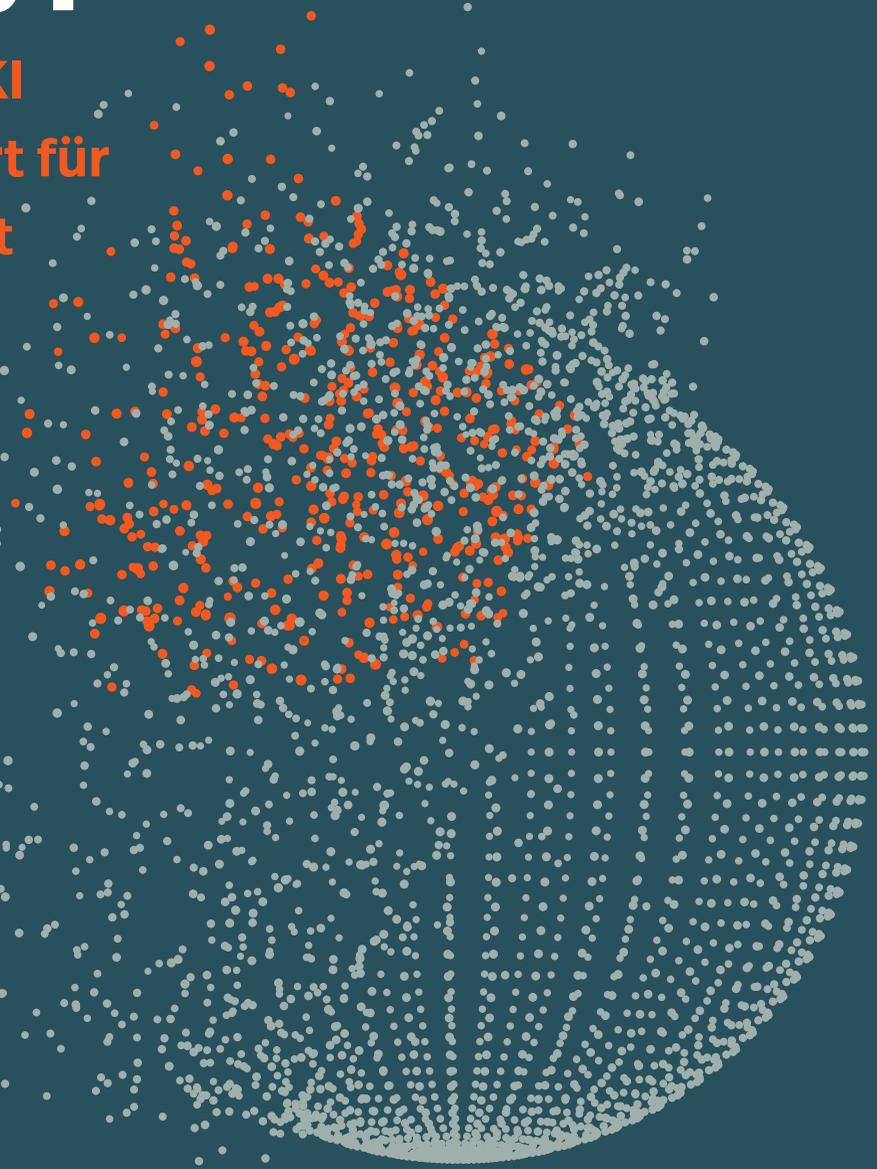


# Jenseits des Hypes

Wie generative KI  
echten Mehrwert für  
Europa generiert





# VORWORT

Die allermeisten von uns dürften den folgenden Satz kennen: „KI ist eine überbewertete Technologie mit wenigen Anwendungsfällen für Unternehmen. Der KI-Markt ist eine Blase.“ Ersetzt man „KI“ durch „das Internet“, dann landet man ironischerweise bei der Debatte von vor 25 Jahren. Mit der Dotcomblase gab es zwar in der Tat einen Dämpfer, doch allen Unkrenrufen zum Trotz gehören Alphabet, Amazon oder Meta heute zu den wertvollsten Unternehmen der Welt. Nur dass sie der KI-Chiphersteller NVIDIA inzwischen eingeholt hat.

