

Frankfurter Allgemeine PRO DIGITALWIRTSCHAFT



Ausgabe 38 / 2024

**AI ist der Anfang eines
neuen Paradigmas**

Frankfurter Allgemeine PRO DIGITALWIRTSCHAFT

Happy Birthday, Digitalwirtschaft (aka D:ECONOMY)

Eigentlich haben wir durchgearbeitet, seitdem unser Briefing vor genau einem Jahr zum ersten Mal erschienen ist.

Wir haben die rasende Entwicklung in der KI begleitet, sahen Sam Altman gehen und kommen. Haben beinahe alle KI-Modelle dieser Welt unter die Lupe genommen und mit praktischen Beispielen erklärt, was man in der (Unternehmens-)Praxis damit anstellen kann. Wir haben die Entwicklung bei den Elektroautos verfolgt und immer wieder eingeordnet, warum Deutschland hier und in Sachen KI „Gas geben“ muss. Vielleicht hat das ein wenig dazu beigetragen, dass sich langsam ein Bewusstsein für den Ernst der Lage in Europa entwickelt. Das würde uns freuen.

Wir haben die Plattformökonomie auf allen Ebenen durchleuchtet und gezeigt, welche Branchen und Berufsbilder als Nächstes von der Digitalisierung betroffen sein werden. Dabei haben wir stets eine internationale Perspektive beibehalten und nicht alles durch eine deutsche oder europäische Brille gesehen.

Also feiern wir heute unseren ersten Geburtstag! Aber es gibt noch mehr Grund zur Freude, und darauf sind wir besonders stolz: Die PRO-Familie hat gerade Zuwachs bekommen. FAZ PRO Weltwirtschaft und FAZ PRO Einspruch sind an den Start gegangen und bilden jetzt mit der FAZ PRO Digitalwirtschaft eine Kleinfamilie. Weitere PRO-Projekte sind in der Pipeline, um Sie – liebe Leserinnen und Leser – jede Woche auf den Punkt und in die Tiefe über ein Thema zu informieren.

Als FAZ PRO Digitalwirtschaft werden wir in den kommenden Wochen die zweite Stufe unseres Informationsangebotes starten. Konkret werden wir mit Live-Sessions beginnen, in die sich unsere Leser einwählen können, um mit uns und unseren Gästen die zentralen Fragen der Digitalisierung zu diskutieren. Zusätzlich veranstalten wir am 6. und 7. November den ersten KI-Kongress in unserem Hause. Die Vorbereitungen laufen auf Hochtouren, gute Redner haben schon zugesagt. Wir würden uns freuen, viele unserer Leser in der Pariser Straße 1 in Frankfurt begrüßen zu dürfen. Wir möchten ins Gespräch kommen. Und wir versprechen: Es lohnt sich.

Bei all der Feierei ist der erste Geburtstag natürlich auch eine Gelegenheit, um Danke zu sagen. Allen Beteiligten in der Redaktion und im Verlag, unseren freien Mitarbeitern und natürlich unseren Fachleuten, die uns mit ihren Gastbeiträgen so wunderbar unterstützt haben: DANKE!

Mit der aktuellen Ausgabe sind wir wieder mittendrin im Getümmel. In der KI war in den vergangenen Tagen Open AI o1 das bestimmende Thema. Wir haben dort mal unter die Motorhaube geschaut und erklären, was das neue Modell besser kann als seine Vorgänger und Wettbewerber. Und wir analysieren, wie die Reise zur allgemeinen Künstlichen Intelligenz weitergeht.

Wir würden uns freuen, wenn Sie uns auf dieser Reise durch die digitale Wirtschaft auch im kommenden Jahr weiter begleiten würden. Bleiben Sie uns gewogen!

Und wenn Sie uns Anregungen oder Kritik zukommen lassen wollen: Nur zu! Unter digitalwirtschaft@faz.de sind wir per Mail erreichbar - und antworten sogar.

Mit den besten Grüßen

Ihr Digitalwirtschaft-Team (okay, an den Namen müssen wir uns auch erst noch gewöhnen)

Mandy Jarry, Nina Müller, Holger Schmidt und Johannes Winkelhage

Künstliche Intelligenz

- 4 o1 ist der Anfang eines neuen Paradigmas
- 10 Die J-Kurve der KI-Revolution: „Das beste Jahrzehnt der Menschheitsgeschichte“
- 14 Wie KI militärisch verantwortungsvoll eingesetzt werden kann
- 18 Microsoft bringt zweite Welle seines Copiloten
- 20 Schon 15 Prozent der Marketing-Fachleute können ohne KI nicht mehr arbeiten
- 22 Die beliebtesten KI-Anwendungen: Viel Bewegung im Markt
- 24 Must Read

Mobility

- 26 BYD baut Dominanz auf dem chinesischen Elektroautomarkt aus
- 29 Must Read

Prompt der Woche

- 30 o1 verändert Kulturtechnik des Prompts: Es wird einfacher und präziser

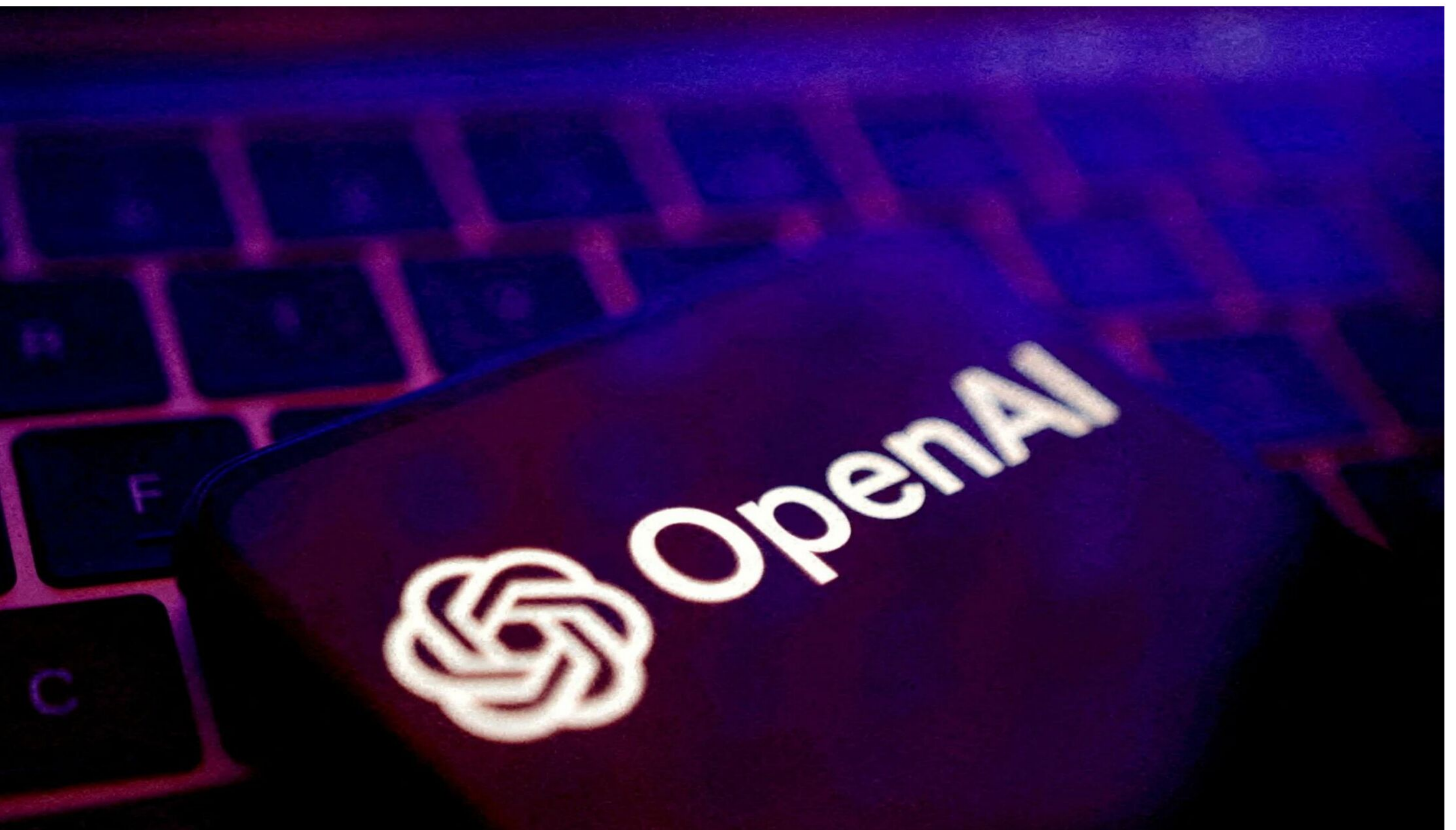
Gadgets

- 34 Videogeneratoren arbeiten immer präziser

Zitat der Woche

- 37 Clemens Fuest, Präsident des Ifo Instituts in München

o1 ist der Anfang eines neuen Paradigmas



Das Modell Open AI o1 gibt sich mehr „Nachdenkzeit“. (Reuters)

Das neue Modell o1 von Open AI stellt einen Bruch bei den LLMs dar. Statt alle Rechenleistung in das Training zu stecken und die Nutzung so schnell und günstig wie möglich zu machen, ist das Modell auf die Lösung komplexer Aufgaben optimiert. Das bedeutet langsamere und teurere Antworten. Das bedeutet: Wir haben erstmals ein LLM, das nicht stur sequenziell antwortet. Und es bedeutet, dass wir eine neue, vielversprechende Architekturnrichtung bei den LLMs sehen.

Es gibt zwei zeitliche Perioden, in denen die großen KI-Sprachmodelle Rechenzeit in Anspruch nehmen. Diese Perioden im Lebenszyklus eines LLMs könnten nicht unterschiedlicher sein. Die erste Periode ist der Modellaufbau, den man gemeinhin Training nennt. In dieser Zeit entstehen aus der Verarbeitung großer Datenmengen die internen Token-Strukturen der Modelle. Diese sind die Basis für die Modelle, Muster und Zusammenhänge in der Sprache erkennen und Sprache generieren zu können. Die zweite Periode ist die sogenannte Inferenz. Inferenz ist ein anderes Wort für Schlussfolgerung oder Konklusion. Im Grunde ist die Inferenz der Moment, in dem das LLM die im Training aufgebauten internen Strukturen anwendet, um eine Eingabe zu verarbeiten und eine hoffentlich sinnvolle Ausgabe zu produzieren.

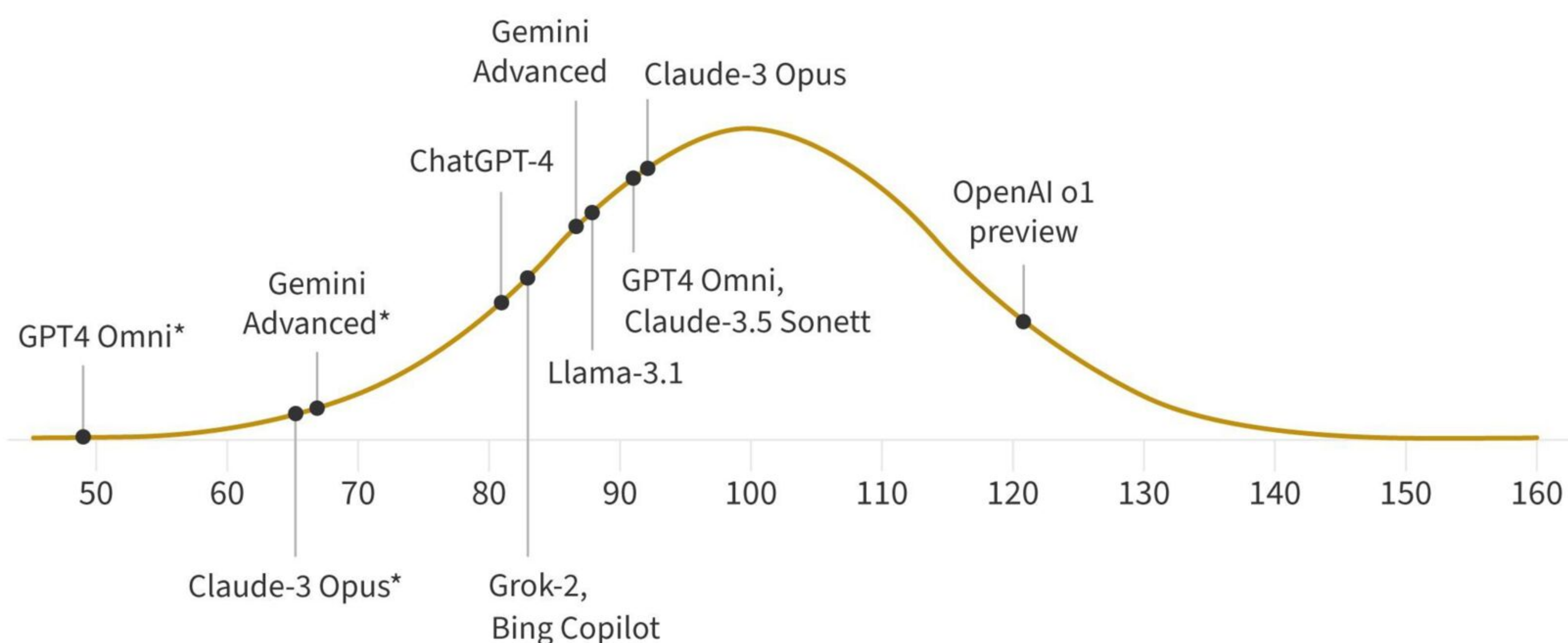
„Herkömmliche“ LLMs: Training wird teurer, Inferenz günstiger

Ein Vergleich aus der klassischen Industrie: Die Trainingsphase ist wie Konzeption und Aufbau einer neuen Fabrik mit neuen Produktionsanlagen. Inferenzen sind die nach erfolgreichem Aufbau stattfindende Produktion in dieser Fabrik, also die eigentliche Nutzung der Produktionsanlagen.

Bis dato hat sich die Branche bei der Verbesserung der Modelle auf die Trainingsphase konzentriert. Diese wurde über die Jahre länger und teurer, weil die Modelle größer wurden. Denn größere Modelle mit mehr Parametern, sprich Variablen, welche die Modelle während des Trainings erlernen, führen in der Regel zu besseren, fähigeren Modellen.

OpenAI o1 gewinnt IQ-Test

Angaben des durchschnittlichen IQs



* (Vision)

Grafik: sacco. / Quelle: TrackingAI.org

Es gibt keine offiziellen Zahlen zu den Trainingskosten. Experten schätzen allerdings, dass das Training von Claude 3.5 zwischen 100 und 200 Millionen Dollar gekostet hat. GPT-4 soll bei ungefähr 78 Millionen Dollar gelegen haben laut unter anderem Stanford AI Index Report 2024. ¹ Gemini Ultra, der Vorgänger von Googles aktuellem Gemini 2, soll 191 Millionen Dollar im Training gekostet haben. Dario Amodei, der CEO von Anthropic, dem Unternehmen hinter Claude, prognostiziert ² öffentlich sogar, dass Trainingskosten bis 2025 oder 2026 auf bis zu zehn Milliarden Dollar pro Modell anwachsen könnten. Ob diese Aussage zutreffen wird oder heute nur potentielle neue Angreifer abschrecken soll, wird sich zeigen. Der Trend im Training ist aber deutlich: Die größten und damit in der Regel besten Modelle werden größer und damit im Training immer teurer.

