Fundament. Der einzige Ausweg ist, dass die Investition nicht nur Kosten senkt, sondern im selben Zug ein neues Gate aufbaut. Davon handelt These V.

Aus der Doppelbewegung folgt die zentrale Filterung für den Bewertungsprozess. Jeder Brutto-Hebel aus KI – Effizienz, Wachstum, strategische Integration – muss durch die Frage gefiltert werden, wer den entstehenden Wert einfängt. Drei Konstellationen sind zu unterscheiden.

Der Hebel kann durch ein *Gate* geschützt sein. Das ist der Regelfall in Branchen mit proprietären Daten, die ein Flywheel antreiben, mit regulatorischen Positionen, die durch KI-gestützte Validierung verstärkt werden, mit physischer Kapazität, die sich durch KI effizienter nutzen lässt. In diesen Konstellationen bleibt der Gewinn beim Unternehmen.

Der Hebel kann durch *persistente Friktion* innerhalb der Halteperiode monetarisierbar sein. Das ist ein Trade auf Zeit, keine Transformation. Der First-Mover-Vorteil bei KI-Adoption gehört hierher: kein Gate, aber nutzbar, solange der Wettbewerb langsamer ist. Die Frage lautet, ob die Friktion schneller erodiert als die Halteperiode dauert.

Der Hebel kann *allgemein verfügbar* sein. Dann entsteht kein Wert für den Investor, weil der Wettbewerb denselben Hebel erreicht; die Ersparnis fließt zu den Kunden. Hoher Brutto-Hebel bei null Netto-Hebel ist eine der häufigen Fallen in KI-Investment-Thesen, und sie lässt sich nur durch disziplinierte Prüfung vermeiden: Welche Position des Unternehmens macht diesen Effizienzgewinn zu *seinem* Gewinn?

Eine eigene Gefahr entsteht dort, wo der Effizienzgewinn nicht nur an die Kunden wandert, sondern an die Infrastruktur-Ebene – an den Foundation-Model-Anbieter, den Cloud-Betreiber, den Plattformbetreiber. Wenn K durch KI sinkt, sinken die Lohnkosten nicht einfach auf null; ein Teil der

Kosten verlagert sich in Substanzaufwendungen: API-Gebühren, Compute-Kosten, Modellpflege, Lizenzen. Das ist nicht Kostensenkung, sondern Kostenverlagerung von arbeitsintensiv zu kapitalintensiv, mit veränderter Risikostruktur (Fixkosten, Abschreibungszyklen, Abhängigkeit von externen Anbietern). Die Marge, die aus dem K-Block entweicht, kann beim Unternehmen bleiben – oder in den S-Block eines Technologielieferanten wandern. Der Unterschied ist keine Bilanzposition, er ist eine vertragliche und technische Frage.

Der *Jevons-Mechanismus* wirkt parallel und in die entgegengesetzte Richtung. Wenn die Stückkosten einer Leistung unter KI fallen, steigt in vielen Märkten die Nachfrage mehr als proportional. Juristische Erstberatung, die bisher nur oberhalb einer Streitwertschwelle tragfähig war, wird für Alltagsfragen zugänglich. Medizinische Diagnostik, die episodisch in Anspruch genommen wurde, wird zu kontinuierlichem Monitoring. Softwareentwicklung wird für Projekte tragfähig, die bisher unterhalb der Rentabilitätsschwelle lagen. In jedem dieser Fälle entsteht ein Markt, der vorher nicht existierte, weil die alten Kosten ihn unzugänglich machten. Die Frage für den Investor ist zweiteilig: Existiert ein solcher latenter Markt in der Umgebung des Portfoliounternehmens – und ist das Unternehmen in der Lage, ihn zu besetzen, oder entsteht dort ein KI-nativer Angreifer, der ihn besser bedient?

Eine eigene, oft übersehene Dynamik betrifft Unternehmen, die selbst kaum K-Exposition haben. Ein Logistiker mit dominantem P-Block, ein Hafenbetreiber, ein Spezialinfrastruktur-Anbieter – sie werden von der unmittelbaren K-Substitution nicht getroffen und betrachten sich unter KI oft als sicher. Sie sind es nicht. Wenn K in den umgebenden Schichten der Wertschöpfungskette drastisch fällt, steigt der relative Anteil aller verbleibenden Kostenblöcke. Der eigene, bisher unauffällige Kostenblock wird zum sichtbarsten Posten in der Kundenrechnung und damit zum offensichtlichen Optimierungsziel, obwohl am Vorgang selbst nichts verändert wurde. Wir bezeichnen diesen Mechanismus als *Druckeffekt*. Er ist kein Preisdruck, er ist Aufmerksamkeitsdruck – auf Kostenblöcke, die der Kunde bisher als feste Größe akzeptierte, weil seine Optimierungsenergie den größeren Blöcken galt. Sobald die größeren Blöcke verschwinden, verschwindet auch diese Akzeptanz. Der Druckeffekt wirkt dort am stärksten, wo Kunden Aufmerksamkeit nach Kostenanteil allozieren – typischerweise in standardisierten B2B-Käufen. Ein KI-Screening, das nur die direkte K/P-Exposition eines Unternehmens prüft, produziert systematisch falsche Sicherheit, weil es diesen Mechanismus übersieht.

5. These III – Die unsichtbarsten Angriffe umgehen die Kette

Die unsichtbarsten Angriffe kommen nicht aus der Substitution innerhalb der eigenen Kette, sondern aus Pfaden, die die Kette umgehen.

Die Wertschöpfungskette, in der ein Unternehmen heute operiert, ist für die Bewertung zunächst selbstverständlich. Kostenvorteile, Verhandlungsmacht, Eintrittsbarrieren – die gängige Analyse operiert auf dieser Kette. Sie übersieht, was KI an ihr tut: Sie greift nicht primär in die Kette, sie ersetzt sie.