Der Hype um (generative) KI fühlt sich darum wie ein Déjà-vu an. Ja, KI revolutioniert bereits die Geschäftswelt. Aber gleichzeitig sind die tatsächlichen Vorteile und Anwendungsfälle für viele Unternehmen noch schwer greifbar.

Unsere Publikation beschreibt deshalb, wo Europa in Sachen KI aktuell steht und erklärt, wie Unternehmen und Behörden die Technologie besser nutzen können. Wir konzentrieren uns dabei besonders auf Europas oft unterschätzte Stärken: seine Daten und das enorme Fachwissen, das in unseren Unternehmen schlummert. Gelingt es, beides stärker zu hebeln, kann die EU ihren Platz in der Weltwirtschaft zurückerobern.

Die Veröffentlichung markiert auch eine Veränderung innerhalb unserer Organisation. Die Internet Economy Foundation (IE.F) wurde 2016 gegründet, um Europas Agenda für die digitale Wirtschaft mitzugestalten und dabei Startups und Scaleups, mutige Investitionen, digitale Infrastruktur und fairen Wettbewerb zu fördern. Doch trotz vieler Fortschritte ist Europa weiter zurückgefallen. Deshalb haben wir unsere Stiftung umbenannt und wollen uns verstärkt als unabhängige Stimme für die europäische technologische Souveränität einsetzen.

Der Weg ist klar: Während wir neue Vorschriften durchsetzen, um die Internetwirtschaft wieder zu öffnen, müssen die Europäer ihre Stärken besser ausspielen. Wir sind weiter überzeugt, dass die EU eine große Zukunft vor sich hat. Nur muss sie dafür Innovationen künftig schneller und mutiger vorantreiben. Deshalb haben wir uns in Innovate Europe Foundation (IE.F) umbenannt und laden Sie ein, sich gemeinsam mit uns für ein widerstandsfähiges und wohlhabendes Europa stark zu machen.



**Clark Parsons**  
Geschäftsführer  
Innovate Europe  
Foundation

# „Stärken stärken“

Drei Handlungsfelder für den Aufbau eines unabhängigen EU-Ökosystems

## 1. Wissenstransfer

Wir müssen das Domänenwissen der europäischen Industrie nutzen.

## 2. Abbau von Abhängigkeiten

Wir müssen souveräne Technologielösungen aus Europa gezielt stärken.