Die steigenden Investitionskosten liegen bei den Modellanbietern, sie sind aber pro Modell einmalig und stellen aufgrund der wachsenden Größe zunehmend eine Markteintrittsbarriere dar. Nur wenige Unternehmen können es sich überhaupt leisten, 100 Millionen Dollar und mehr in das Training von Modellen zu investieren. Neben dem Kapital braucht man außerdem Zugang zur raren Rechenleistung und die ebenso raren Fachkräfte. In ganz Deutschland gibt es aktuell kein Unternehmen, welches das macht. In Europa gibt es lediglich Mistral aus Paris, das bei diesem Spiel mitmischen kann.

Teureres Training bedeutet nicht automatisch teurere Nutzung, also teurere Inferenz. Zwar sind größere Modelle sowohl im Training als auch in der Nutzung rechenintensiver.



Marcel Weiß

Unabhängiger Analyst
und Strategieberater
in Berlin

¹ Beitrag:
faz.net/pro24381

² Beitrag:
faz.net/pro24382

Aber wir sehen seit fast einem Jahr ebenso eine Tendenz bei den großen Modellanbietern, die Nutzung ihrer Modelle günstiger und damit attraktiver zu machen. Dieser Trend begann mit GPT-4 Turbo im November 2023, das erstmals ein Spitzenmodell mit geringeren Nutzungskosten darstellte. Seitdem sinken die Nutzungskosten bei den Topmodellen kontinuierlich.

Das ist auch naheliegend. Hohe Trainingskosten sind zwar eine Belastung für die Bilanz, aber wie angemerkt auch mit Wettbewerbsvorteilen versehen. Hohe, oder zumindest als hoch wahrgenommene Inferenzkosten dagegen halten LLMs davon ab, sich als Produkte zu verbreiten.

o1: Ein stärkerer Blick auf die Inferenz

o1 ist in vielerlei Hinsicht ein signifikanter Bruch mit dem oben beschriebenen LLM-Trend. Die Inferenz dauert bei o1 länger. Das Modell ist also gefühlt langsamer. Sehr viel langsamer. Open AI vermenschlicht die längere Rechenzeit mit „Nachdenken“. Aber warum stellt o1 einen Bruch dar? Erstens, das Modell ist nicht für regelmäßige Allerweltsanfragen optimiert wie „formuliere diese E-Mail um in einen professionelleren Ton“. Die im Vergleich zu anderen Modellen nun längere und teure „Nachdenkzeit“ verschafft o1 neue Fähigkeiten. Es ist besser in der Erfüllung von Logikaufgaben, etwa Mathematik oder Programmierung, als jedes andere Modell. Gleichzeitig ist es nicht besser und oft sogar schlechter bei Textformulierungen als klassische LLMs wie Claude oder GPT-4o.

o1 ist erstmals ein LLM, das besser komplexe Aufgaben erfüllen kann als einfache Aufgaben, selbst wenn man als Nutzer die Aufgaben versehentlich in den gleichen Bereich stecken würde. Gibt man o1 eine einfache Aufgabe, warnt Open AI, kann es passieren, dass das Modell zu viel über die Lösung „nachdenkt“ und das Ergebnis verkompliziert. Die LLM-Landschaft als Ganzes ist nicht intuitiv, und mit o1 verstärkt sich dieser Umstand.

o1 stellt zweitens einen Bruch dar, weil das Modell sehr deutlich zeigt, dass das Inkaufnehmen von erhöhter Inferenzzeit neue Optionen offenbart. Bis dato war die einzige Achse für Durchbrüche bei LLMs auf der Trainingsebene. Sei es mehr

Rechenleistung, seien es mehr oder bessere Daten oder andere Architekturansätze, alles konzentrierte sich auf die Trainingsbeziehungswise Aufbauphase der Modelle. Die Inferenzzeit wird mit o1 vom lästigen Kostenfaktor erstmals zum potentiellen Wegbereiter neuer Ansätze bei den Sprachmodellen.

Vorausgesetzt, die User bringen etwas Geduld mit. Die maximale Rechenzeit zwischen Input und generiertem Output, bevor das Modell abbricht, scheint bei o1 bei etwas mehr als 3 Minuten zu liegen.

„Denke Schritt für Schritt“ als Modellarchitektur

Aber warum würde o1 abbrechen?
Was passiert hier genau?

Hier kommen wir zum dritten Aspekt, warum o1 einen Bruch bei den LLMs darstellt. LLMs arbeiten bis dato strikt pfadabhängig. Das heißt, sie analysieren die Eingabe und beginnen dann, „vorherzusagen“, welche Wörter der Reihe nach als Antwort auf den Input am wahrscheinlichsten sind. Aus diesem Vorgehen entstand im vergangenen Jahr die irreführende Bezeichnung des „stochastischen Papageis“, welche die Komplexitätsstufe von LLMs und die daraus folgende Outputqualität ausblendete. Fehler in der Ausgabe von LLMs entstehen nicht nur, aber auch aus der sequenziellen Erstellung der Sprachausgabe. Ein einmal erstellter Token (Wörter oder Wortteile) legt fest, aus welcher Richtung die nachfolgenden Token kommen können.

In einfachen Worten bedeutet das, einmal falsch abgebogen und das LLM rennt für den Rest des Outputs in die falsche Richtung. User haben diese Pfadabhängigkeit mit ein paar Prompting-Tricks ein bisschen abmildern können. „Durchdenke deine Antwort Schritt für Schritt“ im Prompt und ähnliche Ansätze für Gedankenketten scheinen LLMs in eine Richtung zu schubsen, die eine systematischere Ausgabe zu befördern scheint. Damit lassen sich spürbar bessere Ergebnisse erzeugen. Es mindert aber wie die Modellgröße lediglich das Problem, statt es zu eliminieren. Große Modelle reduzieren die Wahrscheinlichkeit eines falschen Tokens, aber auch hier verschwindet sie nicht.



Open-AI-CEO Sam Altman (AFP)

Auf o1 trifft die Bezeichnung des stochastischen Papageis nun noch weniger zu. Denn hier ist Open AI erstmals über diese sequenzielle Erzeugung in der Inferenz hinausgegangen. Open AI sagt nicht, wie genau sie o1 gebaut haben. Aber so viel wissen wir: Semafor **3** berichtete im Januar 2023, dass Open AI weltweit über 1000 Softwareentwickler als Subunternehmer anstellte, um mehrstufige Programmierprojekte in einzelne Stufen herunterzubrechen. Das Ergebnis dieser Bemühungen dürften Datensätze sein, die LLMs im Training helfen, Patterns zu erstellen, um mehrstufige Aufgaben zu erledigen.

Im Mai 2023 hat Open AI ein Paper mit dem Titel Let's Verify Step by Step **4** veröffentlicht. Darin beschreiben sie unter anderem, wie sie Daten-Etikettieren, wie im Semafor-Artikel angedeutet, Schritt-für-Schritt-Lösungen für Mathematikaufgaben präsentieren und diese die einzelnen Schritte bewerten. Das Ziel des Papers: ein „process-supervised Reward Model“ (PRM) zu bauen. Das PRM soll die Wahrscheinlichkeit für die Richtigkeit eines einzelnen Schritts nach dessen letztem erstellten Token überprüfen.

In Summe lässt sich also Folgendes festhalten: o1 wurde mit Blick auf das Lösen mehrstufiger Logikprobleme trainiert. o1 wurde mit diesem Fokus so gestaltet **5**, mehrere Problemlösungsabläufe innerhalb der Inferenzzeit zu erstellen, dabei jeden Schritt einzeln zu

bewerten und so festzustellen, wenn es falsch „abgebogen“ ist und neu ansetzen muss.

o1s „Nachdenkzeit“ ist länger, weil das Modell mehrere Lösungsrichtungen durchläuft und Fehler selbständig erkennen kann. Deshalb kann es passieren, dass o1 die Berechnung abbricht. Das Modell stellt fest, dass das bisherige Ergebnis falsch ist, die ihm zugeteilte maximale Rechenleistung aber abgelaufen ist.

Wo die Reise für Open AI hingeht

Open AI hat zehn Millionen Abonnenten. Das erst ein Jahr alte, höherpreisige Enterprise-Angebot für Unternehmen liegt bereits bei einer Million Abonnenten. o1 bietet hier enormes Potential. Die Lösung mehrstufiger Herausforderungen erhöht die Einsatzarten. Besonders in der Forschung dürfte o1 zum Einsatz kommen. Aber auch Programmierung mit einem LLM erreicht hier ein neues Level. Man denke an unseren Text zur KI-gestützten Programmierung **6**. Open AI kann o1 auch mit seinen anderen Modellen koppeln. o1 erstellt einen Arbeitsweg, und die günstigeren Modelle machen die „Laufarbeit“. Die größte Herausforderung bleibt auf der eigentlichen Produktseite des Modells. Open AI muss besser kommunizieren, was mit diesem Modell erreicht werden kann und was nicht. LLMs sind schwer zu erfassen, und o1 scheint diese schwere Einordbarkeit noch einmal zu verstärken.

3 Beitrag:
faz.net/pro24383

4 Beitrag:
faz.net/pro24384

5 Beitrag:
faz.net/pro24385

6 Beitrag:
faz.net/pro24386

o1 zeigt aber gleichzeitig, dass die Zeit autonomer und semiautonomer Agenten nah ist. o1 kann die Grundlage für erste, gut arbeitende Agenten sein.

Interessant ist, dass Open AI die zusätzliche Rechenzeit in der Inferenz für API-Nutzer in unsichtbare Tokens **7** gesteckt hat. API-Nutzung wird errechnet aus Inputtokens und Outputtokens. Nun kommt eine unvorhersehbare Variable bei den Kosten hinzu. Open AI sagt nicht, warum sie das tun. Aber unsere Vermutung ist, dass Open AI verhindern will, dass auf der Basis von o1 andere Modelle trainiert werden. Diese Nutzung ist zwar laut AGBs untersagt, findet aber trotzdem über die API statt. o1 zeigt dem Nutzer nicht die Schritte, die es vor dem Output unternommen hat. Man sieht nicht, welche Richtungen das System genommen und verworfen hat. Alle diese Berechnungen kosten Geld, wollen aber von Open AI nicht offengelegt werden.

Wo die Reise für LLMs hingeht

Wenn uns die aktuellen LLMs auf GPT-4-Basis etwas lehren, dann dass Open AI in der Regel nur die Ersten, aber nicht die Einzigen bei LLM-Durchbrüchen sind. Wir werden in den nächsten Monaten weitere Modelle sehen, die ähnlich wie o1 arbeiten. Open-Source-Modelle von Meta oder Mistral könnten die internen Vorgänge im Gegensatz zu o1 offenlegen, was weitere Einsätze erschließen dürfte.

KI-Agenten werden nun ebenso greifbar wie ausgefuchste Modellmixe mit Arbeitsteilung zwischen den LLMs, wie wir es hier **8** beschrieben haben.

Fazit

o1 zeigt, dass das Ende der Fahnenstange bei der LLM-Entwicklung noch weit weg ist.

Mit stärkerem Blick auf die Inferenz bei dieser neuen, nachdenklichen Modellart werden allerdings auch die Chips und die Rechenleistung, **9** die uns zur Verfügung stehen, noch wichtiger.

Zu guter Letzt zeigt o1 auch, wie Regulierung schneller technologischer Entwicklung hinterherhinkt. Der AI Act der EU hat sich auf die Rechenleistung bei der Trainingsphase eingeschossen, um „gefährliche“ und „ungefährliche“ KI voneinander unterscheiden zu können. Der AI Act legt einen Schwellenwert von 1025 FLOPs für die Rechenleistung fest, die zum Training von KI-Modellen verwendet wird. Modelle, die diesen Wert überschreiten, werden als Systeme mit „hohem systemischen Risiko“ eingestuft.

Mit einem simplen, leichten Prioritätenschwenk hat o1 diesen bereits fragwürdigen Regulierungsansatz noch fragwürdiger gemacht. Denn nach o1 werden wir nun in naher Zukunft auch Modelle sehen, die weitaus weniger im Training benötigen, dafür mehr Rechenzeit in der Inferenz nutzen und in ihren Fähigkeiten alles übersteigen werden, was wir heute kennen. Auch in Open Source. Und auch lokal.

7 Beitrag:
faz.net/pro24387

8 Beitrag:
faz.net/pro24388

9 Beitrag:
faz.net/pro24389

Frankfurter Allgemeine
ZEITUNG FÜR DEUTSCHLAND

Freiheit beginnt im Kopf.

Die Frankfurter Allgemeine steht für die Freiheit im Denken. Sie steht für die Vielfalt der Perspektiven. Für die Kraft der Fakten. Mit Tiefe und Intelligenz, mit sachlichem Blick und besonnenem Stil analysiert die Frankfurter Allgemeine das Geschehen und ordnet es ein. Demokratie beruht auf Freiheit. — Freiheit beginnt im Kopf.

Mehr erfahren auf **[freiheitimkopf.de](https://www.freiheitimkopf.de)**

„Das beste Jahrzehnt der Menschheitsgeschichte“



Erik Brynjolfsson warnt vor dem, was er als „Konzept der Turing-Falle“ bezeichnet. (Picture Alliance)

Stanford-Ökonom Erik Brynjolfsson erforscht seit Jahrzehnten die ökonomischen Auswirkungen digitaler Technologien. „Die KI ist eine der größten und effektivsten Technologien zur Veränderung der Arbeitsweise, die je erfunden wurden“, sagt er. Doch die Einführung von KI-Tools reicht allein nicht aus.

In der sich rasant entwickelnden Welt der Technologie steht die generative Künstliche Intelligenz im Mittelpunkt einer neuen digitalen Revolution. Doch wie bei jeder bahnbrechenden Innovation stellt sich die Frage: Wie können Unternehmen diese Technologie effektiv nutzen, um echte Produktivitätssteigerungen zu erzielen? Die Antwort liegt in einem Phänomen, das Fachleute als „Produktivitäts-J-Kurve“ bezeichnen.

Die J-Kurve beschreibt den typischen Verlauf der Produktivität bei der Einführung neuer Technologien, erklärt Erik Brynjolfsson, Ökonom und Professor an der Stanford-Universität, dieses Konzept in einem Podcast. Zunächst kann es zu einem Rückgang oder einer Stagnation der Produktivität kommen, da Unternehmen Zeit und Ressourcen investieren müssen, um ihre Prozesse anzupassen, Mitarbeiter umzuschulen und möglicherweise sogar ihre Geschäftsmodelle zu überdenken. Dies entspricht dem abfallenden Teil des „J“.