Pfadkonkurrenz liegt nicht darin, dass eine Schicht der Kette billiger wird, sondern dass ein alternativer Pfad tragfähig wird, der die Kette insgesamt umgeht. Das Bedürfnis des Endkunden bleibt; die Kette, die es erfüllt, wird durch eine andere ersetzt. Der Parkhausbetreiber aus Exkurs 2 verliert nicht an ein besseres Parkhaus. Er verliert an eine Mobilitätskette, die keinen

innerstädtischen Parkraum mehr braucht. Das Gate bleibt physisch intakt und ökonomisch wertlos.

→ Exkurs 2 – Gate-Entwertung ohne Gate-Bruch: Parkhaus

Der Betreiber eines Innenstadt-Parkhauses verfügt über ein physisches Gate: Standortexklusivität, langfristige Flächennutzungsrechte, hohe Wiederbeschaffungskosten. Im Rahmen klassischer Wettbewerbsanalysen ist die Position robust. Die Prognose – und es bleibt eine Prognose, keine abgeschlossene Beobachtung –, dass autonome Fahrdienste in städtischen Kernen in einem mittelfristigen Horizont tragfähig werden, verändert das Bild nicht beim Wettbewerb, sondern bei der Nachfrage. Das Fahrzeug setzt ab und verlagert sich an einen kostengünstigen Standort außerhalb. Das Gate bleibt physisch intakt; die Marge, die es trug, entfällt, weil das Gut, das es schützt – innerstädtischer Parkraum –, nicht mehr nachgefragt wird. Das Muster tritt strukturähnlich in weiteren Branchen auf und ist der analytische Kern von These III.

Wir nennen diesen Mechanismus *Pfadkonkurrenz*. Pfade entstehen dort, wo die Absenkung eines Kostenfaktors einen Alternativweg tragfähig macht, der zuvor unökonomisch war. Wenn die K-Kosten für Fahrzeugsteuerung auf nahezu null fallen, wird ein Mobilitätspfad ohne menschlichen Fahrer ökonomisch. Wenn K-Kosten für juristische Erstbewertung unter die Kanzleisätze fallen, wird ein Markt für Alltagsrechtsberatung tragfähig, der zuvor nicht existierte. Wenn die Fertigung durch KI-gesteuerte Roboter an den Kunden heranrückt, wird ein Beschaffungspfad tragfähig, der den klassischen Zulieferer umgeht.

Die Analyse wird präziser, wenn man fragt, *wo* in einem Gesamtsystem die kognitive Arbeit überhaupt sitzt, die KI angreifen kann. Sie sitzt an drei Stellen. *Im Produkt* – ein Beratungsbericht ist Kognition, Software ist kodifizierte Kognition. *In der Nutzung* – das Auto enthält keine, sein Fahrer schon; die Werkzeugmaschine enthält keine unmittelbar substituierbare Produktkognition, ihr Einrichter und ihr Programmierer aber schon. *Im System* – die Einspritzpumpe enthält weder noch, sitzt aber in einem Motor in einem Auto, das einen Fahrer braucht. Wenn KI den Fahrer ersetzt, beschleunigt sich die Elektrifizierung, der Verbrennungsmotor verschwindet, und die Einspritzpumpe verliert ihr System, obwohl sie in sich keine KI-Exposition hat.

Ein Unternehmen, das in sich keine direkte KI-Exposition aufweist, kann über die Nutzung seines Produkts oder über das System, in dem es operiert, voll exponiert sein. Jede Prüfung eines Portfoliounternehmens muss deshalb alle drei Ebenen ausdrücklich abfragen.

Wenn ein Pfad kippt, betrifft das alles, was von ihm abhängt. Wir bezeichnen als *Kaskade* die Folge von Pfadverschiebungen, die eine ursprüngliche Verschiebung in benachbarten Branchen auslöst. Autonomes Fahren kippt nicht nur den Pfad „Taxi mit Fahrer“; es kippt parallel die Kfz-Schadenversicherung (weniger Unfälle, sinkende Prämien), den Parkraumbetrieb, die Führerschein-Ausbildung, den Einzelhandel in Innenstädten, die Einspritzpumpen-Zulieferer. Kaskaden laufen nicht nur in eine Richtung: Jede Pfadverschiebung erzeugt neue Nachfrage, die neue Wertschöpfung trägt – neue Haftungsprodukte, neue Wartungssegmente für Flotten, neue Logistikmodelle. Die Kaskadenanalyse ist keine Spekulation, sondern die notwendige zweite

Ordnung der Bewertung. Ein auf erster Ebene stabiles Unternehmen kann auf zweiter Ebene in einem sterbenden System stehen.

Substanz, die ein Unternehmen über Jahrzehnte aufgebaut hat, verhält sich unter Pfadwechsel unterschiedlich – je nachdem, ob sie *pfadgebunden* oder *pfadübertragbar* ist. Der historische Vergleich zwischen Kodak und Fujifilm (Exkurs 3) ist das klarste Beispiel.

