

Das Schlaraffenland der quersubventionierten KI-Tools ist vorbei

Mit autonomen KI-Systemen rollt die nächste große Technologiewelle auf die IT-Branche zu. Ohne einen strukturierten Fahrplan laufen Systemhäuser Gefahr, dabei den Anschluss zu verlieren. Benjamin Eidam analysiert den aktuellen KI-Preisschock und liefert einen pragmatischen 3-Hebel-Fahrplan für Systemhaus-Chefs.

Welche Veränderung erleben wir aktuell beim Thema KI und unseren IT-Budgets?

Wir stecken mitten in einer Phase der absoluten Umorientierung. Bislang herrschte im Bereich KI ein echtes Schlaraffenland-Feeling: Die Tech-Labs haben unendlich viel Geld in ihre Systeme gepumpt und die Nutzung massiv quersubventioniert – ein genialer Marketing-Schachzug, um uns alle an die Tools heranzuführen. Jetzt, im Jahr 2026, stellen die Anbieter ihre Geschäftsmodelle um. Unternehmer müssen sich auf steigende Kosten einstellen.

Auf den ersten Blick ist das paradox, denn die reinen Recheneinheiten im Hintergrund – sprich die Tokens – werden für die Anbieter durch bessere Komprimierung eigentlich effizienter und günstiger. Aber genau hier greift ein bekanntes Paradoxon der Wirtschaft: Wenn etwas billiger wird, explodiert die Nachfrage. Und da reagieren die Anbieter jetzt mit ihrer Bepreisung.



Unser Umgang mit der KI hat sich radikal gewandelt.



Liegt das nur an der reinen Masse der Nutzung oder hat sich auch die Art und Weise verändert, wie wir KI im Betrieb einsetzen?

Auch unser Umgang mit der KI hat sich radikal gewandelt. Früher hat man mit einer KI ein bisschen hin und her geschattet – das verbraucht kaum Leistung. Heute nutzen wir immer häufiger agentische Systeme. Die KI antwortet nicht mehr nur, sondern arbeitet völlig autonom im Hintergrund komplexe Prozesse ab. Sie erstellt quasi eine ganze Armee von digitalen Unter-Agenten, die parallel Aufgaben lösen, sich gegenseitig kontrollieren und die Ergebnisse zusammenführen. Das liefert herausragende Ergebnisse, vervielfacht aber auch den Datenverbrauch.

Genau deshalb stoßen die klassischen Monats-Flatrates an ihre Grenzen und weichen immer mehr einer verbrauchsabhängigen Abrechnung. Wer die Tools jetzt unkontrolliert wie im letzten Jahr dicke Berichte oder Code am Fließband generieren lässt, wird bei den nächsten Abrechnungen eine böse Überraschung erleben.

Ein Begriff, der in diesem Zusammenhang immer öfter fällt, ist die »rekursive Selbstverbesserung«. Welche Dimension hat diese Entwicklung?

Das ist im Grunde der Moment, in dem der Mensch schrittweise aus der Entwicklungsschleife der KI verschwindet. Rekursive Selbstverbesserung bedeutet, dass KI-Systeme autonom anfangen, noch bessere und schnellere KI-Systeme zu bauen. Wir sehen aktuell bereits die ersten echten Vorboten dafür: Große Tech-Unternehmen bestätigen, dass ein erheblicher Teil ihres eigenen Codes nicht mehr von Programmierern, sondern von der KI selbst geschrieben wird.

Das Faszinierende und gleichzeitig Herausfordernde daran ist, dass unsere menschliche Reaktionszeit dabei komplett kollabiert. Bisher bewegen wir uns noch in der gewohnten »Menschenzeit«: Eine technologische Neuerung wird am Freitag veröffentlicht und wir haben als Unternehmer das Wochenende Zeit, uns für den Montag eine Strategie zu überlegen. Wenn diese automatisierte Selbstverbesserung aber



»Welche Daten dürfen in die KI?« – Benjamin Eidam auf der Bühne der SYNAXON-Partner-Geschäftsführertagung

richtig Fahrt aufnimmt, sprechen wir plötzlich von massiven Entwicklungssprüngen innerhalb von 48 Minuten statt 48 Stunden.

Wie bereitet man sich als Unternehmer auf so ein Tempo vor?