Früher dauerte die Beschleunigung Jahrzehnte – heute geht es schneller

Erst wenn diese Anpassungen greifen, setzt der aufsteigende Teil der Kurve ein, und die Produktivität beginnt deutlich zu steigen. Brynjolfsson betont, dass dieser Prozess bei früheren technologischen Revolutionen oft Jahre oder sogar Jahrzehnte dauerte. Bei der generativen KI erwartet er jedoch eine Beschleunigung dieses Zyklus.

Die Gründe für diese Beschleunigung sind vielfältig. Zum einen ist die generative KI laut Brynjolfsson „eine der größten und effektivsten Technologien zur Veränderung der Arbeitsweise, die je erfunden wurden“. Zum anderen ist sie vergleichsweise einfach zu implementieren, da normale Mitarbeiter oft schon nach wenigen Stunden damit arbeiten können. Hinzu kommt, dass die notwendige Infrastruktur – wie Cloud-Computing und schnelle Internetverbindungen – bereits weit verbreitet ist, was eine rasche Skalierung ermöglicht.

„Großartige Technologie allein reicht nicht aus“

Diese Faktoren erklären, warum Anwendungen wie ChatGPT in nur 60 Tagen 100 Millionen Nutzer erreichen konnten – ein Tempo, das bei

früheren Technologien undenkbar war. Dennoch warnt Brynjolfsson davor, zu glauben, dass die bloße Einführung von KI-Tools ausreicht, um Produktivitätssprünge zu erzielen. „Großartige Technologie allein reicht nicht aus. Was man wirklich braucht, ist die Aktualisierung der Geschäftsprozesse, die Umschulung der Belegschaft und manchmal sogar die Änderung der Geschäftsmodelle und der Organisation in großem Umfang.“

Diese umfassende Transformation sei der Schlüssel, um die anfängliche Produktivitätsdelle zu überwinden und in die Wachstumsphase der J-Kurve einzutreten. Unternehmen, die dies verstehen und umsetzen, werden in der Lage sein, die Vorteile der generativen KI voll auszuschöpfen.

Ein weiterer wichtiger Aspekt sei die Auswirkung der generativen KI auf die Arbeitswelt. Entgegen weit verbreiteter Befürchtungen sieht er die Technologie nicht primär als Jobkiller, sondern als Werkzeug zur Erweiterung menschlicher Fähigkeiten. Seine Forschung zeigt, dass fast alle Berufe von KI beeinflusst werden, aber kaum einer vollständig automatisiert wird. Stattdessen werden viele Aufgaben innerhalb der Berufe durch KI unterstützt oder verbessert.

Diese Sichtweise wird durch eine Fallstudie in einem Callcenter untermauert. Dort führte die Einführung eines KI-Systems zu dem, was Brynjolfsson als „dreifachen Gewinn“ bezeichnet:

1. Die Produktivität der Mitarbeiter stieg erheblich, in manchen Fällen um bis zu 35 Prozent.
2. Die Kundenzufriedenheit verbesserte sich deutlich, was durch Sentiment-Analysen der Gesprächsprotokolle bestätigt wurde.
3. Die Mitarbeiterzufriedenheit nahm zu, und die Fluktuation ging zurück.

Besonders bemerkenswert war, dass die am wenigsten qualifizierten Mitarbeiter am meisten von der KI-Unterstützung profitierten. Das System fungierte als Coach und ermöglichte es ihnen, schnell von den besten Praktiken im Unternehmen zu lernen. Dies zeigt das Potential von KI, nicht nur zur Effizienzsteigerung, sondern auch zur Kompetenzentwicklung beizutragen.



Holger Schmidt

Redaktionsleiter
„Newsletter und
Verticals“

Das „Konzept der Turing-Falle“

Brynjolfsson warnt jedoch vor dem, was er als „Konzept der Turing-Falle“ bezeichnet. Benannt nach dem berühmten Turing-Test, beschreibt dieses Konzept die Gefahr, dass Unternehmen und Entwickler sich zu sehr darauf konzentrieren, menschliche Arbeit durch Maschinen zu ersetzen, anstatt Mensch und Maschine synergetisch zu ergänzen. Er argumentiert, dass der Versuch, Menschen einfach zu ersetzen, nicht nur die einzigartigen Stärken der KI vernachlässigt, sondern auch zu einer dystopischen Gesellschaft führen könnte, in der der Wohlstand sich bei wenigen Technologiebesitzern konzentriert.

Stattdessen plädiert Brynjolfsson für einen Ansatz, bei dem KI genutzt wird, um menschliche Fähigkeiten zu erweitern und zu ergänzen. Dies ermöglicht es Menschen, Dinge zu tun, die sie zuvor nicht konnten, und schafft so nicht nur mehr Wert, sondern verteilt diesen auch gerechter. Für Unternehmen bedeutet dies, dass sie bei der Implementierung von KI-Systemen darüber nachdenken sollten, wie sie neue Werte schaffen und ihre Mitarbeiter befähigen können, anstatt nur bestehende Prozesse zu automatisieren.

Trotz aller Herausforderungen bleibt Brynjolfsson optimistisch für die Zukunft. Er bezeichnet sich selbst als „achtsamen Optimisten“ und betont, dass Technologie nicht unser Schicksal bestimmt, sondern wir die Macht haben, sie zu gestalten. Mit dem richtigen Ansatz und einer verantwortungsvollen Nutzung von KI sieht er das Potential für „das beste Jahrzehnt in der Menschheitsgeschichte“.

Handlungsempfehlungen für Führungskräfte

Für Unternehmen und Führungskräfte ergeben sich aus diesen Erkenntnissen klare Handlungsempfehlungen:

1. Bereiten Sie sich auf die J-Kurve vor: Verstehen Sie, dass die Einführung von KI zunächst Investitionen und Anpassungen erfordert, bevor sich die Vorteile voll entfalten.
2. Denken Sie ganzheitlich: KI ist nicht nur ein technologisches Upgrade, sondern erfordert eine Neugestaltung von Prozessen, Fähigkeiten und manchmal ganzen Geschäftsmodellen.
3. Fokussieren Sie auf Erweiterung, nicht Ersetzung: Suchen Sie nach Möglichkeiten, wie KI die Fähigkeiten Ihrer Mitarbeiter erweitern und verbessern kann, anstatt sie zu ersetzen.
4. Investieren Sie in Weiterbildung: Die Fallstudie des Callcenters zeigt, wie KI zur Kompetenzentwicklung beitragen kann. Nutzen Sie dieses Potential, um Ihre Belegschaft zukunftsfähig zu machen.
5. Messen Sie den wahren Wert: Berücksichtigen Sie bei der Bewertung von KI-Initiativen nicht nur harte Finanzkennzahlen, sondern auch weichere Faktoren wie Kundenzufriedenheit und Mitarbeiterengagement.
6. Bleiben Sie flexibel und lernbereit: Die KI-Landschaft entwickelt sich schnell. Kultivieren Sie eine Unternehmenskultur, die offen für Veränderungen ist und kontinuierliches Lernen fördert.

Indem Unternehmen diese Prinzipien beherzigen, können sie die Herausforderungen der KI-Revolution meistern und die Chancen nutzen, die sich daraus ergeben. Die J-Kurve mag zunächst steil erscheinen, aber mit der richtigen Strategie und Einstellung können Organisationen den Aufstieg bewältigen und in eine produktivere und innovativere Zukunft steuern, erklärt der Stanford-Professor.

**75****Jahre**
Frankfurter
Allgemeine

Zukunft braucht Denkanstöße — wir liefern sie seit 1949

Feiern Sie mit uns 75 Jahre Qualitätsjournalismus und lesen Sie 6 Monate die Frankfurter Allgemeine Zeitung und Sonntagszeitung im **Jubiläumspaket ab 75 €**.

Unser Jubiläumsgeschenk für Sie:

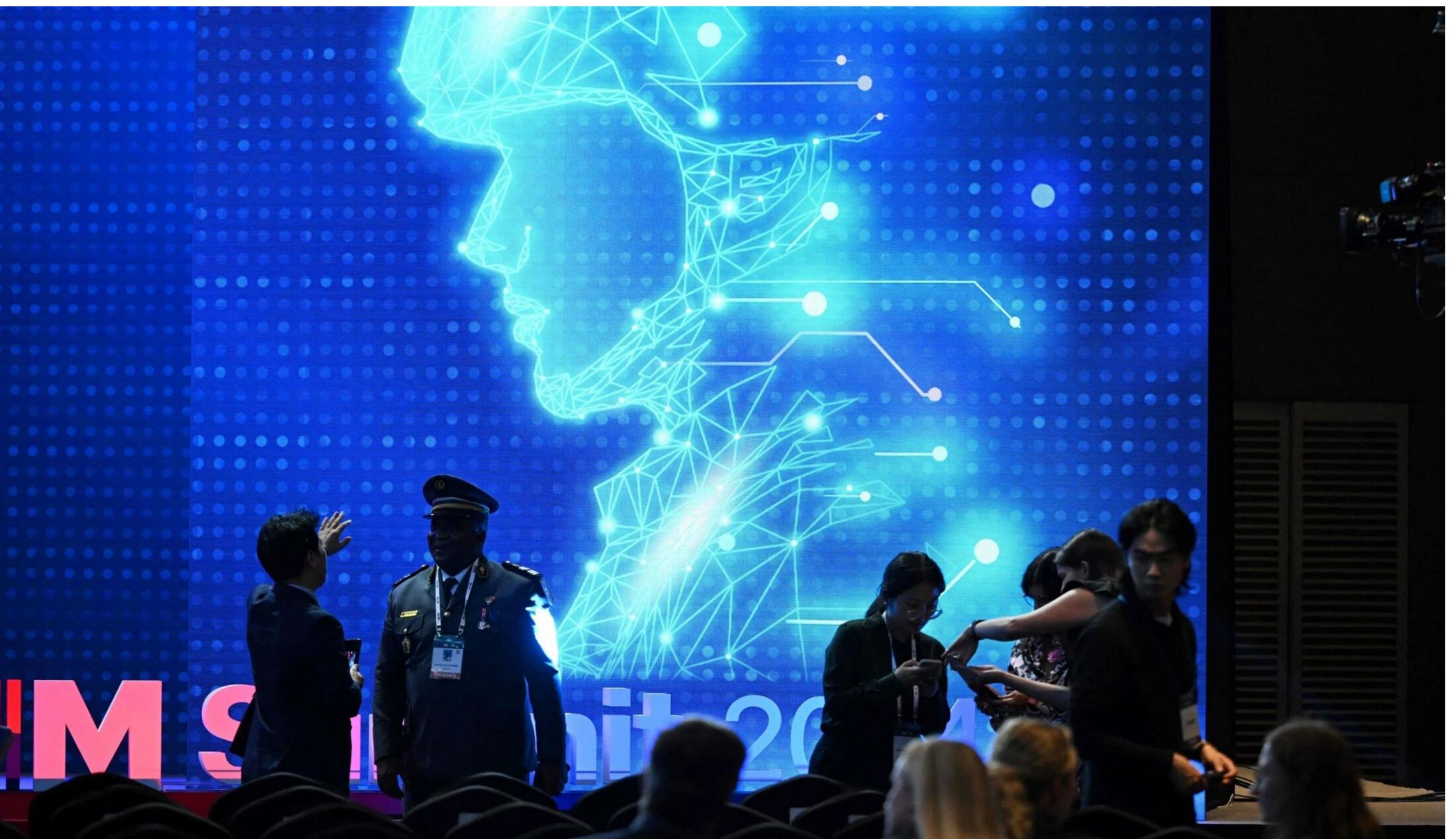
- ✓ **Bis zu 75 % Rabatt:** Lesen Sie die F.A.Z./F.A.S. im Jubiläumspaket für 6 Monate ab 75 €.
- ✓ **75 Preise:** Sichern Sie sich die Gewinnchance auf 75 tolle Preise wie den **vollelektrischen Crossover Kia Niro EV***, eine exklusive Malediven-Reise ins Velassaru Maldives sowie 2 x 60.000 Miles & More Meilen oder ein ETF-Depot im Wert von 7.500 € bei Trade Republic.
- ✓ **Initiative „Zukunft gestalten“:** Erfahren Sie mehr über die bedeutenden Themen unserer Zeit in der redaktionellen Serie der F.A.Z. zum Jubiläumsjahr.

**75 Jahre –
75 Preise**Partner von
Miles & More**TRADE REPUBLIC*** Kia Niro EV 64,8-kWh-Batterie (Strom/Reduktionsgetriebe); 150 kW (204 PS); Stromverbrauch kombiniert 16,2 kWh/100 km; CO₂-Emission kombiniert 0 g/km; CO₂-Klasse A.

Jetzt Jubiläumsangebot sichern: **(069) 75 91-33 59** oder **faz.net/75jahre**

PR24039

Wie KI militärisch verantwortungsvoll eingesetzt werden kann



Es muss eine verstärkte Aufklärung geleistet werden, damit KI-Systeme für den militärischen Einsatz so entwickelt werden können, dass sie eine sichere Nutzung ermöglichen. (AFP)

Der Einsatz von KI in militärischen Anwendungen und autonomen Waffensystemen führt aktuell zu umfangreichen Diskussionen in der Presse, Politik und Wissenschaft. Hier der aktuelle Stand der Dinge.

Im vergangenen Jahr fand der erste Responsible AI for the Military Domain Summit (REAIM) in den Niederlanden statt. Es folgte ein zweiter REAIM Summit in Seoul ¹, an dem Vertreter aus über 90 Staaten die Auswirkungen von KI auf die internationale Sicherheitslage und den Einsatz von autonomen Waffensystemen (AWS) erörterten.

Das öffentliche Meinungsbild zu dieser Frage ist noch vielfach durch von Filmwerken befeuerte Ängste vor einem menschlichen Kontrollverlust über vermeintliche „Killer-Roboter“ und einem mangelnden Verständnis von KI geprägt. Eine im Mai von der Bundesregierung im Deutschen Bundestag beantwortete Kleine Anfrage ² ergab, dass sich die Bundesregierung für die Ächtung vollautonomer letaler Waffensysteme, die außerhalb menschlicher Kontrolle operieren, sowie für die Regulierung von Waffensystemen mit autonomen Funktionen einsetze.

Der kürzlich verabschiedete EU AI Act gilt aber unter anderem ausdrücklich nicht für KI-Systeme, wenn und soweit sie ausschließlich für militärische Zwecke, Verteidigungszwecke oder Zwecke der nationalen Sicherheit in Verkehr gebracht, in Betrieb genommen oder verwendet werden. Potentielle Gegner der demokratischen Staatengemeinschaft arbeiten aktuell jedoch aktiv an Entwicklung und Einsatz von AWS.

Die Folgen einer Überregulierung wären fraglos fatal, da sie die Verteidigungsbereitschaft westlicher Demokratien gegenüber autokratischen Systemen und Terrororganisationen schwächen würde, die vor einem Missbrauch von KI zur Durchsetzung ihrer Interessen nicht zurückschrecken und sich auch nicht an Verbote halten werden, mit denen sich der technische Fortschritt nicht aufhalten lässt.