→ Exkurs 3 – Pfadgebundene und pfadübertragbare Substanz: Kodak und Fujifilm

Beide Unternehmen hatten über Jahrzehnte Substanz in der Filmchemie aufgebaut und hielten technologische Gates im analogen Filmpfad. Die Digitalisierung machte den Pfad obsolet. Kodaks Substanz – Filmchemie, Entwicklungslabore, Distributionsnetze für Filmrollen – war im digitalen Pfad wertlos: Sie blieb in der Bilanz, erzeugte laufende Kosten, trug aber keine Marge mehr. Fujifilms Substanz in Materialwissenschaft, Dünnschichttechnologie, Gelatine- und Kollagenchemie aus der Filmemulsion war auf andere Domänen übertragbar; sie wurde zur Grundlage eigener Gates in Kosmetik, medizinischer Bildgebung und LCD-Optik. Dieselbe Kategorie von Bilanz-Asset, zwei fundamental verschiedene Schicksale. Die Prüffrage für jede Substanz unter KI-Einfluss lautet: Wäre sie in einem alternativen, KI-geprägten Pfad weiterhin Gate-tragend, oder ist sie pfadgebunden?

Die Konsequenz für die Bewertung ist einschneidend. Ein Portfolio lässt sich unter KI-Einfluss nicht asset-für-asset auf direkte KI-Exposition prüfen. Jede Einzelprüfung muss mindestens eine zusätzliche Ebene einbeziehen: die Pfadstabilität der Kette, in der das Unternehmen operiert, und die Pfadstabilität der Systeme, in die sein Produkt einfließt. Ein Zulieferer mit intaktem technologischen Gate kann innerhalb einer Halteperiode in einem System ohne Käufer stehen, und kein Jahresabschluss zeigt dieses Risiko heute an.

6. These IV – Sperren statt Kalender

Zeitpunktprognosen zur KI-Entwicklung scheitern an der eigenen Methode; tragfähig ist die Beobachtung von Sperren und ihrer Vorindikatoren.

Zeitprognosen zur KI-Entwicklung sind notorisch schlecht. Sie überschätzen kurzfristig, weil die ersten Pilotprojekte scheitern und die ersten Demos nicht überzeugen; sie unterschätzen langfristig, weil die schrittweise Fähigkeitszunahme an der kritischen Schwelle in plötzliche Adoption kippt. Für die Portfolio-Entscheidung ist das zweifach unbrauchbar: Falsche Früh-Prognosen erzeugen verfrühte Transformations-Investitionen; falsche Spät-Prognosen erzeugen übersehene Risiken. Die Antwort kann nicht sein, bessere Prognosen zu bauen. Die Antwort ist, den Analyserahmen von Prognosen zu lösen.

Der hier vorgeschlagene Ansatz arbeitet mit beobachtbaren Sperren statt mit Zeitpunkten. Eine *Sperre* ist eine Bedingung – technologisch, regulatorisch oder gesellschaftlich –, deren Nichterfüllung die KI-Durchdringung einer Domäne blockiert. Solange mindestens eine Sperre aktiv ist, bleibt die Domäne weitgehend geschützt. Die Methodik definiert für jede relevante Sperre

Vorindikatoren, deren Eintreten den Fall der Sperre ankündigt; das Monitoring-System reagiert auf Signale, nicht auf Schätzungen.

Drei technische Sperren lassen sich präzise benennen. Die erste betrifft die *Quelldaten*: Existieren Rohdaten, aus denen ein Modell die Struktur der Domäne lernen kann? Die Sperre fällt am schnellsten, weil Foundation Models zunehmend weniger domänenspezifische Daten brauchen und synthetische Datenerzeugung den Bedarf weiter reduziert. Innerhalb der Sperre unterscheidet sich *Beobachtungsdaten* (was passiert in der Welt) von *Interventionsdaten* (was passiert, wenn ich aktiv etwas ändere) und *kontrafaktischen Daten* (was wäre passiert, wenn ich anders entschieden hätte). Beobachtungsdaten reichen für Mustererkennung, Interventionsdaten für Handlungsempfehlungen, kontrafaktische Daten für Kausalanalyse. Domänen, in denen Interventionen teuer und Simulationen unvollständig sind – klinische Therapieentscheidungen, makroökonomische Politik, komplexe chirurgische Eingriffe –, halten diese Sperre auch dann, wenn Beobachtungsdaten in großer Menge vorliegen.

Die zweite Sperre ist das *Validierungssignal*: Existiert ein Signal, das dem Modell zurückmeldet, ob seine Ausgabe korrekt ist? Diese Sperre ist die bestimmende in vielen wirtschaftlich relevanten Domänen. Eine *Outcome-Validierung* (hielt der Vertrag, war die Diagnose richtig) ist billig, aber der Feedback-Zyklus ist lang und die Kausalzuordnung unsicher. Eine *Prozess-Validierung* (war jeder Zwischenschritt korrekt) ist teurer, ermöglicht aber ein um Größenordnungen effizienteres Training, weil der Lernalgorithmus auf jedem Schritt ein Fehlersignal erhält. Domänen mit präziser Prozess-Validierung – Code durch Ausführung, Mathematik durch Beweiskontrolle, Spiele durch Regelbewertung – erreichen in kurzer Zeit Expertenniveau. Domänen mit nur Outcome-Validierung bleiben auf Jahre gebremst. Zusätzlich bestimmt, wer validieren kann: Laien-Validierung durch jeden Kunden oder Nutzer ist kostenfrei und massenhaft; Experten-Validierung ist teuer und in der Menge begrenzt. Daraus folgt eine ökonomische Konsequenz: Nicht das beste Modell gewinnt, sondern die schnellste Validierungsschleife. Wo Modelle schneller produzieren, als menschliche Experten prüfen können, wird die Validierung selbst zum Engpass – und damit zum wichtigsten Investitionsziel im Aufbau eines KI-Vorteils.