## 3. Antwort auf das Wettrüsten

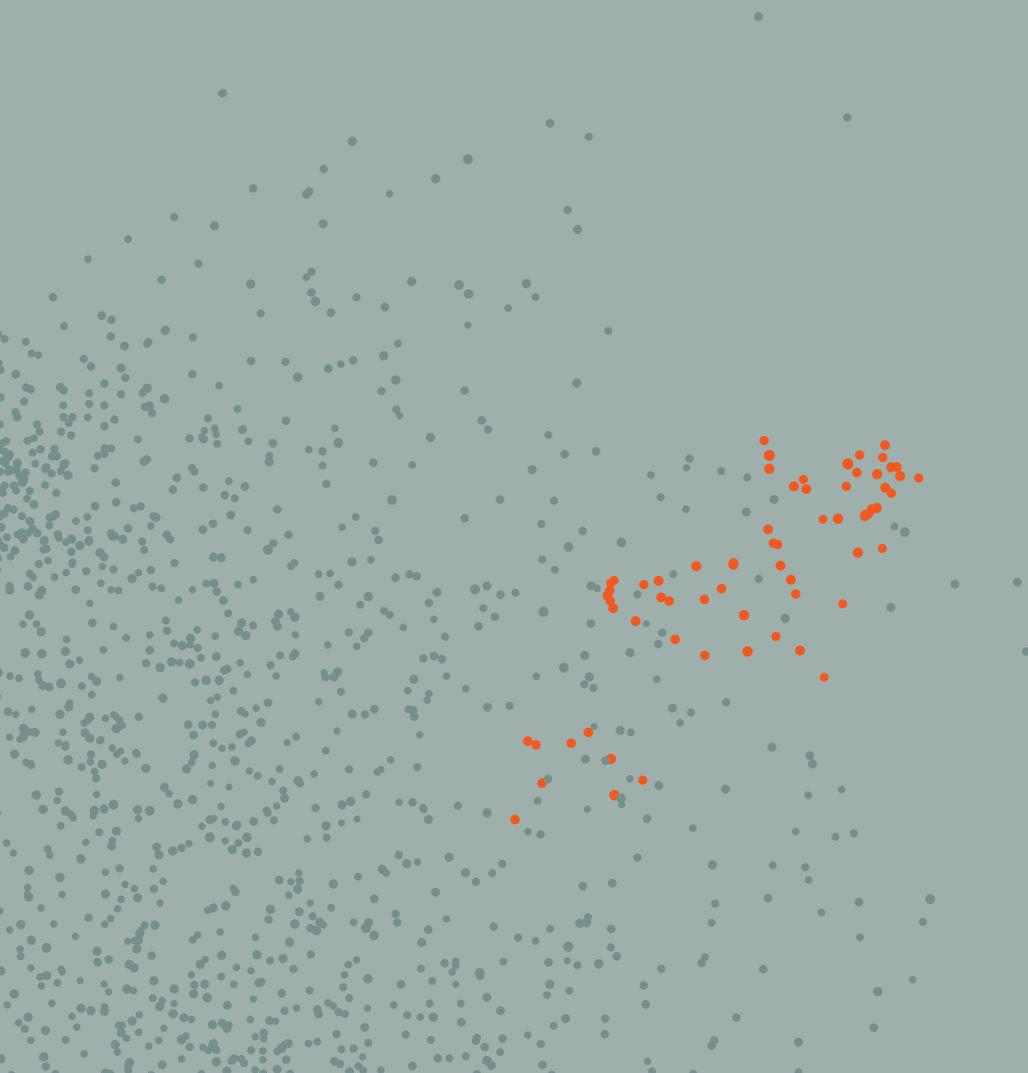
Wir müssen mehr privates Kapital für europäische KI-Innovationen mobilisieren.

# INHALT

<b>VORWORT</b>	<b>3</b>
<b>1</b> <b>EINLEITUNG: DAS GROSSE MISSVERSTÄNDNIS</b>	<b>6</b>
<b>2</b> <b>DIE DOMINANZ VON BIG TECH UND DER FALSCHER FOKUS AUF DAS EINE SUPERSPRACHMODELL</b>	<b>9</b>
<b>3</b> <b>DER PRAXISTEST: WARUM DIE BESTE TECHNOLOGIE OHNE DOMÄNENWISSEN INS LEERE LÄUFT</b>	<b>18</b>
<b>4</b> <b>„STÄRKEN STÄRKEN“: WIE DER AUFBAU EINES UNABHÄNGIGEN ÖKOSYSTEMS GELINGEN KANN</b>	<b>26</b>

# 1

## EINLEITUNG: DAS GROSSE MISSVERSTÄNDNIS



Der Hype um generative KI (Gen KI) ist noch nicht abgeflaut. Im Alltag macht sich aber eine erste Ernüchterung breit. Zwar haben viele Unternehmen und Behörden angefangen, KI-Sprachmodelle in ihre Prozesse zu integrieren. Der erhoffte Durchbruch mit einer breiten Implementierung in Unternehmen (B2B) und Behörden (B2G) ist aber (noch) nicht gelungen. Dies liegt weniger an der Technologie als an den falschen Erfolgskriterien, die an sie angelegt werden.

Viel zu oft stehen die auf den ersten Blick beeindruckenden Fähigkeiten der großen Allroundmodelle von OpenAI & Co im Vordergrund. Insbesondere Anwendungsfälle aus dem B2C-Bereich haben dadurch große mediale Aufmerksamkeit erzeugt. Die praktischen Anforderungen von Unternehmen und der öffentlichen Verwaltung geraten hingegen zunehmend in den Hintergrund. Viele professionelle Anwender sind deshalb bisher nicht über die Experimentierphase hinausgekommen. Oder aber sie leisten sich teure Schaufensterprojekte, von denen klar ist, dass sie auf absehbare Zeit nicht zur erhofften Produktivitätssteigerung führen.