Wir haben auf dem REAIM Summit die Breakout Session „Mission Possible: Deploying responsible and effective autonomous weapon systems without overregulation“ veranstaltet, um die Debatte mitzugestalten. Dabei haben wir die Gefahren einer Überregulierung von AWS aufgezeigt und eine praktikable Alternative skizziert. Angesichts der schnellen Fortschritte im Bereich KI und AWS sind für uns die folgenden Fragestellungen entscheidend:

1. Braucht es zusätzliche Regulierung für militärische KI (und wenn ja, welche Regulierung)?
2. Welche Instrumente gewährleisten die Einhaltung des Humanitären Völkerrechts (International Humanitarian Law, kurz IHL) durch AWS?
3. Wie beeinflussen Beweislasten den Einsatz von AWS?

Im Rahmen der Veranstaltung haben wir unsere Positionen zu diesen Fragestellungen vorgestellt und sie mit den Teilnehmern diskutiert. Wir sind der Ansicht, dass Überregulierung im Stil des EU AI Act vermieden werden muss und bessere Risikomanagementstrategien essentiell sind. Unserer Ansicht ³ nach sind KI-Systeme mit der passenden Methodik grundsätzlich sicher und transparent.

Einsatz von AWS erfolgt nicht im rechtsfreien Raum

Wie aus dem Kreis der Gipfelteilnehmer zu Recht geäußert wurde, muss deshalb zunächst eine verstärkte Aufklärung darüber geleistet werden, dass KI-Systeme für den militärischen Einsatz so entwickelt werden können, dass sie eine sichere Nutzung ermöglichen. Der Einsatz von AWS erfolgt unserer Ansicht nach nicht etwa in einem rechtsfreien Raum, da das geltende Völkerrecht auch für neue AWS gilt und für deren Einsatz bewährte Grundsätze bereithält, die von den Vertragsstaaten entgegen anderer Darstellung gerade auch bei der Nutzung von KI zur Verteidigung eingehalten werden müssen und eingehalten werden können.

Um eine informierte Entscheidung über die Frage der Erforderlichkeit einer weiteren Regulierung von AWS treffen zu können, wird man sich unserer Meinung nach in der sicherheitspolitischen Diskussion jedoch von der Vorstellung lösen müssen, dass in allen realistischen und militärisch sinnvollen Anwendungsfällen für AWS von einer Überlegenheit des menschlichen Geistes und menschlichen Reaktionsvermögens auszugehen sei, die eine fortlaufende Überprüfung sämtlicher Aktionen solcher Systeme durch einen menschlichen Entscheidungsträger („human in the loop“) erfordere.



Prof. Dr. Patrick Glauner

Professor für KI an der TH Deggendorf und hat als Sachverständiger die Parlamente von Deutschland, Frankreich und Luxemburg zu den sich aus KI ergebenden politischen und rechtlichen Fragestellungen beraten



Dr. Andreas Leupold

Mitglied der Arbeitsgruppe Recht der Plattform Industrie 4.0 und ist als Rechtsanwalt mit den Rechtsfragen der KI und neuer Technologien befasst

¹ Beitrag: faz.net/pro243810

² Beitrag: faz.net/pro243811

³ Beitrag: faz.net/pro243812

Würde man dem nachgeben, so würde dies auf ein De-facto-Pauschalverbot von AWS hinauslaufen, die sich nach der aktuellen Definition der Expertengruppe der Vereinten Nationen zur UN-Waffenkonvention gerade dadurch auszeichnen, dass sie nach ihrer Aktivierung ohne weitere Intervention durch einen menschlichen Nutzer selbständig Ziele identifizieren, auswählen und angreifen können.

Dann könnten AWS auch nicht mehr in Gebieten eingesetzt werden, in denen die Netzwerkinfrastruktur durch Angriffe des Gegners so beschädigt oder funktionsuntüchtig gemacht wurde, dass keine Verbindung zu militärisch relevanten KI-Systemen mehr aufgebaut werden kann, obwohl eine Abwehr weiterer Angriffe eben dies erfordern würde. Auch dort, wo schon heute große Datenmengen in Echtzeit ausgewertet werden müssen, um Angriffen wirksam begegnen zu können, und dort, wo Entscheidungen von KI-Systemen in Bruchteilen von Sekunden getroffen werden müssen, weil jede Entscheidung durch einen Menschen zu spät kommen würde, um die Vernichtung von Menschenleben oder Angriffe auf kritische Infrastrukturen (KRITIS) wie die Energie- und Wasserversorgung und medizinische Versorgung verhindern zu können, werden AWS unverzichtbar sein.

EU-Produkthaftung kann Beitrag leisten

Einen wichtigen Beitrag zur Sicherheit von AWS kann unserer Meinung nach zudem das neue EU-Produkthaftungsrecht leisten, das seine Geltung ausdrücklich auf Software erstreckt und die

Fähigkeit von Produkten berücksichtigt, „nach Einsatzbeginn weiter zu lernen“.

Haftungsrechtliche Fragestellungen beim Einsatz von AWS sind somit keine unmögliche Aufgabe, denn das Produkthaftungsrecht verlangt keine absolute Fehlerfreiheit, sondern befreit den Hersteller von der Haftung, wenn der Produktfehler nach dem wissenschaftlichen und technischen Kenntnisstand zum Zeitpunkt des Inverkehrbringens, der Inbetriebnahme des Produkts oder solange sich das Produkt unter der Kontrolle des Herstellers befand, nicht entdeckt werden konnte.

Die Aufgabe, sich mit der Befürchtung einer – tatsächlich nur vermeintlich – unkontrollierbaren KI auseinanderzusetzen, hat der REAIM Summit erfüllt, zugleich aber Raum dafür gegeben, solchen Ängsten ein realistisches Bild von den Chancen und Grenzen einer Technologie entgegenzusetzen, die wie keine andere das Potential hat, künftige Angriffskriege zu verhindern und die internationale Sicherheit zu gewährleisten, wenn sie nicht Opfer einer Überregulierung wird. In ihrem Abschlusspapier haben die teilnehmenden Minister einen Vorschlag für die künftigen Rahmenbedingungen zum militärischen Einsatz von KI unterbreitet, der zeigt, dass ein verantwortungsvoller Umgang mit KI auch in diesem Bereich möglich ist. Erreicht werden soll dies mit einem Fahrplan, der auf die Einhaltung internationalen Rechts, die Verantwortlichkeit von Menschen für den richtigen Einsatz zuverlässiger und vertrauenswürdiger KI, eine angemessene menschliche Beteiligung und eine verbesserte Erklärbarkeit von KI-Systemen setzt.

Jetzt entdecken: die digitale F.A.Z. als bildstarke Edition.

Moderne, bildstarke
Aufbereitung mit ergänzenden
Multimedia-Inhalten.

Hier lesen:
faz.net/edition

Schon ab 19 Uhr
am Vorabend verfügbar.



Schriftgrößen individuell
einstellen, Merkzettel
oder Vorlesefunktion nutzen
und mehr.

Sorgfältig kuratierte
Top-Themen als Einstieg in
die tagesaktuelle Ausgabe.



Microsoft bringt zweite Welle seines Copiloten



Microsoft hat Copilot in seinem Cloud-Speicher OneDrive vorgestellt, mit dem Nutzer den Assistenten nutzen können, um Dateien in ihrem OneDrive-Repository zu finden, Zusammenfassungen zu erhalten und Dateien zu vergleichen. (Mauritius)

Microsoft hat den Copiloten früh auf den Markt gebracht, aber seitdem kaum weiterentwickelt. Die zweite Generation bringt nun Produktivitätstools für Office-Anwendungen und KI-Agenten für die Automatisierung.



Holger Schmidt

Redaktionsleiter
„Newsletter und
Verticals“

Microsoft hat den Copiloten sehr schnell nach dem Aufkommen der generativen KI gestartet und in seine Office-Produkte integriert, aber bisher kaum weiterentwickelt und an Schwung verloren. Nun soll eine zweite Welle mit neuen Funktionen für weitere Kunden sorgen.

Im Zentrum der Neuerungen steht Copilot Pages, eine Funktion, die „Multiplayer-KI-Kollaboration“ ermöglicht. Nutzer können Antworten des Copilot-Chatbots in ein editierbares Dokument überführen, um Teams die gemeinsame Arbeit an KI-generierten Inhalten in Echtzeit zu ermöglichen.

Eine weitere wichtige Innovation sind die Kopilot-Agenten. Das sind KI-Assistenten, die komplexe Aufgabenreihen automatisieren oder eigenständig E-Mail-Postfächer überwachen können. Diese Agenten versprechen, Routinearbeiten zu übernehmen und mehr Zeit für kreative und strategische Tätigkeiten zu schaffen. Diese Agenten gelten als der nächste große Schritt in der Künstlichen Intelligenz, um nicht nur einzelne Arbeitsschritte zu ersetzen, sondern mehrere Arbeitsschritte in einem Schwung auszuführen. Ein Beispiel ist die Planung einer Reise mit Flügen, Restaurants und Museen, für die die Künstliche Intelligenz alle Buchungen in einem Arbeitsschritt erledigt.

Der Copilot wird nun in die Tabellenkalkulation Excel integriert und mit der universellen Programmiersprache Python verbunden. Damit werden komplexe Datenanalysen wie Prognosen möglich – alles über natürliche Spracheingaben, ganz ohne Programmierkenntnisse. Andere Anbieter haben solche Tools schon entwickelt, aber die Integration in Excel könnte die Nutzung dieser KI-Funktionen erheblich beschleunigen.

PowerPoint Narrative Builder: Diese nun allgemein verfügbare Funktion hilft Nutzern, innerhalb weniger Minuten einen ersten Präsentationsentwurf zu erstellen. Der „Brand Manager“ stellt sicher, dass Präsentationen den Unternehmensrichtlinien entsprechen.

Teams mit verbesserter Meetinganalyse: Copilot in Teams kann jetzt sowohl das Meetingtranskript als auch den Chat analysieren, um ein vollständiges Bild der Diskussion zu liefern.

Outlook mit Posteingangspriorisierung: Die neue Funktion „Prioritize my inbox“ analysiert E-Mails basierend auf Inhalt und Kontext, um wichtige Nachrichten hervorzuheben. Der neue Copilot erstellt automatisch Zusammenfassungen und erklärt die Priorisierung.

Schließlich hat Microsoft Copilot in seinem Cloud-Speicher OneDrive vorgestellt, mit dem Nutzer den Assistenten nutzen können, um Dateien in ihrem OneDrive-Repository zu finden, Zusammenfassungen zu erhalten und Dateien zu vergleichen.

Die neuen Funktionen, die schrittweise ausgerollt werden, sollen das KI-Geschäft von Microsoft ankurbeln. Zuletzt hatte es einige Kritik gegeben, zum Beispiel von Salesforce-Gründer Marc Benioff, der Microsoft vorwarf, viele Kunden enttäuscht zu haben. Allerdings konterte das Unternehmen die Kritik nun mit neuen Zahlen, die ein starkes Wachstum suggerieren sollen. So habe Vodafone 68.000 Copilot-Lizenzen erworben, nachdem Tests drei Stunden Zeitersparnis je Woche gebracht hätten. Die Copilot-Kundenzahl sei im vergangenen Quartal um insgesamt 60 Prozent gestiegen, da sich die Zahl der Großkunden verdoppelt hätte, gab das Unternehmen bekannt.

Schon 15 Prozent der Marketing-Fachleute können ohne KI nicht mehr arbeiten

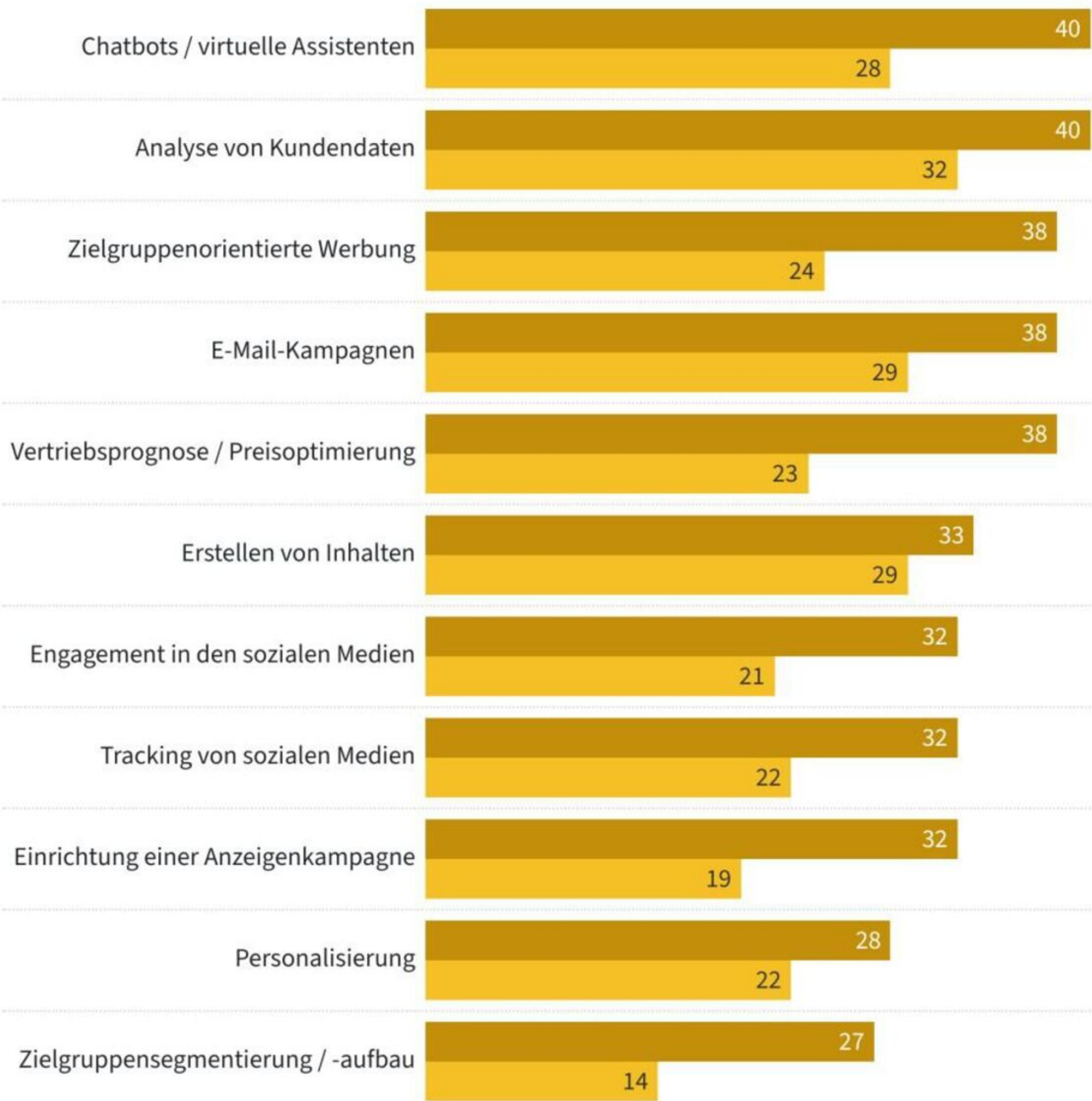
Holger Schmidt

Redaktionsleiter „Newsletter und Verticals“

KI-Einsatzfelder im Marketing

In europäischen Unternehmen in Prozent

● Unternehmen mit überdurchschnittlichem Return on Investment ● Alle anderen Unternehmen



Grafik: sacco. / Quelle: Hubspot 2024

Die Einführung von Künstlicher Intelligenz im Marketing schreitet voran. Gut jeder dritte Beschäftigte in der Branche hat die KI inzwischen fest in seinen Tagesablauf integriert.