Indem man seine Hausaufgaben macht, bevor der große Sturm losbricht. Wer heute noch ohne feste interne Prozesse, ohne klare Messgrößen und ohne ein grundlegendes Verständnis für die Systeme arbeitet, wird es in Zukunft extrem schwer haben, überhaupt noch Schritt zu halten. Das ist so, als würde man ohne Boot in einen reißenden Gebirgsfluss springen. Man muss heute schon wissen: Wie messe ich den Erfolg einer Software? Welches Tool setzen wir für welche Aufgabe ein? Wer sich das erst im Moment des Versuchs überlegen will, verliert wertvolle Zeit und viel Geld.



Wenn Laien anfangen, tiefgreifende IT-Strukturen zu bauen, wird es brandgefährlich.



Was bedeutet diese rasante Entwicklung für das Thema Cybersicherheit?

Die Risiken verschieben sich gerade massiv. Durch die neuen KI-Fähigkeiten kommen wir in eine Phase, in der auch Menschen ohne tiefes IT-Verständnis via Prompting eigene Apps und Skripte bauen. Dabei passieren im Hintergrund fatale Sicherheitsfehler, die ein Laie nicht auf dem Schirm hat. Es gab neulich ein Paradebeispiel auf Social Media: Jemand präsentierte stolz seine erste selbstgebaute App. Keine drei Tage später kam der Hilferuf, weil er plötzlich eine API-Rechnung über 10.000 Dollar auf dem Tisch hatte. Was war passiert? Er hatte seinen API-Key als Klartext im Code hinterlegt. Jemand anderes hat den Key ausgelesen und auf seine Kosten fröhlich rechenintensive KI-Modelle genutzt.

Das zeigt einfach: Durch die enorme Komplexität der Systeme reicht ein reines Oberflächenwissen nicht mehr aus. Wenn Laien anfangen, tiefgreifende IT-Strukturen zu bauen, wird es brandgefährlich. Und genau da liegt die Verantwortung bei uns IT-Experten: Für einen Prototypen reicht heute zwar reines Prompting, aber sobald Software im Netzwerk läuft, braucht es IT-Know-how für Security, Leitplanken und DSGVO. Das kann keine KI ersetzen.



Viele Unternehmen haben bereits jetzt ihr gesamtes KI-Budget aufgebraucht. Wie behalten Systemhäuser im Alltag die Kontrolle über ihr Budget?

Viele von uns nutzen die Tools unbewusst noch im reinen Überfluss-Modus. Nach dem Motto: »Lieber einen 30-seitigen Bericht als drei Stichpunkte eingeben, kostet ja nix.« Genau das rächt sich durch die neuen Reasoning-Modelle bitter. Ich habe das neulich erst bei einem Kunden erlebt: Mitten im Projekt waren plötzlich alle API-Limits erschöpft und die Entwickler gerieten in Panik. Wer hier ohne Leitplanken agiert, steuert direkt in den finanziellen Blindflug.

Wie können diese Leitplanken aussehen?

Das Zauberwort heißt heute Token-Effizienz. Wir brauchen für eine simple E-Mail kein teures High-End-Modell. Zudem sollte man die Feinjustierung der Tools nutzen: Viele moderne Modelle bieten Regler für »Deep Thinking« oder schnelle Antworten – Vollgas sollte man wirklich nur dann einschalten, wenn die Aufgabe es verlangt. Ein weiterer wichtiger Hebel ist die Chat-Hygiene. Wer endlose Chats mit zehn verschiedenen Themen führt, bläht den Kontextspeicher auf und treibt die Kosten pro Nachricht massiv in die Höhe. Macht lieber für jedes Thema einen neuen Chat auf. Zu guter Letzt können wir getrost jegliche Höflichkeitsfloskeln vergessen. Ein simples „Please“ kostet die Tech-Welt weltweit schätzungsweise eine Milliarde Dollar an überflüssigen Token.

Sobald wir einen Prompt in das Chat-Fenster eintippen, ist maximale Präzision und Kontextbezogenheit geboten. Verschwendet weder eure Zeit, noch eure verfügbaren Token für unnötiges Drumherum. Bei Code- oder Datengenerierung hilft der »Caveman-Trick«: Man weist die KI explizit an, auf hochgestochene Formatierungen oder lange Einleitungen zu verzichten und nur das nackte Ergebnis auszuspacken. Das spart massiv Textzeichen und damit am Ende des Monats bares Geld.



KI soll keine Mitarbeiter ersetzen, sondern ihre Schlagkraft multiplizieren.