Das vorliegende Policy Paper wagt vor diesem Hintergrund eine ehrliche Bestandsaufnahme. Wir gehen der Frage nach, warum die Technologie trotz des berechtigten Hypes den Alltag von Unternehmen und Behörden (noch) nicht revolutioniert hat. Und wir zeigen auf, welche Rahmenbedingungen erfüllt sein müssen, damit das enorme Potenzial von Gen KI endlich ausgeschöpft werden kann.

Ein wesentlicher Teil der Erklärung besteht darin, dass zu viele Hoffnungen in die Technologie alleine gesetzt werden. Mindestens so wichtig für den Erfolg in der

**Die praktischen Anforderungen von Unternehmen und der öffentlichen Verwaltung geraten zunehmend in den Hintergrund.**

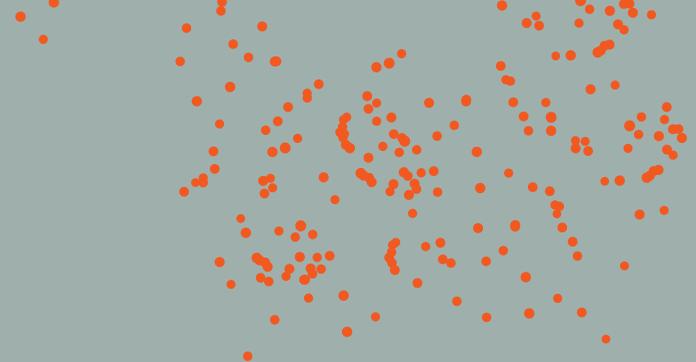
Praxis sind aber die richtige Einbettung in unternehmensinterne Prozesse und die Verknüpfung mit dem Domänenwissen der Mitarbeiter. Unter Domänenwissen verstehen wir das über Jahre hinweg aufgebaute Spezialwissen einer Fachkraft in einer bestimmten Branche oder einem Wissensbereich. Ohne genau diese Rückkoppelung an die Praxis kann auch die beste KI ihre Wirkung nicht entfalten.

In Kapitel 2 beleuchten wir die Ursachen für diese einseitige Fokussierung auf die Technologie. Wir zeigen, wie sich der Wettlauf um das größte Sprachmodell beschleunigt hat und alle Aufmerksamkeit auf sich zieht. Dominiert wird er wie so oft von den großen Tech-Playern aus den USA, die in puncto Investitionen, Partnerschaften und Infrastruktur in einer eigenen Liga spielen. Pessimisten hat diese Dominanz zu der altbekannten Klage über die hoffnungslos abgehangenen Europäer verleitet. Aus unserer Sicht wird diese einseitige Betrachtung der Realität nicht gerecht.

In Kapitel 3 zeigen wir anhand von Praxisbeispielen, dass nicht zwangsläufig das vermeintlich mächtigste, weil größte Large Language Model (LLM) den größten Nutzen stiftet. Oft können kleinere und auf eine bestimmte Anwendung zugeschnittene Modelle einen deutlich höheren Mehrwert in der Praxis erzielen. Allgemeine Benchmarktests führen deshalb leicht in die Irre, weil sie die konkrete Anwendersituation zu wenig oder gar nicht erst berücksichtigen.

Trotzdem muss Europa seine Anstrengungen beschleunigen. Wie wir in unseren Empfehlungen in Kapitel 4 ausführen, muss das Domänenwissen der europäischen

Industrie besser genutzt und einseitige Abhängigkeiten bei der Infrastruktur abgebaut werden. Gleiches gilt für die Kapitalmarktlücke: Sie muss endlich geschlossen werden. Nur ein weitgehend unabhängiges Ökosystem aus Technologie, Domänenwissen und Infrastruktur kann in Europa die Basis für eine KI-getriebene Wertschöpfung bilden und die Zukunftsfähigkeit der Wirtschaft sichern.