Die dritte Sperre ist die *Simulationsumgebung*: Kann die Domäne digital simuliert werden? In Bereichen, in denen die zugrundeliegende Physik vollständig verstanden ist – Mechanik, Strömungslehre, Halbleiterphysik, Proteinfaltung –, kann Simulation so gut werden, dass sie die Validierung ersetzt. In diesem Fall kollabieren Sperre 2 und 3 in eine: Die Simulation *ist* die Validierung. Das erklärt, warum Softwareentwicklung rasch durchdrungen wird: Die Ausführung des Codes ist zugleich die Validierung. Dieselbe Kollaps-Logik gilt in strukturierter Lagerhaltung, mechanischer Konstruktion, Teilen der Materialforschung. Sie gilt nicht in der Chirurgie, nicht in der Wirkstoffentwicklung für komplexe Erkrankungen, nicht in volkswirtschaftlichen Interventionen.

Neben den drei technischen Sperren wirken zwei soziale. *Regulatorische Sperren* folgen politischen Zyklen, Lobby-Aktivität und internationalen Präzedenzfällen. Sie halten nur so lange, wie alle relevanten Jurisdiktionen sie aufrechterhalten; sobald eine Jurisdiktion sie senkt, entsteht regulatorische Arbitrage, und der Druck auf die übrigen Jurisdiktionen steigt. *Akzeptanz-Sperren* sitzen auf der Nachfrageseite: Ein KI-System kann technisch reif, regulatorisch zugelassen und ökonomisch vorteilhaft sein und trotzdem nicht eingesetzt werden, weil Kunden oder Nutzer menschliche Interaktion erwarten. Akzeptanz-Sperren erodieren generationell und durch

kumulative positive Erfahrungen; spektakuläre Fehlschläge können sie punktuell zurückwerfen. Der Wechsel der öffentlichen Diskussion von „ob“ zu „wann“ ist der zuverlässigste Vorindikator ihres Falls – weil eine Diskussion über das Wann eine implizite Einigung über das Dass voraussetzt. Der gesellschaftliche Raum, in dem grundsätzliche Ablehnung noch möglich war, hat sich geschlossen, auch wenn das nicht formuliert wird.

Neben dem Sperrenstatus bestimmt eine zweite Variable das Tempo der Durchdringung: die *Feedback-Zykluslänge* einer Domäne. Sie ist die Zeit, die vergeht, bis nach einer Modellausgabe bekannt wird, ob sie korrekt war. Sie reicht von Millisekunden (Code-Ausführung, Fahrsituation) über Tage (Diagnose-Befund) bis zu Jahren (Wirkstoffentwicklung, Materialermüdung). Bei identischer Sperrenkonstellation führt sie zu Durchdringungsgeschwindigkeiten, die um Größenordnungen auseinanderliegen. Code kippt innerhalb einer Halteperiode, Onkologie nicht. Eine Domäne mit langer Feedback-Zykluslänge hat für den Investor eine doppelte Eigenschaft: Der Vorsprung eines Pioniers hält lange, weil auch Konkurrenten viele Jahre brauchen, um die nötigen Lernzyklen zu durchlaufen – ein einmal etabliertes Gate ist dort ungewöhnlich stabil. Im Gegenzug verlängert sich auch die eigene Aufbauphase, in der noch keine Gate-Wirkung trägt; das Kapital ist über entsprechend lange Zeit gebunden, bevor sich die Schutzwirkung einstellt.

Eine dritte Variable wirkt auf die Schärfe des Falls: der *Staudruck* hinter einer geschlossenen Sperre. Wenn eine Sperre eine technisch längst mögliche Lösung blockiert und hinter ihr substantielle latente Nachfrage wartet, akkumuliert sich über die Zeit ein Disruptionspotenzial. Die Technologie reift weiter, Kapital fließt in Infrastruktur, Anwendungen entstehen in weniger restriktiven Jurisdiktionen. Wenn die Sperre fällt, tritt die zuvor blockierte Adoption in komprimiertem Zeitraum ein. Ein Kippunkt ohne Staudruck ist zeitlich gestreckt; ein Kippunkt mit hohem Staudruck ist komprimiert. Der europäische Regulierungsraum ist ein relevantes Beispiel. Ein Investor, der eine geschlossene Sperre als dauerhafte Sicherheit liest, setzt sich genau diesem Effekt aus. Zusätzlich kann die regulatorische Bewegung selbst Vorindikator sein: Wo staatliche Lenkung aktiv auf den KI-Umbruch reagiert – Schutzregulierung, Förderprogramme, Exportkontrollen –, wird das Regulierungsumfeld zu einer eigenen Dynamik, die beobachtet werden muss.

Domänen sortieren sich nach Sperrenstatus grob in drei Zonen, die jeweils eine andere Strategie verlangen. Eine *Commodity-Zone*, in der alle drei technischen Sperren offen sind – allgemeine Textverarbeitung, generischer Code, Standard-Bildverarbeitung –, wird von den großen Frontier-Laboren für den gesamten Markt gelöst; ein eigenständiger Moat ist hier kaum aufzubauen. Eine *Integrations-Zone*, in der mindestens eine Sperre noch aktiv ist – typischerweise proprietäre Daten oder Experten-Validierung – erlaubt einen Moat-Aufbau durch den Anbieter, der den Datenzugang und die Validierungspipeline kontrolliert. Eine *Sperrzone*, in der alle drei Sperren aktiv bleiben – Chirurgie-Robotik, Werkstoffforschung, seltene industrielle Prozesse –, verlangt tiefe Domänenexpertise, eigene Datenerhebung, physische Infrastruktur; die Zeitlinien sind lang, der Kapitalbedarf hoch, der Moat aber entsprechend massiv.

Das Monitoring der Sperren ist kein Kalender, sondern eine periodische Prüfliste. Wenn keine Vorindikatoren eintreten, passiert nichts, und die Analyse bleibt intakt. Wenn Signale eintreten, sind sie vorab definiert, und die Reaktion kann vorbereitet sein.