KI wird zum Beispiel die Produktionskosten für Texte, Bilder und Videos senken oder das Geschäftsmodell von Google und damit auch das Performance-Marketing unter Druck setzen, wenn Suchmaschinen künftig Antworten statt Links liefern. Und auch die Kundenkommunikation wird sich ändern, weil Chatbots dank KI sehr viel leistungsfähiger geworden sind. Die großen Player in der Werbung, vor allem Google und Meta, sind auch führend in der Entwicklung der Künstlichen Intelligenz, geben Milliarden für neue Rechenzentren und das Training ihrer Modelle aus.

Mit neuen KI-Tools bauen sie ihre starke Stellung als Werbeplattformen aus, da die Kunden mithilfe der Tools sehr einfach neue Kampagnen erstellen können. „KI ermöglicht es Marketingteams, Kampagnen präziser zu personalisieren und die Kundenerfahrung signifikant zu verbessern“, sagt Josephine Wick Fröna, Head of Marketing DACH bei Hubspot. Tatsächlich berichten 76 Prozent der Befragten, dass KI-Lösungen bereits zu einem besseren Kundenerlebnis geführt haben.

F.A.Z. KI-Podcast – Live zum Thema „Einsatz der KI im Marketing“

Die Podcast-Hosts Peter Buxmann und Holger Schmidt sprechen mit Markus Caspari, Head of Performance Marketing bei Dentsu, über die Auswirkungen der KI auf das Marketing.

Ort: DMEXCO, 19. September 2024, 15.50 Uhr, Experience Stage

Nach einer Studie von Hubspot und LinkedIn unter 2.306 Marketingverantwortlichen in Europa setzen bereits 93 Prozent der befragten Unternehmen KI-Lösungen aktiv ein oder erforschen deren Einsatzmöglichkeiten. Die Erwartungen sind hoch: 92 Prozent der Marketingleiter sind zuversichtlich, dass ihr Team mithilfe von KI die Ziele für 2025 erreichen oder übertreffen wird.

Besonders erfolgreich sind Unternehmen, die KI-Technologien strategisch in bestehende Prozesse integrieren. So verzeichnen vollständig integrierte Marketing- und Vertriebsteams mit vierfach höherer Wahrscheinlichkeit einen deutlichen Anstieg ihres Return on Investment. Diese Unternehmen nutzen KI unter anderem für

Zielgruppensegmentierung (51 Prozent), Personalisierung (54 Prozent) und Kampagnenoptimierung (48 Prozent).

Unternehmen in Großbritannien führen

Allerdings variiert der Reifegrad zwischen den Regionen erheblich. Während in Großbritannien und Irland bereits 36 Prozent der Unternehmen KI-Lösungen optimieren oder strategisch einsetzen, befinden sich in der DACH-Region noch 64 Prozent in der Erforschungs- und Testphase. Als größte Herausforderungen bei der KI-Einführung nennen die Befragten Budgetbeschränkungen (43 Prozent), die Auswahl geeigneter Tools (39 Prozent) und fehlendes Fachwissen (36 Prozent).

Die „2024 State of Marketing AI“-Studie des Marketing AI Institute und Salesloft hat ebenfalls eine stark steigende Nutzung der KI im Marketing ergeben. Besonders bemerkenswert ist der Anstieg derjenigen, die KI in ihre täglichen Arbeitsabläufe integriert haben. Dieser Wert stieg von 29 Prozent im Jahr 2023 auf 36 Prozent im Jahr 2024. Zudem gaben 15 Prozent der Befragten an, dass sie ohne KI nicht mehr arbeiten könnten – ein deutlicher Anstieg gegenüber den 6 Prozent im Vorjahr.

15 Prozent können ohne KI nicht mehr arbeiten

Ein Hauptmotiv für den Einsatz von KI ist die Automatisierung wiederkehrender Aufgaben, um Zeit zu sparen. 80 Prozent der Befragten nannten dies als ihr primäres Ziel. Dies zeigt, dass KI vor allem zur Steigerung der Effizienz genutzt wird. Ein weiteres zentrales Ziel besteht darin, aus Marketingdaten verwertbare und umsetzbare Erkenntnisse zu gewinnen. 64 Prozent der Befragten betonten diesen Punkt als einen der Hauptvorteile von KI.

Trotz dieser Fortschritte bleiben jedoch Hindernisse bei der Einführung von KI bestehen. Die häufigste Barriere ist nach wie vor das Fehlen von Schulungen und Ausbildung. 67 Prozent der Befragten gaben an, dass dies das größte Hindernis für eine breitere Nutzung von KI in ihren Unternehmen darstellt. Dies zeigt, dass trotz des zunehmenden Interesses und der Verbreitung von KI-Technologien der Zugang zu Schulungen und der Aufbau von Wissen in diesem Bereich hinterherhinken.

Die beliebtesten KI-Anwendungen: Viel Bewegung im Markt

Die Top 50 der generativen KI auf dem Desktop

Auswahl durch Wagniskapitalfirma Andreessen Horowitz, sortiert nach monatlichen Visits, ermittelt durch SimilarWeb

1	ChatGPT	18	Midjourney	35	Vocal Remover
2	character.ai	19	yodayo	36	PicWish
3	perplexity	20	cutout.pro	37	Chub.ai
4	Claude	21	Viggle	38	Hix.ai
5	Suno	22	Photoroom	39	Vidnoz
6	JenitorAI	23	Gamma	40	Pixelcut
7	QuillBot	24	Veed.io	41	MaxAI.me
8	Poe	25	Pixlr	42	Blackbox AI
9	liner	26	ideogram	43	ChatPDF
10	Civit AI	27	you.com	44	Gauth
11	SpicyChat	28	DeepAI	45	coze
12	IIElevenLabs	29	SeaArt AI	46	Playground
13	Hugging Face	30	invideo AI	47	Doubau
14	Luma AI	31	Pixai	48	Speechify
15	candy.ai	32	Clipchamp	49	Nightcafe
16	Crushon AI	33	udio	50	AI Novelist
17	Leonardo.Ai	34	Chatbot App		

Tabelle: saco. / Quelle: a16z.com 2024

Der Markt für KI-Anwendungen ist weiter stark im Wandel. 30 Prozent Neulinge verzeichnet eine Liste der 100 meistgenutzten KI-Apps im Web und auf Smartphones.

Die Daten stammen von SimilarWeb und basieren auf monatlichen Visits sowie monatlich aktiven Nutzern. Die Liste wurde von der Wagniskapitalfirma Andreessen Horowitz zusammengestellt. Betrachtet wurden nur ausgewählte KI-Anwendungen, die ins Portfolio der Kapitalgeber passen. So lässt sich erklären, dass Googles KI-Dienst Gemini im vorangegangenen Report von März 2024 noch auf Platz 2 rangierte, nunmehr aber aus der Liste verschwunden ist. Vorne liegt im Segment der Webprodukte aus Sicht von Andreessen Horowitz weiterhin ChatGPT, gefolgt von character.ai – einer Website zum Chatten mit simulierten Figuren. Im Schutz der Anonymität können sich Nutzer hier mit einer KI-gesteuerten Psychologin unterhalten, Vorstellungsgespräche bei einem KI-Personalchef üben oder mit virtuellen „Freunden“ intime Geheimnisse erörtern.

Konkurrenz rückt ChatGPT näher

Drittplatzierte KI-Dienst ist die weiter aufstrebende Antwortmaschine Perplexity. Es folgt Claude, die unmittelbare ChatGPT-Konkurrenz.

Claude und Perplexity hatten zuletzt die höchsten monatlichen Steigerungsraten. Bei Perplexity beeindruckt viele Nutzer, wie die KI eine Anfrage in zumeist sinnvolle Einzelteile zerlegt, Teilantworten findet und zu den Quellen verlinkt. Diese schrittweise Annäherung in Gedankenschritten hat Marktführer ChatGPT erst vor einer Woche in seinem neuen Modell o1 eingeführt.

Weiterhin unterlässt es ChatGPT häufig, Quellen zu benennen und zu verlinken. Konkurrent Claude übertrifft nach Meinung einiger Forscher ChatGPT in standardisierten Tests zu Argumentation, Mathematik und Programmierung. Zudem ist Claude oft sprachlich ausgefeilter und präziser. Die Konkurrenz rückt also ChatGPT zusehends auf die Pelle.

Unter den Top 50 sind zunehmend andere KI-Systeme als Chatbots: Videoanwendungen und Soundgeneratoren nutzen Künstliche Intelligenz, um mehr oder minder kreative neue Medien zu erfinden. In Luma AI beispielsweise formuliert man einen Prompt, aus dem die Maschine ein Video schnitzt. Viggie AI erlaubt es, beliebige Fotos mit Tanzsequenzen zu verknüpfen. Da betrat etwa der Rapper Lil Yachty die Bühne, tanzte sich bis zur Vorbühne und eröffnete ein wuchtiges Konzert. Mit Viggie haben wir dann mal den Autor dieses Textes ins Video montiert. Eine Zeitlang

entstanden aus dieser Szene immer wieder neue Memes – Internetphänomene, in denen dann der Joker, Elon Musk oder Donald Trump in die Tanzszene kopiert wurden.

KI generiert Musik

Höchster Neuzugang in diesen Charts ist die Musikanwendung Suno. Damit können beliebige Songs aus der Maschine generiert werden, mitsamt Gesang, anhand der Stilrichtung. Manche vorgeschlagenen Stile dürfte man noch nie gehört haben, nehmen wir nur mal den Bubblegum Dance oder den Crunk Calypso. Das Zeug zum Hit haben die Erzeugnisse eher nicht, irgendwie mischt die Maschine bekannte Muster neu ab. Und eine einst wichtige Funktion, prominente Sänger oder Bands zu imitieren, hat Suno vermutlich aus rechtlichen Gründen unterbunden. Wer aber auf der Suche nach Fahrstuhlmusik oder dem Hintergrundsound fürs Restaurant ist, dürfte hier fündig werden.

Auf Mobilgeräten führt ebenfalls ChatGPT die Liste an, gefolgt vom Browser Microsoft Edge und der Matheanwendung (!) Photomath. Die App löst Matheaufgaben – einfach eine Gleichung abfotografieren, die KI löst sie auf. Bei Bedarf erklärt sie Schritt für Schritt, wie der Weg zum Ziel führt, wenn auch auf Englisch. Auf den weiteren Plätzen sind es vor allem Fotoanwendungen, die häufig genutzt werden. Facemoji macht aus Gesichtern verspielte Emojis, Remove It entfernt Personen oder Gegenstände aus Bildern. Meitu hilft beim Nachbearbeiten von Fotos, Snow erledigt das Gleiche speziell bei Porträts. Hoch eingestiegen in dieser Liste ist eine KI-Assistentin namens Luzia: eine künstliche Figur, die dem Handybesitzer bei Alltagsfragen zur Seite steht.

Chinesen nehmen westliche Märkte ins Visier

Wachsende Bedeutung erlangen auf dem Smartphone Anwendungen der Tiktok-Mutter ByteDance. Die Chinesen haben eine eigene Abteilung für neuartige KI-Apps gegründet und nehmen damit nicht nur den chinesischen Markt ins Visier, sondern auch den amerikanischen. Dazu gehören als Apps neben einem allgemeinen Assistenten namens Doubao eine Lernplattform und ein Chatbot-Ersteller. Andreessen Horowitz hat zudem ein neues Genre ausgemacht: Schönheit und Dating. Da beurteilt die KI eigene aufgenommene Fotos auf ihre Ästhetik und gibt Tipps für bessere Aufnahmen. Eine andere App namens Rizz hilft beim Chatten in der Dating-App: Sie macht Vorschläge für sinnvolle Antworten, ein Screenshot der bisherigen Konversation genügt.



Marcus Schwarze

Freier Journalist und
Berater Digitales in
Koblenz

Must Read

Was sonst noch wichtig ist

- Blackrock und Microsoft planen einen **30-Milliarden-Dollar-Fonds** für KI-Infrastruktur. Ziel ist der Bau von Rechenzentren und Energieprojekten, um den steigenden Bedarf an Rechenleistung für KI zu decken. MGX aus Abu Dhabi ist Gesellschafter, Nvidia bietet Expertise. Mit Fremdfinanzierung sollen bis zu 100 Milliarden Dollar mobilisiert werden.
[faz.net/pro243813](https://www.faz.net/pro243813)
- KI-Systeme, die von **Personalabteilungen** für Einstellungen genutzt werden, benachteiligen Menschen mit Behinderungen, Einwanderer und – trotz „geschlechtsblinder“ Trainings – sogar Frauen, hat eine Studie ergeben. Zudem werden den Bewerbern bevorzugt Positionen mit geringer Entlohnung angezeigt.
[faz.net/pro243814](https://www.faz.net/pro243814)
- **Medien- und Rundfunkunternehmen** sehen KI positiv, implementieren die Technik aber nur vorsichtig. 85 Prozent glauben, AI wird die Branche signifikant beeinflussen. KI wird hauptsächlich für Routineaufgaben wie Transkription und Clip-Erstellung genutzt und soll Mitarbeiter von monotonen Aufgaben entlasten.
[faz.net/pro243815](https://www.faz.net/pro243815)
- Die **Medienbranche** erlebt durch KI einen Wandel, der sich auf die Anforderungen an Fachkräfte auswirkt: 72 Prozent der Medienführungskräfte sehen die Notwendigkeit, Talente und Fähigkeiten in ihren Organisationen neu zu gestalten. 66 Prozent wollen Talente außerhalb der Medienbranche rekrutieren, zeigt eine Kantar-Studie.
[faz.net/pro243816](https://www.faz.net/pro243816)
- Bytedance, das Mutterunternehmen von **Tiktok**, verstärkt seine Bemühungen in der Chipentwicklung. Das Unternehmen hat dafür ein Team von etwa 100 Ingenieuren aufgebaut. Bytedance will mehrere Hunderttausend Chips bestellen, die von TSMC gefertigt werden.
[faz.net/pro243817](https://www.faz.net/pro243817)

Alles Wichtige zu KI, Transformation und Plattformen.

F.A.Z. PRO Digitalwirtschaft

Ihr Experten-Update rund um digitale Wirtschaft als wöchentliches Briefing und als Website. Inklusive freiem Zugang zu allen Artikeln auf unserem Onlineportal FAZ.NET.