# 2

## **DIE DOMINANZ VON BIG TECH UND DER FALSCHER FOKUS AUF DAS EINE SUPER- SPRACHMODELL.**



Angesichts des gigantischen Wettlaufs um die schlagkräftigste generative KI ist es nicht verwunderlich, dass sich aktuell die ganze Aufmerksamkeit auf die Riesensprachmodelle richtet. Im Zentrum stehen dabei – wie so oft – die Tech-Giganten, und ganz vorne: Alphabet, Amazon, Meta und Microsoft. Sie kämpfen erbittert um die Vormachtstellung und verkünden beinahe täglich neue Fortschritte. Zusammen wollen die vier Unternehmen allein in 2024 200 Milliarden US-Dollar für Datenzentren, Chips und anderes Equipment für die Anwendung und den Betrieb von Gen-KI-Modellen investieren – ein neuer Rekord und eine Steigerung um 45 Prozent im Vergleich zum Vorjahr. Es gilt als wahrscheinlich, dass die Investitionsausgaben (CAPEX) des Quartetts in den nächsten fünf Jahren auf bis zu einer Billion US-Dollar ansteigen werden. →A

Die Investitionen beschränken sich aber nicht nur auf den Ausbau der eigenen Infrastruktur. Big Tech betätigt sich immer stärker auch als Wagniskapitalgeber und sichert sich dadurch Zugang zur neuesten Technologie aufstrebender Startups. Die Beteiligungen folgen dabei weniger der Logik klassischer Finanzinvestitionen, sondern zielen auf eine Festigung der Marktdominanz ab. Die Partnerschaft zwischen Microsoft und OpenAI beziehungsweise zwischen Microsoft und dem französischen Startup Mistral sind nur zwei Belege für diesen Trend. Auch Google und Amazon unterstützen das kalifornische Startup Anthropic mit Milliardensummen.

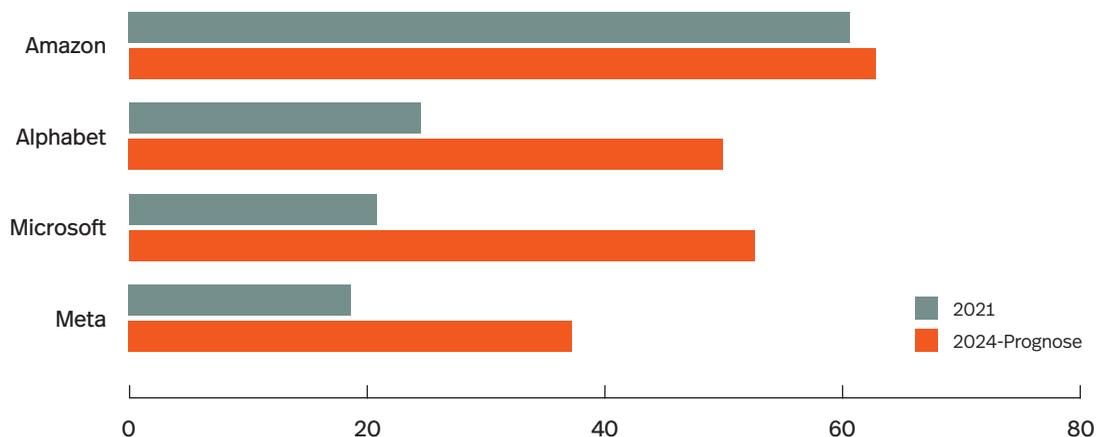
Das neue Engagement schlägt sich auch in den Venture-Capital(VC)-Statistiken nieder. Laut der Analysefirma

---

## A 200-Mrd.-US-Dollar-Wette

---

Investitionsausgaben der vier KI-Giganten [USD Mrd.]