7. These V – Die neue Klasse von Gate-Quellen

KI erzeugt eine neue Klasse verteidigbarer Gate-Quellen. Die mächtigste ist der geschlossene Lernkreislauf – das KI-Flywheel –, der nur dann Marge trägt, wenn der Kreislauf-Betreiber technisch und rechtlich die Kontrolle über seine Daten hält.

Die bisherigen Thesen behandeln die Bedrohung: Friktion fällt, Effizienz wird Standard, Pfade verschieben sich, Sperren fallen. Die fünfte These behandelt die Chance. Aus dem Umbruch entstehen nicht nur Risiken; es entstehen zwei neue Klassen von Margenquellen, die nebeneinander wirken.

Die mächtigste ist der geschlossene Lernkreislauf – das *KI-Flywheel*. Der Begriff Flywheel ist nicht neu; wir verwenden die spezifische Form, in der ein KI-System durch den Betrieb seines eigenen Produkts mit Trainingsdaten gespeist wird, um es von Plattform-Flywheels im klassischen Sinn (Netzwerkeffekte, mehrseitige Märkte) abzugrenzen. Innerhalb des Flywheels liegt die wertvollste Position auf der Validierungs- und Haftungsschicht: Wer dem Markt die Verantwortung für eine KI-gestützte Leistung abnimmt – durch nachweisbare Validierung und vertragliche Haftungsübernahme –, kontrolliert den Punkt, an dem KI-Output zu Marktwert wird. Die Mechanik dahinter ist nicht primär technisch, sondern ökonomisch: KI-Ausgaben sind ohne externe Verantwortungszusage nur begrenzt am Markt platzierbar; wer diese Zusage glaubhaft erbringt, hält die preissetzende Position im Kreislauf.

Die zweite Quelle sind *First-Mover-Vorteile in neuen Pfaden*: Geschäftsmodelle, die durch fallende KI-Kosten erst tragfähig werden und in denen bisher keine etablierte Konkurrenz existiert. Wer einen solchen Pfad früh besetzt, baut sein Gate aus dem Vorsprung heraus auf – temporär entstanden, aber in Kundenbeziehungen, Marktposition oder Plattformwirkung konvertierbar. Pfadsubstitution ist dabei nicht nur Bedrohung (These III), sondern auch eine Chance, sich aus einer anderen Richtung in einen fremden Markt einzubringen.

Wer sich allein auf das KI-Flywheel verlässt, verpasst die Pfadschöpfung; wer auf die Pfadschöpfung allein setzt, verpasst die strukturelle Tiefe des Flywheels. In den meisten Fällen wird der Investor beide Quellen prüfen müssen.

Der Flywheel-Mechanismus lässt sich als Sequenz von sechs Stufen beschreiben, die im Betrieb des Produkts zusammenwirken müssen. Die Sequenz ist eine Referenz-Architektur, nicht die einzig mögliche; in der Praxis sind mehrseitige Plattformen oder mehrstufige Lernkreisläufe ebenso gangbar, solange die Funktionen erfüllt sind.

Die erste Stufe: Das Produkt ist selbst *Datensammler*. Der Kunde finanziert die Datenerhebung als Nebenprodukt seiner Nutzung, ohne dass eine separate Datenakquise betrieben werden muss. Dieser Punkt trennt tragfähige KI-Strategien von solchen, die Daten als eigene Kostenstelle aufbauen müssen.

Die zweite Stufe: Eine *Simulationsumgebung* ergänzt die realen Daten um synthetische Szenarien. In vielen Domänen sind seltene, aber sicherheitskritische Situationen – der *Long Tail* der Ereignisse – in realen Daten unterrepräsentiert. Simulationen skalieren diese Fälle und erzeugen zusätzlich *adversariale Szenarien*, also gezielt konstruierte Grenzfälle, an denen das Modell seine Schwächen offenbart. Beides verkürzt den Weg zur Robustheit.

Die dritte Stufe: Eine *Testbasis* wird aus realen Fehlern, Grenzfällen und Diskrepanzen abgeleitet. Sie wächst kumulativ – jeder neue Fall erweitert die Prüfkriterien, gegen die künftige Modellversionen bestehen müssen. Dadurch entsteht Regressionshärte: Fortschritt an einer Stelle darf keine Verschlechterung an einer anderen nach sich ziehen.

Die vierte Stufe: Ein *Freigabeprozess*, in dem kein neues Modell in den Einsatz geht, ohne sich gegen die gesamte Testbasis zu bewähren. Das macht die Testbasis strategisch zentral – sie ist nicht Begleitmaterial, sondern Torwart.

Die fünfte Stufe: *Shadow-Inferenz* im Betrieb. Neben dem aktiven Modell läuft ein Beobachtungsmodus, der Abweichungen zwischen Modellerwartung und realem Verlauf erfasst, ohne selbst zu steuern – eine Diskrepanzerfassung ohne Eingriffsrisiko.

Die sechste Stufe: *Rücklauf*. Die erfassten Abweichungen fließen in Trainingsdaten und Testbasis zurück. Der Kreislauf schließt sich und beginnt mit jeder Iteration schneller zu drehen. Die Drehzahl dieses Kreislaufs – wie schnell eine vollständige Runde aus Datenerhebung, Modellaktualisierung, Validierung und Rücklauf abgeschlossen wird – bezeichnen wir als *Iterationsgeschwindigkeit*. Sie ist neben Datenkontrolle und der Validierungs- und Haftungsschicht die dritte Größe, die darüber entscheidet, ob der Kreislauf Marge trägt: In adversarialen Domänen wächst der Vorsprung mit jeder Iteration, in stationären Domänen schließt sich das Konvertierungsfenster mit ihr.