**Jetzt 2 Monate
kostenfrei testen!**

F.A.Z. PRO Digitalwirtschaft vereint relevante Beiträge rund um Digitalisierung, digitale Ökonomie und Technologie in einem neuen Informationsangebot. Sichern Sie sich Ihren Wissensvorsprung:

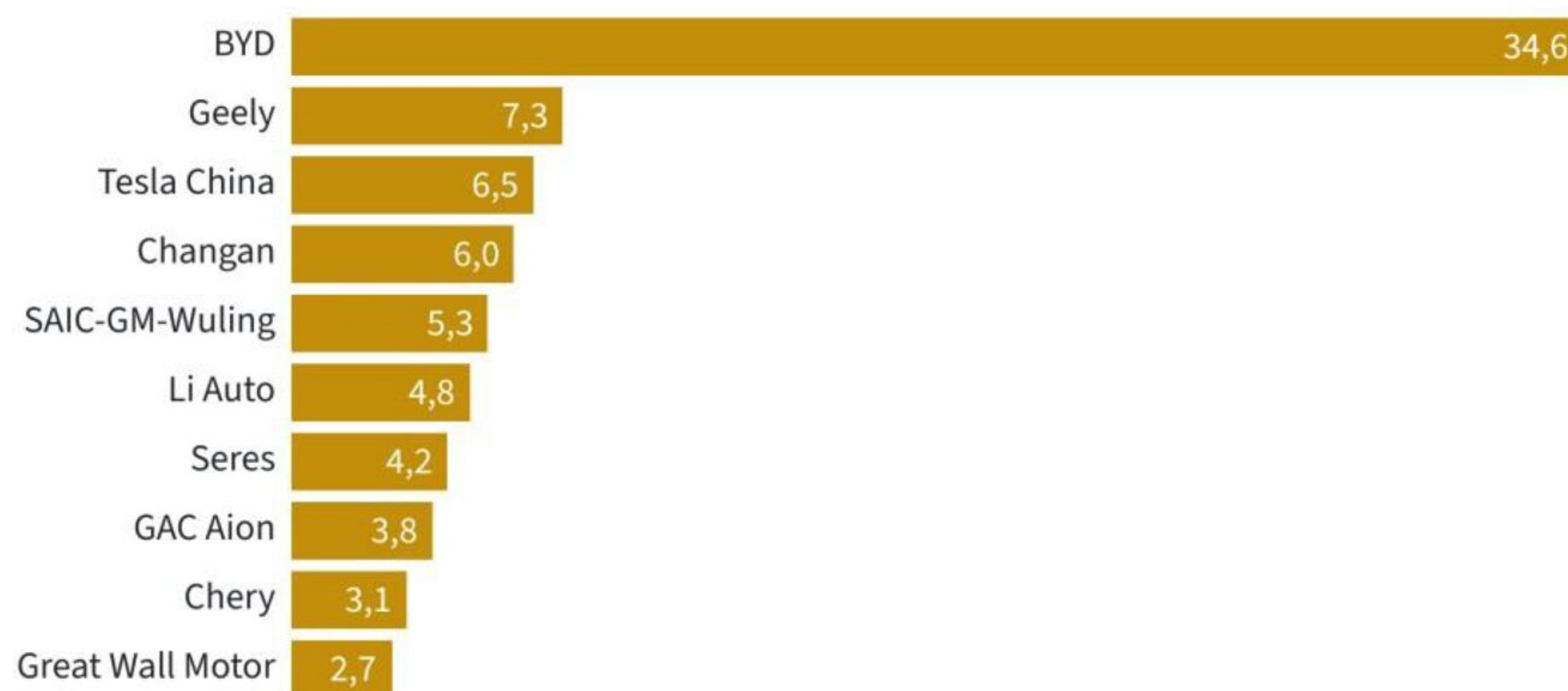
**Jetzt F.A.Z. PRO Digitalwirtschaft und freien Zugang zu FAZ.NET
2 Monate kostenfrei testen unter faz.net/pro-digitalwirtschaft**



BYD baut Dominanz auf dem chinesischen Elektroautomarkt aus

Top-10 Elektroautohersteller in China

Reine Elektroautos und Plug-in-Hybride, Anteile in Prozent zwischen Januar und August 2024



Grafik: sacco. / Quelle: CPCA 2024

Während die europäischen Hersteller in die Krise geraten, fährt BYD in China von einem Rekord zum nächsten. Im August verkaufte das Unternehmen 373.000 Elektroautos, das sind 36 Prozent mehr als im Vorjahr.

Chinas konsequenter Ausbau der Elektromobilität zeigt Wirkung: Im August wurden erstmals mehr als eine Million „New Energy Vehicles“ verkauft, die sich aus reinen Elektroautos und Plug-in-Hybriden zusammensetzen. Das bedeutet einen Anstieg um 43 Prozent im Vergleich zum Vorjahr. Gleichzeitig wird die Dominanz des Marktführers BYD immer größer: Der Anteil an diesen Elektroautos kletterte auf 37 Prozent; der Anteil von BYD am Gesamtmarkt (inklusive Verbrenner) verdoppelte sich sogar auf 20 Prozent.

BYD hat inzwischen nicht nur die deutsche Konkurrenz deklassiert, sondern hängt auch den Rest der Branche ab. Geely erreicht als Zweitplatzierte lediglich einen Marktanteil von 7,2 Prozent, Tesla folgt mit 6,2 Prozent auf dem dritten Rang. Deutsche Hersteller sind auch im August nicht unter den Top 10 vertreten. Nach Berechnungen der Beratungsgesellschaft von PWC liegt der gemeinsame Anteil von VW, BMW und Mercedes am chinesischen Elektroautomarkt bei nur rund 6 Prozent.

900.000 Beschäftigte bei BYD

Bei BYD stehen alle Zeichen weiter auf Wachstum. Die Zahl der abgesetzten Fahrzeuge könnte in diesem Jahr auf vier Millionen steigen. Das Unternehmen hat dafür seine Belegschaft deutlich aufgestockt und beschäftigt inzwischen mehr als 900.000 Mitarbeiter, was BYD zu einem der größten Arbeitgeber Chinas macht. Die Einstellungsoffensive passt zu den Bestrebungen

der chinesischen Regierung, mehr Arbeitsplätze zu schaffen, insbesondere für Hochschulabsolventen.

Mit fast 110.000 Beschäftigten in den Bereichen Technologie und Forschung & Entwicklung bezeichnet sich BYD als weltweit größter Autohersteller im Hinblick auf das F&E-Personal. In den vergangenen zwei Jahren wurden fast 50.000 Hochschulabsolventen eingestellt. Während BYD und einige andere einheimische Unternehmen Personal aufstocken, bauen ausländische Automarken in China aufgrund sinkender Verkaufszahlen Stellen ab.

Erfolgsfaktor ist vertikale Integration

Laut dem jüngsten Geschäftsbericht von BYD machten die Personalkosten 17,5 Prozent des Umsatzes im ersten Halbjahr aus, was durchschnittlichen monatlichen Kosten von 11.700 Yuan je Mitarbeiter entspricht. Trotz eines aggressiven Preiskriegs, bei dem hohe Rabatte auf die meistverkauften Modelle gewährt wurden, konnte BYD im zweiten Quartal einen verbesserten Nettogewinn verzeichnen.

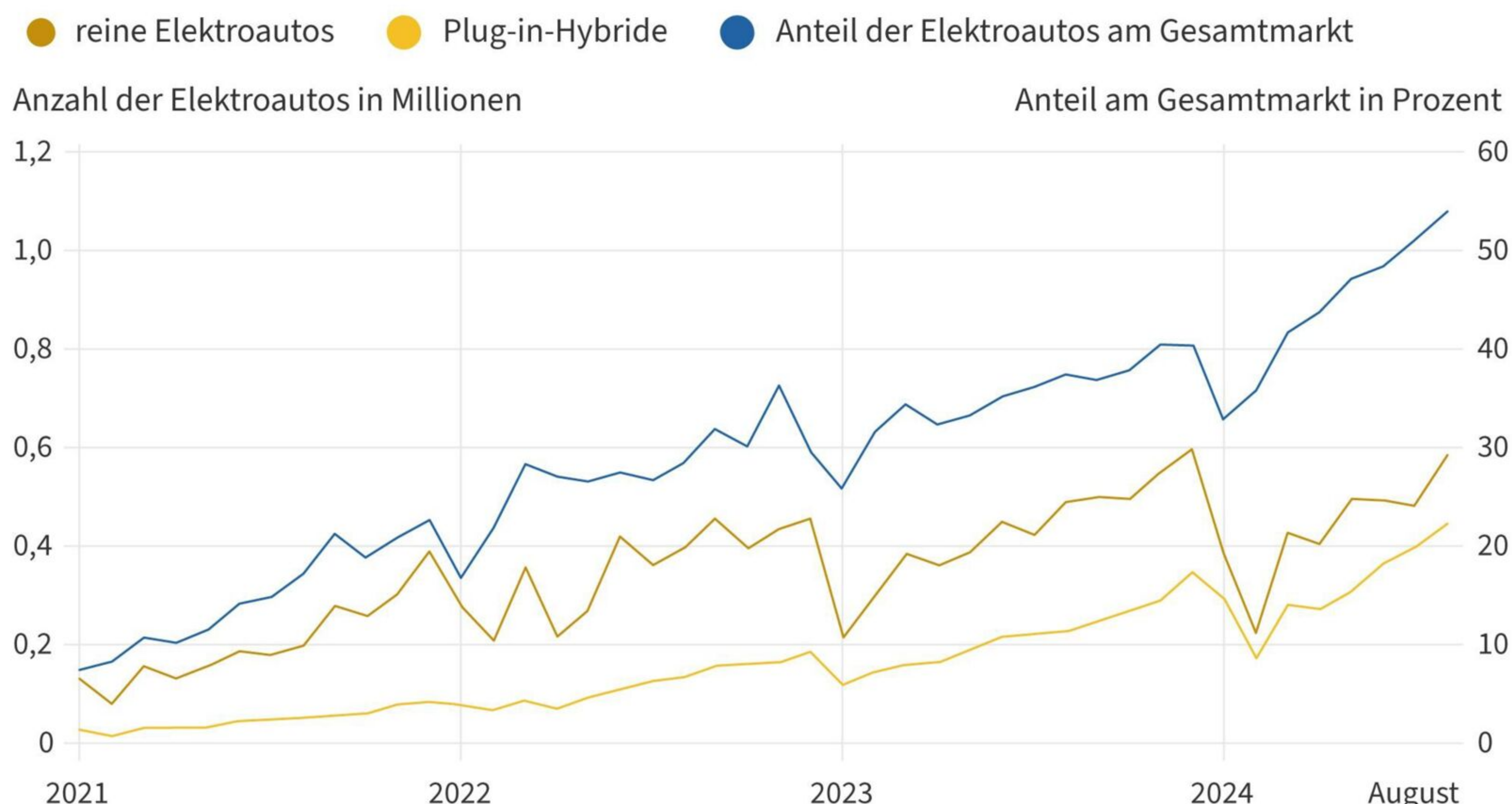
BYDs Erfolg basiert auf einer vertikalen Integration, die von der Batterieproduktion bis zur Fahrzeugherstellung reicht, was dem Unternehmen signifikante Kostenvorteile verschafft. Denn BYD ist nach CATL inzwischen auch der zweitgrößte Batteriehersteller der Welt. In diesem Geschäftsfeld liegen auch die Wurzeln des Unternehmens. Die Autos wurden erst später um die Batterien als zentrale Komponente eines Elektroautos herumgebaut.



Holger Schmidt

Redaktionsleiter
„Newsletter und
Verticals“

Elektroauto-Verkäufe in China



Grafik: sacco. / Quelle: CPCA 2024

BYD steht im harten Wettbewerb mit rund 130 Marken auf dem chinesischen Markt. Diese intensive Wettbewerbssituation treibt Innovationen voran und drückt die Preise, was die Attraktivität von Elektrofahrzeugen für chinesische Verbraucher weiter steigert. Zudem hat die chinesische Regierung durch eine gezielte Förderpolitik maßgeblich zum Aufstieg der Elektromobilität beigetragen. Dazu gehören Anreize wie die weitgehende Befreiung der meisten Elektromodelle von der Mehrwertsteuer. Zusätzlich setzt die Regierung auf restriktive Zulassungsregeln für Verbrenner in Großstädten, während die Bedingungen für Elektroautos verbessert wurden.

30 Prozent Kostenvorteil

Die dominante Stellung chinesischer Marken stellt europäische Automobilhersteller vor enorme Herausforderungen. Sie kämpfen mit mehreren strukturellen Nachteilen. Chinesische Hersteller haben Kostenvorteile von bis zu 30 Prozent gegenüber westlichen Konkurrenten. Die hohe Wettbewerbsintensität in China führt zu schnelleren Innovationszyklen, bei denen europäische Unternehmen oft nicht mithalten können. Zudem bevorzugen chinesische Verbraucher zunehmend einheimische Marken, die als innovativer und preiswerter wahrgenommen werden.

Der Erfolg chinesischer Elektroautohersteller beschränkt sich jedoch nicht mehr nur auf den Heimatmarkt. Mit wachsenden Überkapazitäten im eigenen Land richten chinesische Unternehmen ihren Blick verstärkt auf internationale Märkte, insbesondere Europa.

2023 exportierte China erstmals mehr Autos als Japan und wurde damit zur weltweit größten Autonation. Von den vier Millionen exportierten Fahrzeugen gingen aber nur 141.000 Einheiten nach Europa. Die Einfuhrzölle auf Elektroautos, die nach den USA nun auch die EU erhebt, haben den Marktanteil der chinesischen Hersteller zuletzt wieder sinken lassen.

Regierung will E-Auto-Technik im Land halten

Die chinesische Regierung drängt die Autohersteller dazu, fortschrittliche Elektrofahrzeugtechnologien im Land zu behalten, auch wenn sie weltweit Fabriken errichten. Peking ermutigt zur Ausfuhr sogenannter „Knock-down-Kits“, bei denen Schlüsselkomponenten in China produziert und zur Endmontage ins Ausland geschickt werden. Bei einem Treffen im Juli wurden Autohersteller angewiesen, keine autobezogenen Investitionen in Indien zu tätigen und Investitionen in der Türkei vorab zu melden. Diese Direktive könnte die Globalisierungsbemühungen chinesischer Unternehmen behindern, die neue Märkte erschließen wollen.

Europäische Länder, die chinesische Autobauer anlocken wollen, könnten enttäuscht werden. Die EU warnt, dass bloße Montagefabriken nicht ausreichen, um Zölle zu umgehen. In Brasilien wollen BYD und Great Wall Motor den Anteil lokaler Komponenten erhöhen, um Zollvorteile zu nutzen. Diese Entwicklungen zeigen Chinas Bemühungen, seine Elektroautoindustrie zu schützen und regulatorische Risiken zu mindern, während chinesische Autohersteller ihre globale Präsenz ausbauen.