---

Quelle: Capital IQ, Forbes

---

2  
Die Dominanz von Big Tech  
und der falsche Fokus auf das  
eine Supersprachmodell

Pitchbook stammen zwei Drittel der insgesamt 27 Milliarden US-Dollar, die 2023 in aufstrebende Gen-KI-Firmen flossen, von Alphabet, Amazon und Microsoft. Traditionelle Wagniskapitalgeber geraten demgegenüber zunehmend ins Hintertreffen. →B

Die Entscheidung für eine VC-Beteiligung und gegen eine Akquisition markiert dabei einen wesentlichen Unterschied im Vergleich zu früheren Innovationswellen. Große Übernahmen, wie sie Meta vor einigen Jahren mit WhatsApp oder Instagram tätigte, stehen aktuell nicht auf der Tagesordnung – wohl auch aus Furcht vor den Kartellwächtern. Die Sorge um die Unabhängigkeit der von Big Tech finanzierten Gen-KI-Startups bleibt trotzdem bestehen.

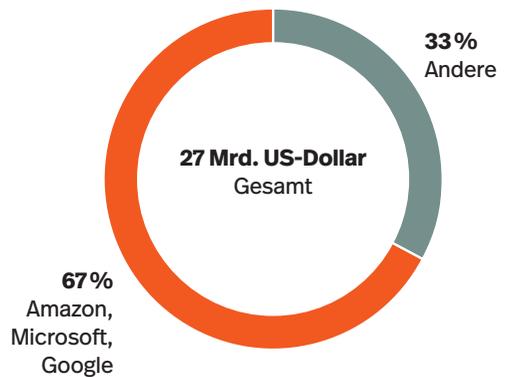
Angesichts ihrer dominanten Marktposition gelingt es den Tech-Giganten auch, die raren KI-Talente im großen Stil abzuwerben. Laut der Forschungsfirma Live Data Technologies stellten Alphabet, Amazon, Apple, Meta und Microsoft seit 2022 mindestens 30 führende KI-Experten von OpenAI, Anthropic und Cohere ein. Erst im März 2024 gab Microsoft außerdem bekannt, fast die ganze Belegschaft von Inflection AI zu übernehmen. Darunter befindet sich auch Mustafa Suleyman, Mitbegründer und ehemaliger Leiter der Abteilung für angewandte KI bei DeepMind. Nach seinem Ausscheiden bei DeepMind gründete er 2022 gemeinsam mit weiteren Partnern das Unternehmen Inflection AI, das sich auf maschinelles Lernen und generative KI spezialisiert hat. Im Juni 2023 erhielt Inflection AI eine der größten KI-Finanzierungen, die ein Volumen von 1,3 Milliarden US-Dollar erreichte. Zu den Hauptinvestoren zählte Microsoft. Suleyman übernimmt die Leitung des neuen