Ob ein solches Flywheel ein Gate erzeugt, hängt nicht primär von seiner technischen Konstruktion ab. Die entscheidende Bedingung ist *Datenkontrolle*. Ein Flywheel, dessen Daten dem Kunden gehören oder deren Nutzungsrechte bei einem Hyperscaler liegen, erzeugt kein Gate bei dem Unternehmen, das die Daten produziert. Die Wertschöpfung aus der Dateninterpretation fließt zum Rechteinhaber. Ein Maschinenbauer, der über seine Anlagen Betriebsdaten sammelt, dessen Lieferverträge aber vorsehen, dass die Daten dem Kunden gehören, hat kein Flywheel, sondern eine Kostenstelle. Ein Laborgerätehersteller, dessen Messdaten in einer Cloud liegen, deren Auswertungsinfrastruktur von einem Hyperscaler bereitgestellt wird, hat den Teil der Wertschöpfung, der aus der Interpretation stammt, an den Hyperscaler abgegeben.

Datenkontrolle ist kein rechtlicher Binärzustand. Sie hat eine rechtliche Seite (Nutzungsrechte, Exklusivität, Übertragbarkeitsklauseln) und eine technische Seite (Durchsatz, Latenz, Zugriff in der nötigen Menge und Frequenz). Wer rechtlich Eigentümer der Daten ist, aber nur über ein API-Kontingent eines Drittanbieters Zugriff hat, hat formal Kontrolle, aber ökonomisch nicht die Kontrolle, die für einen drehenden Kreislauf nötig ist. Umgekehrt kann ein Unternehmen über latente Kontrolle verfügen – technisch im Besitz der Daten, rechtlich schwach abgesichert –, und diese Position erodiert beim ersten Vertragsbruch. Due-Diligence-Prozesse prüfen diesen Spektrumscharakter selten systematisch; der häufigste Datenrechteverlust entsteht nicht durch aktiven Rechtsstreit, sondern durch Standardklauseln in Service-Agreements, die bei Abschluss nicht als strategisch wichtig wahrgenommen wurden. Umgekehrt lässt sich die eigene Position aktiv schützen – durch sorgfältig formulierte Klauseln zu Daten-Ownership, Nutzungsrechten und technischem Zugriff in Bestands- wie Neuverträgen. Die DD-Frage ist deshalb zweiseitig: Was wurde bisher eingeräumt, und was lässt sich in zukünftigen Vertragsrunden zurückholen oder absichern?

Nicht jedes Flywheel ist gleich dauerhaft. Die Dauerhaftigkeit des resultierenden Gates hängt von einer Eigenschaft der Domäne ab: Reagiert die Umgebung auf das Modell. In *stationären* Domänen

reagiert die Umgebung nicht – Physik bleibt Physik. Der Problemraum ist endlich; mehrere Spieler mit hinreichend Daten und Compute konvergieren auf eine vergleichbare Lösung. Das bedeutet nicht, dass First-Mover-Vorteile in stationären Domänen immer temporal bleiben: Wo Netzwerkeffekte, Plattformposition oder regulatorische Position sich mit dem Datenvorsprung verbinden, kann auch in einer stationären Domäne ein struktureller Moat entstehen – Google Maps ist ein bekanntes Beispiel. Wo diese Konvertierung unterbleibt, verschwindet der Zeitvorsprung, sobald die Wettbewerber aufschließen. In *adversarialen* Domänen reagiert die Umgebung – Angreifer adaptieren gegen Verteidiger, Gegenparteien optimieren gegen die KI, Wettbewerber reagieren auf Preissignale. Der Problemraum ist offen. Der Marktführer muss permanent retrainieren, aber er formt die Umgebung mit, und ein Nachfolger müsste die gesamte Koevolutionsgeschichte nachbilden. Der Moat ist hier strukturell, nicht nur temporal. Cybersecurity, Finanzmarktalgorithmen und dynamisches Pricing gehören zu dieser Kategorie.

Manche Domänen wechseln ihr Regime, sobald KI sie durchdringt. Die Rechtsstreit-Domäne ist heute weitgehend stationär; sobald eine Partei massenhaft KI-generierte Schriftsätze einreicht, beginnt die Gegenseite mit eigener KI zu optimieren, und aus stationär wird adversarial. Für die Bewertung eines KI-Vorsprungs ist das künftige Regime die primäre Frage; das heutige Regime ist das schwächere Signal.

→ **Exkurs 4 – Tesla: Der bewusst gebaute Flywheel**

Tesla verbaut in jedem ausgelieferten Fahrzeug die vollständige Sensorik und Rechnerkapazität für das autonome Fahrsystem, ohne Aufpreis. Jedes Fahrzeug ist primär Datensammler, erst nachrangig Produkt im klassischen Sinn. Die Simulation ergänzt reale Fahrten in so großem Umfang, dass synthetische Szenarien einen wesentlichen Teil des Trainingsvolumens stellen. Validierung erfolgt durch jeden Fahrer mit Führerschein – die Einschätzung „diese Situation ist gefährlich“ erfordert kein Spezialwissen, was die Validierungskosten pro Datenpunkt nahe null drückt. Jedes Fahrzeug läuft im Shadow-Mode, jede Diskrepanz zwischen Modellerwartung und tatsächlicher Realität wird erfasst. Alle drei technischen Sperren sind offen; die Feedback-Zykluslänge liegt im Bereich von Minuten bis Tagen. Die Entscheidung, auf die Hardware-Marge zu verzichten, ist nicht Opportunismus, sondern die konstitutive Entscheidung für einen Flywheel. Autonomes Fahren ist eine stationäre Domäne – der First-Mover-Vorteil wäre deshalb ohne Konvertierung temporär, und Tesla arbeitet aktiv an der Transformation vom Fahrzeughersteller zum Mobilitätsanbieter, um den temporären Vorteil in eine Plattformposition zu überführen.