Must Read

Was sonst noch wichtig ist

- **Volvo** stellt sein Auto-Abo-Programm in den USA und Europa ein. Bestehende Verträge werden erfüllt, aber Neukunden nicht mehr angenommen. Volvo wolle sich auf die Einführung neuer Produkte konzentrieren. Offenbar ist der Schritt gedacht, um die Kosten zu senken und den Umsatz kurzfristig zu erhöhen.
[faz.net/pro243824](https://www.faz.net/pro243824)
- Die Investitionen in **Mobilitätstechnologien** steigen wieder. Im zweiten Quartal 2024 dominierte autonomes Fahren, mit großen Investitionen von Nvidia und anderen. Die Branche erlebte einen Aufschwung mit 6,7 Milliarden Dollar an Investitionen, aber kleinere Start-ups hatten es schwerer. Besonders Waymo und Zoox expandieren.
[faz.net/pro243825](https://www.faz.net/pro243825)
- **Apples** neues Betriebssystem iOS 18 bringt einige Neuerungen für Carplay, zum Beispiel Apple Intelligence mit verbessertem Siri (ab iOS 18.1), Farbfilter für Sehbehinderte, Sprachsteuerung für freihändige Bedienung, Klangerkennung für Hörgeschädigte und einen automatischen Stumm-Modus.
[faz.net/pro243826](https://www.faz.net/pro243826)
- **Festkörperbatterien** versprechen eine Revolution für Elektrofahrzeuge. Sie bieten höhere Energiedichte und schnelleres Laden im Vergleich zu Lithium-Ionen-Batterien. Die Massenproduktion wird für die zweite Hälfte des Jahrzehnts erwartet, was Reichweite und Ladezeiten von E-Autos deutlich verbessern könnte.
[faz.net/pro243827](https://www.faz.net/pro243827)

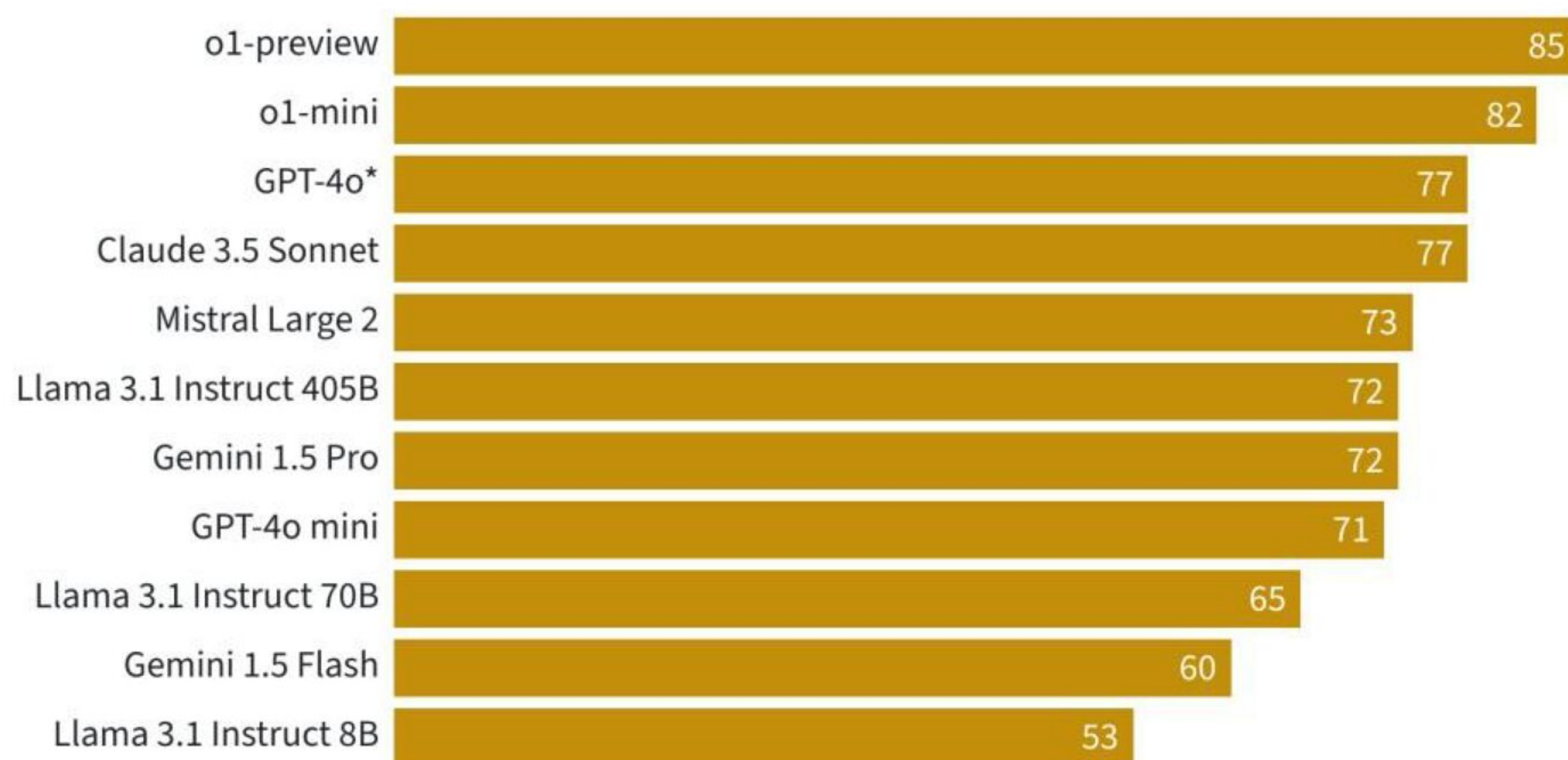
o1 verändert Kulturtechnik des Promptens: Es wird einfacher und präziser

Marcus Schwarze

Freier Journalist und Berater Digitales in Koblenz

KI-Systeme im Vergleich – Qualität

Qualitätsindex von Artificial Analysis, höher ist besser



* 6. August 2024

In Sachen Qualität der Antwort haben sich die KIs o1-Preview und o1-Mini vor GPT-4o und Claude geschoben

Grafik: sacco. / Quelle: artificialanalysis.ai

Die nächste Generation der generativen Künstlichen Intelligenz (KI) beflügelt die weltweite KI-Szene. Die von Open AI vorgestellte Version o1 übertrumpft in einzelnen Anwendungen die menschliche Intelligenz – etwa bei einem IQ-Test von Mensa, dem Hochbegabten-Netzwerk.

Onlinetests zum Prüfen des persönlichen Intelligenzquotienten (IQ) geben regelmäßig nur Anhaltspunkte über die menschliche Intelligenz, sie sind kein Ersatz für professionelle Intelligenztests von Psychologen. Mensa Norwegen, ein Netzwerk von Hochbegabten in dem skandinavischen Land, bietet den Test dennoch im Netz an, auch auf Deutsch **1**. Abgefragt werden kognitive Fähigkeiten, bei denen Menschen meist einen IQ-Wert zwischen 85 und 145 erhalten, im Schnitt 100 Punkte.

Schlauer als der Mensch

Der Reporter Maxim Lott hat der neuen KI o1 von Open AI die Fragen des Intelligenztests gestellt **2**. o1 erzielte 120 Punkte – 20 mehr als der Mensch im Durchschnitt, bis zu 30 mehr als bisherige KI-Versionen von Claude 3 und GPT 4. Hintergrund ist ein sogenanntes „Reasoning“-Modell, bei dem sich die Maschine mehr Zeit für eine gestellte Aufgabe nimmt. Sie bildet Gedankenketten **3** und ist so zwar langsamer als bisherige Versionen, löst die Probleme aber mehrstufig. Noch ist o1 nicht in der Lage, Bilder entgegenzunehmen, um darin Muster zu erkennen. Lott behalf sich durch textliche Beschreibungen der Rätsel – was strenggenommen den Test verfälschte, aber dennoch die Richtung aufzeigt: KI wird schlauer.

Das zeigt sich besonders bei Fragen der Mathematik und Physik. So lautet ein kniffliges Rätsel in Tests, in Zahlenfolgen ein Muster zu erkennen – beispielsweise folgende Aufgabe:

Vorgegeben ist die folgende Folge:

1; 5; 12; 22; 35; ...

a) Für Normalsterbliche:

Nennen Sie die drei nächsten Folgeglieder.

b) Für Ambitionierte:

Wie lautet das 24. Glied dieser Folge ($a_1 = 1$)?

c) Für Schüler der Oberstufe:

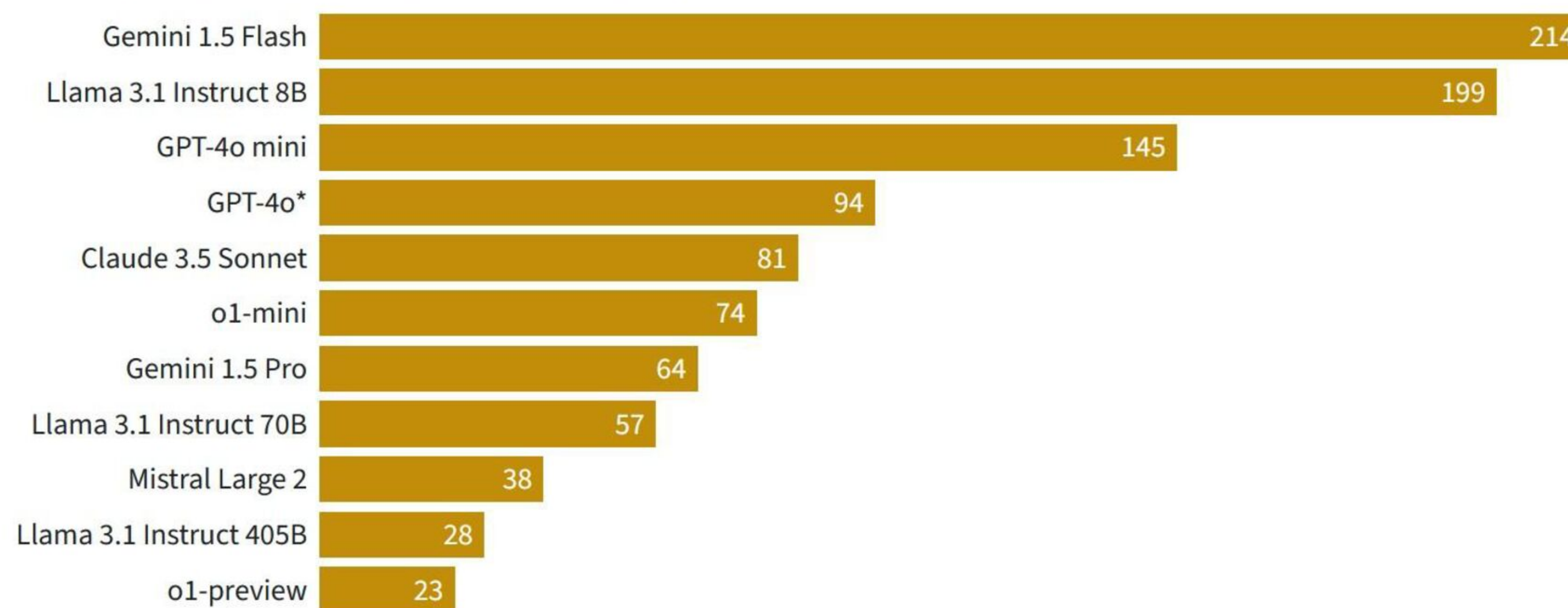
Geben Sie eine rekursive und explizite Darstellung an und rechnen Sie ein Beispiel als Probe.

Bessere Antwort durch Schritt-für-Schritt-Vorgehensweise

Die o1-Maschine kommt auf das richtige Ergebnis **4** – und ist zudem in der Lage, die mathematischen Formen als rekursive und explizite Darstellung anzugeben. Sie braucht dann durchaus eine halbe Minute bis zur Antwort, zerlegt die Anfrage nach und nach in Einzelschritte. Durch diese Schritt-für-Schritt-Erklärungen werden die Berechnungen vertrauenswürdiger und leichter nachvollziehbar – eine Vorgehensweise, die man in früheren Versionen von ChatGPT explizit im Prompt einbauen musste. Nun nimmt einem die KI-Assistenz diese Zerlege-Aufgabe ab. Und dürfte so das Prompten vereinfachen.

KI-Systeme im Vergleich – Geschwindigkeit

Anzahl ausgegebener Tokens pro Sekunde, mehr ist besser



* 6. August 2024

Die KI nimmt sich Zeit: Das Modell o1-preview landet bei der Geschwindigkeit der Antwort am Schluss der Rangliste

Grafik: sacco. / Quelle: artificialanalysis.ai

1 Beitrag:
faz.net/pro243828

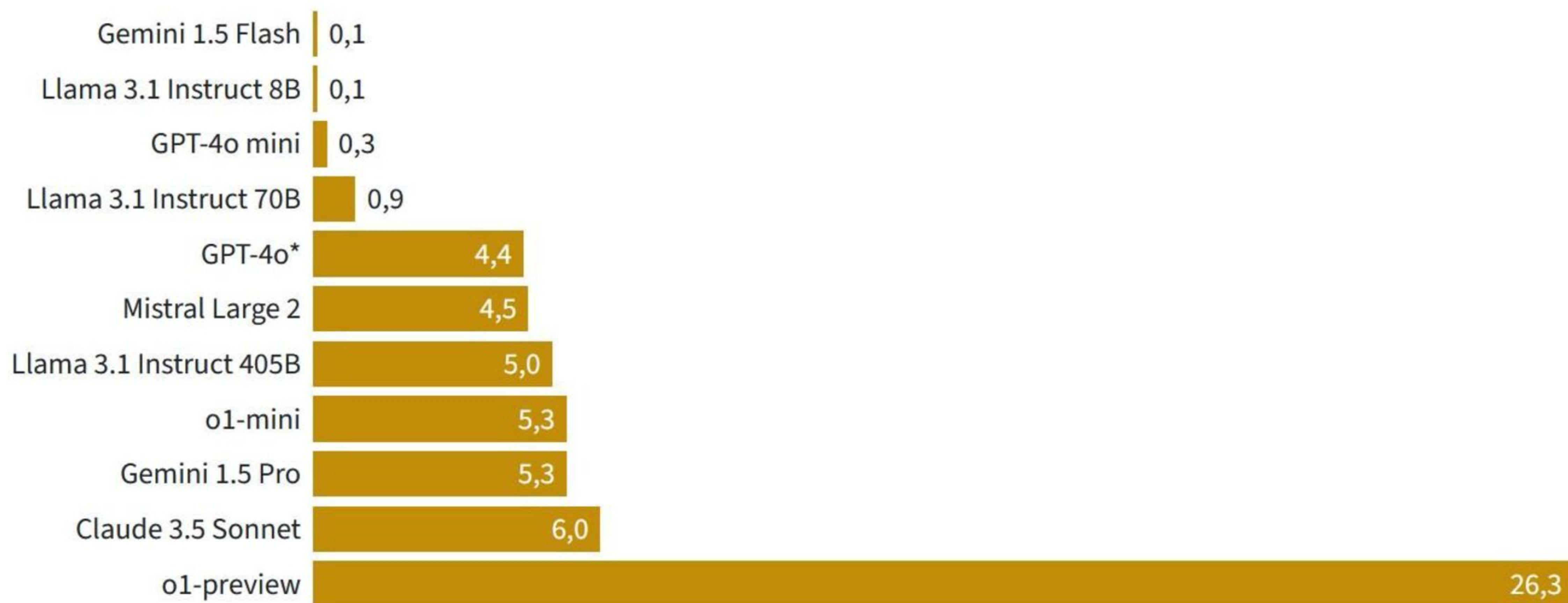
2 Beitrag:
faz.net/pro243829

3 Beitrag:
faz.net/pro243830

4 Beitrag:
faz.net/pro243831

KI-Systeme im Vergleich – Preis

US-Dollar per eine Million Tokens, kleiner ist besser



* 6. August 2024

Das Modell o1-preview ist von allen verglichenen Modellen das mit Abstand teuerste

Grafik: sacco. / Quelle: artificialanalysis.ai

In der KI-Szene machten in den vergangenen Tagen zahlreiche Beispiele zur Leistungsfähigkeit von o1 die Runde. Ein Programmierer ließ in sechs Prompts eine Aufgabe der Astrophysik auswerten, ⁵ die KI schrieb ihm den Code binnen einer Stunde – für eine Aufgabe, die nach seinem Bekunden sonst bis zu einem Jahr gedauert hätte. Ammaar Reshi kombinierte o1 mit der Programmierumgebung Cursor und kreierte nach seinen Worten ⁶ innerhalb von zehn Minuten eine Wetter-App. Yunyu Lin gab der KI die Rätselaufgabe ⁷, vier vorgegebene Zahlen mathematisch zu verknüpfen, damit am Ende die Zahl 24 herauskommt. o1 löste die Aufgabe, GPT-4o scheiterte daran.