»Wenn die KI morgen alles 50 Mal schneller erledigt, kann ich ja bald 50 Mal weniger Mitarbeiter einstellen oder mein Team verkleinern.« Was antwortest du einem Geschäftsführer, der so kalkuliert?

Ich antworte ihm, dass das strategisch zu kurz gedacht ist und man sich damit das eigene Wachstumspotenzial nimmt. Natürlich kann man kurzfristig Kosten sparen, indem man Positionen streicht. Aber in

der Praxis beschneidet man damit die eigene Zukunft. Wenn man ein Team hat, das funktioniert und die Firmenkultur lebt, dann trennt man sich nicht von diesen wertvollen Fachkräften, sondern stattet sie mit der besten Technologie aus. Man muss sich das eigene Team eher wie eine absolute Spezialeinheit vorstellen: Die KI übernimmt die stupide Masse an Routinearbeit. Dadurch bleibt der Kopf der Mitarbeiter frei für echte Qualität. Während dein Team bereits blind funktioniert, kommt die KI wie eine überragende Ausrüstung dazu, die den Weg für dich und dein Team freihält.

KI ist also kein Werkzeug, um Mitarbeiter zu ersetzen, sondern ein Katalysator, um ihre Schlagkraft zu multiplizieren. Daher besteht die einzige richtige Entscheidung darin, deine Mitarbeiter zu befähigen, mit der KI zu arbeiten, statt wertvolles Personal abzubauen.



Der größte Fehler ist jetzt, jedem neuen Feature hinterherzujagen.



Wie stellt ein Systemhaus sicher, dass es nicht den Anschluss verliert?

Das ist das Kernproblem, vor dem wir aktuell alle stehen. Der größte Fehler ist, jedem neuen Feature hinterherzujagen. Das ist, als würde man versuchen, das Wasser aus einem Feuerwehrschauch mit dem Mund aufzufangen – man wird einfach weggeschwemmt. Wir dürfen uns als Unternehmer nicht mehr primär auf die einzelnen Tools fokussieren, sondern müssen den Fokus auf unsere Prozesse und deren Validierung legen.

Um da sicher durchzusteuern, sehe ich vor allem drei entscheidende Hebel. Zuallererst muss die Führungsebene zwingend ein fundiertes technologisches Verständnis aufbauen. Wenn die Geschäftsführung im Kern nicht versteht, wie diese Systeme funktionieren, basieren alle strategischen Entscheidungen am Ende nur auf reiner Spekulation.

Der zweite Hebel ist der Aufbau einer echten internen Test-Kultur. Wenn ein neues Modell auf den Markt kommt, sollten wir nicht fragen: »Was kann das theoretisch?« Stattdessen müssen wir sprichwörtlich die Stoppuhr nehmen und messen: Wie lange dauert ein Prozess manuell, wie lange liefert mit dem alten Tool und welchen echten Hebel bringt das neue System?

Wer lernt, die Effizienz von KI intern knallhart datenbasiert zu messen, behält die Kontrolle. Wir müssen immer im Hinterkopf behalten: KI ist kein Zauberstab.

Wie nimmt man seine Mitarbeiter bei diesem rasanten Tempo erfolgreich mit?

Genau das ist der dritte, entscheidende Hebel: Wir müssen im Endeffekt jeden einzelnen Mitarbeiter mitnehmen und befähigen. Packt die Techniker mit ins Boot, formuliert gemeinsam die größten Probleme im Betrieb und testet die Systeme gezielt dagegen. Aber prinzipiell kann jeder einzelne Mitarbeiter von KI unfassbar profitieren, weil du im Bestfall deine Quantität multiplizierst, indem du mehr Ergebnisse pro Zeit erzielst. Wer diese drei Schritte konsequent geht, sorgt dafür, dass die Technologie im eigenen Betrieb einen echten, messbaren Mehrwert bringt, statt einfach nur Geld zu verbrennen.



Das Interview als Podcast:

Du willst das Gespräch lieber unterwegs hören? [Hier gelangst du direkt zur Podcast-Folge.](#)

Benjamin Eidam

Benjamin Eidam ist seit 16 Jahren als selbstständiger Strategieberater, Trainer und Software-Entwickler an der Schnittstelle von Technologie und Business aktiv. Als Experte für KI-Implementierung begleitet er Führungskräfte und Systemhaus-Chefs durch ihre KI-Projekte: Er übersetzt den rasanten technologischen Fortschritt in messbare Produktivität und sorgt dafür, dass aus vagen KI-Chancen echte wirtschaftliche Ergebnisse entstehen.