→ **Exkurs 5 – Ocado: Der nicht geschlossene Kreislauf**

Ocado betreibt hochautomatisierte Fulfillment-Center mit Roboterschwärmen in einer stationären Umgebung. Alle drei technischen Sperren sind offen: Jedes Zentrum erzeugt permanent Betriebsdaten als Nebenprodukt. „Wurde der richtige Artikel gegriffen?“ ist ein binäres, kostenloses Validierungssignal. Lagerumgebungen sind vollständig simulierbar – bekannte Geometrie, bekannte Physik, kontrollierte Umgebung. Die Ausgangslage ist dem Tesla-Fall ungewöhnlich ähnlich. Ocado optimiert stattdessen Robotergewicht, Batterielaufzeit und Grid-Effizienz: KI als Werkzeug innerhalb des bestehenden Geschäftsmodells, nicht als Mechanismus, der das Modell transformiert. Die Entscheidung, Hardware-Marge zugunsten eines aggressiven Flywheel-Aufbaus zu opfern, wurde nicht getroffen. In einer stationären Domäne mit parallelen Akteuren – Amazon baut einen vergleichbaren Kreislauf – hat die Nichtkonvertierung konkrete Konsequenzen: Der Problemraum ist endlich, das Zeitfenster schließt sich.

→ **Exkurs 6 – Validierungs-Arbitrage: Chirurgie-Robotik über die Veterinärmedizin**

Die bestimmende Sperre für autonome chirurgische Systeme ist die Validierung. Chirurgische Outcome-Validierung dauert Monate, erfordert Fachärzte und unterliegt ethischen Grenzen, die kontrafaktische Daten ausschließen. Ein möglicher Weg zur Senkung dieser Sperre führt über die Veterinärmedizin: niedrigere regulatorische Hürden, breite Nutzung von Chirurgie-Robotik in Tierkliniken, Tierärzte als Validatoren im Nebenprodukt ihrer normalen Praxis. Die ethische Bremse fiel damit weitgehend weg, weil Tiere Patienten sind, die behandelt werden, nicht Versuchsobjekte. Der Autonomiegrad könnte in der Tiermedizin wesentlich schneller hochgefahren werden; die Erkenntnisse fließen mit regulatorischem Rückenwind in die Humanmedizin zurück. Das Muster ist übertragbar – Kleintierforschung für Wirkstoffe, Mediationsplattformen mit KI-Vorentscheid für juristische Vorarbeit, interne Continuous Assurance für Wirtschaftsprüfung –, und es schafft eine eigene Kategorie von Gate: eine Datenposition, die durch Pfad-Topologie, nicht durch Produkt-Datenkontrolle entsteht.

Der Gegensatz zwischen Tesla und Ocado ist eine der schärfsten Illustrationen der These. Beide operieren in stationären Domänen mit offenen Sperren. Beide haben technischen Zugang zu denselben Werkzeugen. Der Unterschied ist die strategische Entscheidung, bereits realisierte Marge gegen den Aufbau eines Flywheels einzutauschen. Diese Entscheidung ist nicht technisch, sie ist organisatorisch. Sie setzt voraus, dass die Führung die Mechanik versteht, ihre Konsequenzen tragen will und das operative Umbauprogramm durchsetzen kann.

Damit verbindet sich eine organisatorische Dimension, die in rein marktstrukturellen Analysen fehlt. Die zentrale Frage an die Führung ist nicht, ob sie KI in ihrer Tiefe versteht – das tun die wenigsten Vorstände, auch viele Jahre nach der ersten Welle –, sondern ob sie bereit ist, die Konsequenzen zu tragen: laufende Marge gegen Datenerhebung eintauschen, bestehende Rollen neu zuschneiden, Machtstrukturen verschieben, Plattformabhängigkeiten vertraglich auf

Datenkontrolle hin neu verhandeln. Der häufigste Führungsfehler besteht darin, KI als Kostenwerkzeug im bestehenden Geschäft zu optimieren, während parallel ein alternativer Pfad das Geschäft eliminiert. Diese Führungsprüfung gehört in die Due Diligence – nicht als Nebensache, sondern als eigenständige Dimension.

Für den Investor entsteht aus den genannten Mechanismen eine gestufte Prüfung der Flywheel-Möglichkeit. Sie beginnt nicht beim Flywheel, sondern bei der Kette. Zunächst werden über Ketten- und Pfadanalyse die Stellen lokalisiert, an denen K in der Wertschöpfung tatsächlich wirkt – im Produkt, in der Nutzung, im System. Anschließend wird geprüft, welches Flywheel diese K-Anteile tragen könnte – gibt es ein Produkt, das im Betrieb Daten erzeugt, deren Interpretation zur Margenquelle werden kann? Erst dann stellt sich die Frage nach der Verteidigbarkeit: Liegen die Datenrechte rechtlich und technisch beim Unternehmen? Gibt es Wettbewerber, die ähnlich gute oder bessere Voraussetzungen für ein Flywheel haben? Und schließlich die organisatorische Frage: Ist die Führung bereit und fähig, laufende Marge in den Aufbau des Kreislaufs zu investieren? Wer in einer der Stufen eine negative Antwort erhält, steht nicht zwingend vor einem Abbruch – im Mittelstand kann die First-Mover-Position in einem regional begrenzten Pfad bereits ausreichen, wenn die Konvertierung in ein anderes Gate gelingt. Ein Flywheel, das andere Wettbewerber besser bauen können, ist aber kein Flywheel – es ist ein Wettrennen, in das ein Mittelständler in der Regel nicht gehört.