Best Practices fürs neue Prompten

Die Kulturtechnik des Promptens verändert sich mit der neuen Maschine. Hersteller Open AI empfiehlt für den Umgang mit o1 möglichst einfache Aufforderungen: „Einige Prompt-Engineering-Techniken, wie die Aufforderung, ‚Schritt für Schritt‘ zu denken, verbessern die Leistung möglicherweise nicht, sondern können sie manchmal sogar behindern.“ Als „Best Practices“ empfiehlt ⁸ Open AI:

- Halten Sie den Prompt einfach und direkt: o1 zeichnet sich durch das Verständnis und die Beantwortung kurzer, klarer Anweisungen aus, ohne dass umfangreiche Anleitungen erforderlich sind.
- Vermeiden Sie Aufforderungen zur Denkkette: Da diese Modelle intern argumentieren, ist es unnötig, sie dazu zu bringen, „Schritt für Schritt zu denken“ oder „ihre Vorgehensweise zu erklären“.
- Verwenden Sie Trennzeichen zur Übersichtlichkeit: Verwenden Sie Trennzeichen wie dreifache Anführungszeichen, XML-Tags oder Abschnittstitel, um eindeutige Teile der Eingabe anzuzeigen, und helfen Sie dem Modell so, verschiedene Abschnitte zu verstehen.
- Begrenzen Sie den zusätzlichen Kontext: Wenn Sie zusätzlichen Kontext oder Dokumente bereitstellen, fügen Sie nur die relevantesten Informationen hinzu, um zu verhindern, dass das Modell seine Antwort zu kompliziert macht.

⁵ Beitrag: faz.net/pro243832

⁶ Beitrag: faz.net/pro243833

⁷ Beitrag: faz.net/pro243834

⁸ Beitrag: faz.net/pro243835

Alles Wichtige aus der Welt des Rechts.

Jetzt neu: F.A.Z. PRO Einspruch

Ihr Experten-Update rund um Recht und Gesetz als wöchentliches Briefing und als Website. Inklusive freiem Zugang zu allen Artikeln auf unserem Onlineportal FAZ.NET.



**Jetzt 2 Monate
kostenfrei testen!**

F.A.Z. PRO Einspruch vereint alle juristisch relevanten Beiträge aus der F.A.Z., ergänzt um exklusive Analysen und Einordnungen in einem neuen Informationsangebot. Sichern Sie sich Ihren Wissensvorsprung:

**Jetzt F.A.Z. PRO Einspruch und freien Zugang zu FAZ.NET
2 Monate kostenfrei testen unter faz.net/pro-einspruch**



Videogeneratoren arbeiten immer präziser



Deutschland in Form eines
erleuchteten Puddings
(Marcus Schwarze/Dall-E3,
KI-generiert)

Marcus Schwarze

Freier Journalist und Berater Digitales in Koblenz

Video-KIs erleben einen neuen Boom. Aus den anfänglichen Spielereien werden immer präzisere Umsetzungen. Die Unternehmen der Künstlichen Intelligenz im Genre Video entwickeln eine neue Generation von Text-zu-Video-Erzeugern.

Die „Tagesschau“ hat es vorgemacht: „Das ist ja echt cringe“, sagt Sprecherin Susanne Daubner am Ende eines Videos, in dem sie ankündigt, beim nächsten „Jugendwort des Jahres“ nicht mehr dabei zu sein. „Was machen wir denn jetzt?“, fragt eine Redakteurin aus dem Off, und schon lädt eine Künstliche Intelligenz ein künstliches Video nach. Darin spricht plötzlich eine KI-generierte Figur mit dem Aussehen von Daubner und ihrer Stimme die Nachrichten. Ist das „cringe“?

Lippensynchron, mit Gestik und Mimik

Was die „Tagesschau“-Redaktion als Gag auf Instagram und anderen Kanälen veröffentlichte, ist bei manchen Sendern in Asien und den USA bereits Realität. Und das Format kann jedermann für kleines Geld kopieren. Während die „Tagesschau“ Daubner erkennbar als Sprechpuppe gestaltete – ein wenig wie aus einem Comic –, stellen KI-Dienste im Netz real aussehende Menschen in den Mittelpunkt.

Zum Beispiel bei Vidnoz ¹. Sie heißen Olivia und Mila, Ali und Paul – und lesen dann die eingegebenen Texte. Lippensynchronität und eine sparsame Gestik speist die KI in die Figuren, etwas Mimik gehört dazu. Wer die Augen ein bisschen zusammenkneift und nicht so genau hinhört, könnte glauben, es seien echte Menschen.

Die Tücke liegt im Detail. Nicht jedes Wort kennt die KI, zum Beispiel „cringe“. Das spricht die Maschine wie „Inge“ aus, nun müsste man ihr mit einer Umformulierung in „krinsch“ auf die Sprünge helfen. Auch wirken manche Wortübergänge unnatürlich. Wer sich traut, klont bei Vidnoz AI seine eigene Stimme. Fortan ist die KI in der Lage, jeden beliebigen getippten Text zu sprechen, ohne „Äh“ und „Hmm“. Das funktioniert nicht nur mit deutschen Texten, sondern auch mit übersetzten. Der persönliche Avatar spricht dann nicht nur die erlernte Stimme mit dem persönlichen Duktus, sondern auch spanisch, chinesisch oder norwegisch, je nach eingetipptem Text. Die persönliche Stimme bleibt erhalten.

Das gelingt auch mit realen Videos: Wer als internationale Unternehmenschefin die Belegschaft in Frankreich, Italien und Spanien in der Landessprache ansprechen möchte, kann bei Vidnoz sein deutsches Video entsprechend übersetzen lassen. Die KI wandelt die Videos lippensynchron um und, wenn man das will, mit einer synthetischen Nachbildung der Originalstimme. Vidnoz gibt es ab 12,48 Euro im

Monat, für den ersten Test kann man drei Minuten am Tag kostenlos vertonen. Bei der Konkurrenz von HeyGen ² gibt es ähnliche Funktionen ab 24 Dollar pro Monat.

Text-zu-Video-Maschinen erzeugen Clips

Sprechende Avatare sind nur ein Genre der Video-KIs. Immer beliebter werden Text-zu-Video-Maschinen. Die Luma Dream Machine ³ erstellt aus einem kurzen Prompt ein Fünf-Sekunden-Video. Ein Prompt lautet beispielsweise: „A teddy bear in sunglasses playing electric guitar, dancing and headbanging in the jungle in front of a large beautiful waterfall“. Das generiert einen Teddybär, der vor einem Wasserfall Gitarre spielt. Einen Tiger im Schnee erzeugt der Prompt „Low-angle shot of a majestic tiger prowling through a snowy landscape, leaving paw prints on the white blanket“. Danach kann man das Video immer wieder um fünf Sekunden erweitern. Auch Änderungen der Kameraposition sind möglich. Nicht immer gelingen die Videos perfekt, im Beispiel wächst unserem Tiger plötzlich ein zweiter Schwanz.

Doch ergeben sich für Kreative manche faszinierende neue Möglichkeiten. Nach einer Idee des auf X aktiven Künstlers Umesh haben wir einmal Flüssigkeiten in Form von Ländern Europas ineinander zerfließen lassen. Durch Aneinanderreihen entstand so ein Video wie geschaffen für den nächsten Eurovision Song Contest. Als Prompt für ein zunächst statisches Bild verwendeten wir folgende Anweisung:

„A top view of a photorealistic red liquid spill on a flat shiny surface, the puddle forming the exact shape of Germany, with light reflecting through it to enhance its translucent and glossy appearance. format 16:9“

Luma transformiert zwischen zwei Einzelbildern

Das Ergebnis entstand bei der KI Dall-E3. Es erinnert an einen Pudding auf einer matten Fläche, stark beleuchtet, im Umriss des gewünschten Landes. In der Videoanwendung Luma AI verwendeten wir nun den Deutschland-Pudding als Eingangsbild und einen Umriss Europas als Schlussbild. „Slow seamless transformation“ lautete der Prompt für Luma, um die beiden Flüssigkeiten langsam ineinander zu transformieren. Mit einem normalen

¹ **Vidnoz:**
faz.net/pro243836

² **HeyGen:**
faz.net/pro243837

³ **Lumalabs:**
faz.net/pro243838

Schnittprogramm reihten wir am Ende mehrere solcher Sequenzen aneinander. Es fehlt nur noch die Europahymne, wir haben mal die Musikmaschine Sumo darauf angesetzt. „Freude schöner Götterfunken“ hat die Maschine nach unserem Geschmack allerdings zu stark abgewandelt. An das Original reicht die KI-generierte Musik noch nicht heran.

Welche Möglichkeiten sich bei den weiter entwickelten Video-KIs ergeben, zeigt der beliebte Dienst Runway **4** mit der dritten Generation seiner Videomaschine, „Gen-3 Alpha“. Auch hier können aus Textprompts Kurzvideos hergestellt werden, jeweils von fünf oder zehn Sekunden Länge. Anderthalb Minuten später ist das Video erstellt. Im Prompt wählt man das gewünschte Motiv, die darzustellende Szene, die Beleuchtung, Einzelheiten zur Kamera und ihren Einstellungen, Bewegungen und mehr. Je ausführlicher der Prompt, desto präziser stellt die KI das Video her. Dafür lassen sich auch Einzelbilder als Ausgangsmaterial verwenden. Ebenso kann man bereits erstellte Videos als Quelle wählen und dann per Prompt aus einer dunklen Dschungellandschaft eine schön beleuchtete machen oder aus einer persönlichen Aufnahme einen Comic.

Videogenerierung wird künftig Teil des Adobe-Pakets

„Video to Video“ heißt die Funktion. Manche auf X sichtbare **5** Beispiele wirken etwas cringe **6**, würde Daubner wohl sagen. Doch taugt das Werkzeug vermutlich auch als professionelles, wie ein Blick auf einen Porsche **7** zeigt, der mal in Weiß aufgemotzt, mit Bundeswehrtarnung und letztlich im Schneegestöber platziert wird. Das ruft auch die Gestaltungsprofis von Adobe auf den Plan. Firefly, das Adobe-Werkzeug zum Generieren von Bildern, wird künftig auch Videos beherrschen. Adobe legt dabei großen Wert auf die Achtung von Urheber- und Verwertungsrechten. Damit bekommen etwa auch Werbeagenturen eine gewisse Rechtssicherheit beim Generieren von Medien. Die Videofunktion will Adobe bis Ende des Jahres einbauen.

Die Krux freilich bleibt stets die Kürze von erzeugten Clips – für „richtige“ Videos artet die Produktion auch weiterhin in Arbeit und hohem Stromverbrauch aus. Doch haben zumindest die Bildgeneratoren schon einen wichtigen Part weiterentwickelt: die konsistente Charakterreferenz. Dabei stellt die KI sicher, dass ein Protagonist auch in der nächsten Szene so aussieht wie zuvor in einem Referenzbild.

4 Runway:
faz.net/pro243839

5 Beitrag:
faz.net/pro243840

6 Beitrag:
faz.net/pro243841

7 Beitrag:
faz.net/pro243842

PRO WELTWIRTSCHAFT

Alles Wichtige zu Globalisierung, Sicherheit und Ressourcen.



Jetzt neu: F.A.Z. PRO Weltwirtschaft
Ihr Experten-Update rund um globale Ökonomie als wöchentliches Briefing und als Website. Inklusive freiem Zugang zu allen Artikeln auf unserem Onlineportal FAZ.NET.



Jetzt 2 Monate
kostenfrei testen!

F.A.Z. PRO Weltwirtschaft beleuchtet in einem neuen Informationsangebot relevante Ereignisse der Geoökonomie. F.A.Z.- und Branchen-Experten analysieren die Entwicklungen gezielt hinsichtlich ihrer Bedeutung und der daraus resultierenden Herausforderungen in Deutschland – für ein besseres Verständnis der Zusammenhänge globaler Entwicklungen.

Jetzt F.A.Z. PRO Weltwirtschaft und freien Zugang zu FAZ.NET
2 Monate kostenfrei testen unter faz.net/pro-weltwirtschaft



Zitat der Woche

„Dass Intel sein Projekt in Magdeburg auf Eis legt und vielleicht ganz aufgeben wird, ist enttäuschend für die Region. Für Deutschland insgesamt ist es aber kein Verlust, weil dieses Projekt zehn Milliarden Euro an Subventionen benötigt hätte.“



Clemens Fuest, Präsident des Ifo Instituts in München (dpa)

Frankfurter Allgemeine PRO DIGITALWIRTSCHAFT



<https://www.faz.net/pro/digitalwirtschaft/>

VERANTWORTLICHER REDAKTEUR:

Dr. Holger Schmidt

KONTAKT:

deconomy@faz.de

COPYRIGHT:

Frankfurter Allgemeine Zeitung GmbH,

Pariser Straße 1,
60486 Frankfurt am Main,
Hessen, Germany

+49 (0)69 7591-0

info@faz.net

HRB 7344, Amtsgericht Frankfurt am Main,
USt.-IDNr.: DE 114 232 732

GESCHÄFTSFÜHRER:

Thomas Lindner (Vorsitzender),

Dr. Volker Breid

HERAUSGEGEBEN VON

Gerald Braunberger,

Jürgen Kaube,

Carsten Knop,

Berthold Kohler