Teil III – Synthese

8. Wo die Marge anfällt

Unter stabilen Bedingungen fallen die juristische Zurechnung eines Geschäftsmodells und die ökonomische Zurechnung der Marge zusammen. Ein Unternehmen produziert etwas, verkauft es, realisiert die Marge. Der Jahresabschluss zeigt das Bild vollständig. Unter den in den fünf Thesen beschriebenen Mechanismen trennt sich diese Zuordnung.

Ein Beratungshaus, dessen kognitive Leistung aus dem Foundation Model eines Dritten kommt und dessen Modell-Infrastruktur bei einem Hyperscaler läuft, produziert juristisch weiterhin dieselbe Leistung. Die Marge, die dabei entsteht, verteilt sich nach dem Prinzip der Flywheel-Ökonomie – also der Wertverteilung in einer KI-durchdrungenen Wertschöpfungskette – auf drei Positionen: den Inhaber des Foundation Models (Leistungserbringung), den Inhaber der Infrastruktur (Betrieb), den Inhaber des Kundeninterfaces (Nachfrage). Vom Gesamtbetrag, den der Mandant zahlt, realisiert das Beratungshaus nur den Teil, der seiner eigenen Position in dieser Verteilung entspricht – im Grenzfall, als bloßer Zwischenhändler zu einem externen KI-System, deutlich weniger, als der Mandatsvertrag in seiner juristischen Form suggeriert.

Das Muster tritt in unterschiedlichen Branchen auf. Ein Maschinenbauer mit KI-gestützter Prozessoptimierung, dessen Betriebsdaten durch eine Service-Agreement-Klausel beim Kunden oder beim Cloud-Anbieter landen, produziert die Maschinen weiter und verkauft sie weiter – die Marge aus der Dateninterpretation realisiert ein anderer. Ein Laborgerätehersteller in derselben Konstellation erlebt dasselbe. Eine Kanzlei, deren Arbeit zunehmend durch ein externes KI-System gestützt wird, erbringt juristisch weiter die gleiche Dienstleistung; welcher Teil der Marge

tatsächlich bei ihr bleibt, hängt davon ab, ob sie das Interface zur KI-Nutzung kontrolliert oder nur ihr Endkunde ist.

Die Konsequenz für die Bewertung ist präzise: Der Jahresabschluss zeigt, bei welchem Unternehmen heute welche Marge gebucht wird. Die Kettenanalyse, der Sperrenstatus und die Flywheel-Frage zeigen, an welcher Stelle der Kette in den kommenden Jahren die Marge anfällt. Wo beide Bilder auseinanderfallen, entsteht der analytische Mehrwert einer KI-spezifischen Due Diligence. Der Fehler, den sie verhindert: ein Investment auf Basis des heutigen Jahresabschlusses, obwohl die ökonomische Schicht bereits bei einem anderen Akteur liegt.

Die Frage an jedes Portfoliounternehmen lässt sich in einem Satz formulieren: Realisiert das Unternehmen im relevanten Investitionshorizont noch die Marge der Wertschöpfungskette, in der es steht, oder wandert die Marge an einen anderen Akteur in dieser Kette, obwohl die juristische Struktur unverändert bleibt?

9. Grenzen dieses Rahmens

Der vorgeschlagene Rahmen hat drei Grenzen, die benannt sein müssen.

Die erste betrifft die Quantifizierung. Der Zerstörbarkeitstest, die Sperren-Analyse, die Flywheel-Prüfung produzieren qualitative Urteile, keine Zahlen. Zwei Analysten können in der konkreten Einordnung eines Gates abweichen. Das ist kein Defekt des Rahmens, sondern die Folge einer Problemlage, in der numerische Präzision Scheinsicherheit produziert. Der Rahmen erzeugt strukturierte Urteile, keine Scores.

Die zweite betrifft die Pfadidentifikation. Pfadkonkurrenz (These III) verlangt die Fähigkeit, einen noch nicht existierenden Pfad als plausibel zu erkennen. Das ist keine Methodik, das ist ein Urteil, das auf Branchenkenntnis und analytischer Fantasie beruht. Der Rahmen liefert die Struktur, in der das Urteil geprüft werden kann; er liefert nicht das Urteil selbst.

Die dritte betrifft die Nachfrageseite. Der Rahmen analysiert Angebotsstrukturen: Kostenfaktoren, Margenquellen, Pfadverschiebungen. Die Zahlungsfähigkeit und -bereitschaft der Endkunden wird als Rahmen behandelt, nicht als Variable. Innerhalb des Papiers kommt der Jevons-Mechanismus in These II als Markt-schaffende Kraft vor, weil er unmittelbar aus der Kostensenkung folgt und die Angebotsseite selbst neue Märkte eröffnet. Weiterreichende Nachfrageeffekte – etwa durch Arbeitsmarktverwerfungen, die Einkommensverteilung und damit Zahlungsbereitschaft verändern – sind real, aber in ihrer Richtung nicht prognostizierbar. Sie hängen von der relativen Position der eigenen Wirtschaft, von staatlicher Lenkung und von politischen Entscheidungen ab, die außerhalb des Rahmens liegen. Der Rahmen lässt diese Dimension offen, weil er sie nicht schließen kann, ohne Annahmen zu treffen, die er nicht begründen würde.

Das Papier endet an der Stelle, an der die operative Anwendung beginnt. Welche Fragen im Screening zu stellen sind, wie ein CEO-Gespräch die nötigen Informationen erhebt, wie Unternehmen in typisierte Investitionssituationen eingeordnet werden – das sind eigenständige Aufgaben.