---

**B Startup-Finanzierung statt Startup-Übernahme**

---

Big-Tech-Anteil an US-Wagniskapitalinvestitionen  
in generative KI, 2023



---

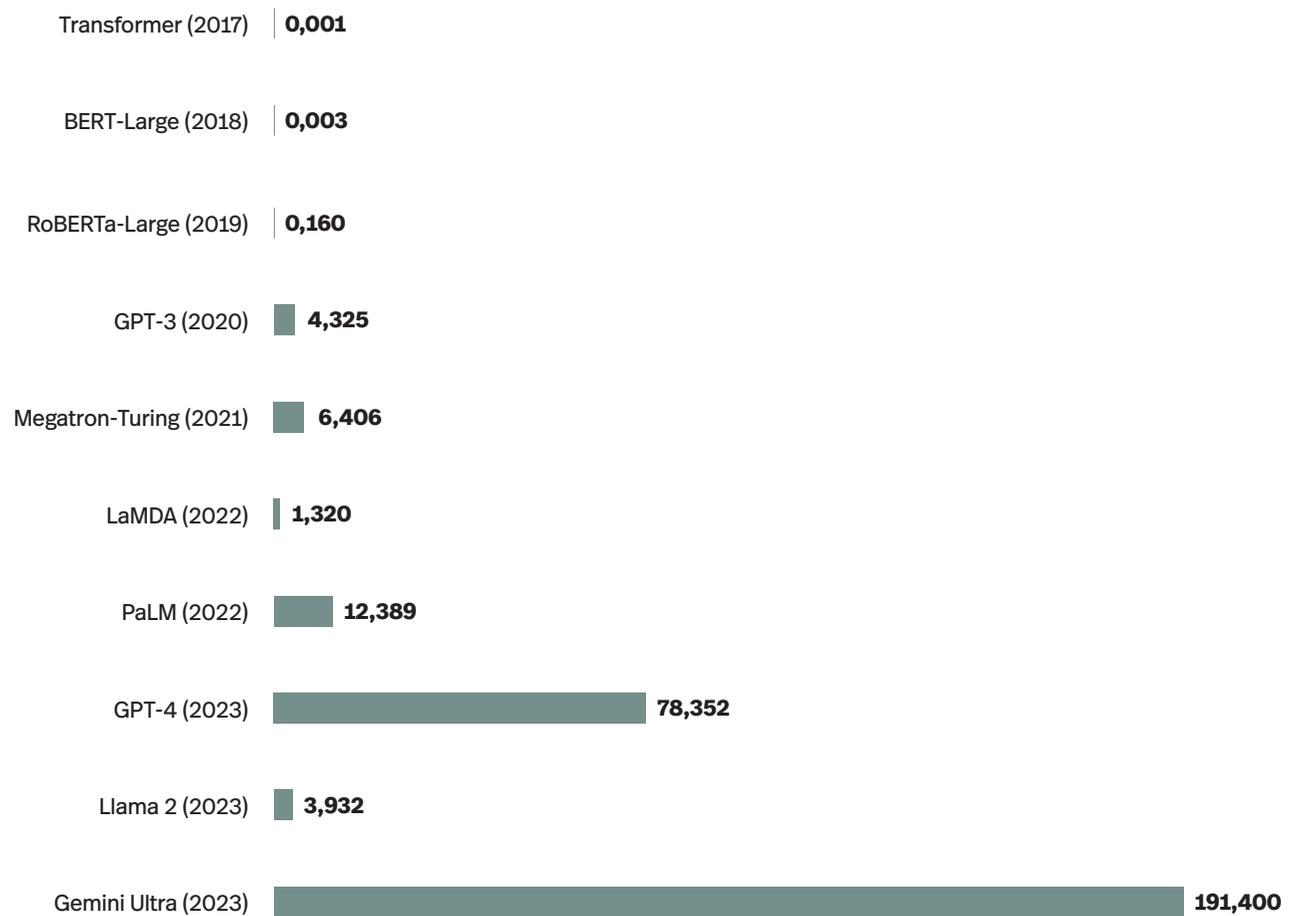
Quelle: Pitchbook

---

## C Teures Unterfangen

---

Geschätzte Trainingskosten für ausgewählte LLMs [USD Mio.]



---

Quelle: Stanford University

Konzernbereichs Microsoft AI. Somit kommt es zu einer weiteren Machtkonzentration in der internationalen KI-Forschung und -Entwicklung.

Die Kartellbehörden wollen dieser Entwicklung nicht tatenlos zusehen. Jonathan Kanter, der oberste Monopolwächter im US-Justizministerium, erklärte zum Beispiel erst kürzlich in einem Interview mit der Financial Times, bei sogenannten „Acqui-Hires“ künftig genauer hinsehen zu wollen. Gemeint ist das akquisitionsähnliche Abwerben von Angestellten, mit denen dominante Unternehmen das geistige Eigentum von Startups einkaufen. Gleiches gilt für den Zugang zu Hardware und anderen kritischen Ressourcen, die den freien Wettbewerb einschränken.

Nicht nur die Ausgaben für Personal und Beteiligungen verschlingen immer größere Summen. Auch die Entwicklungskosten für die Super-LLMs schießen in die Höhe. Zwar geben einzelne Unternehmen nur selten Auskunft über die exakten Trainingskosten ihrer Modelle. Die Investitionen lassen sich aber zumindest annäherungsweise anhand der Trainingsdauer sowie des Typus, des Umfangs und der Nutzungsrate der Trainingshardware schätzen. Auf dieser Basis kommt die Stanford University in Zusammenarbeit mit dem KI-Forschungsinstitut Epoch AI beispielsweise auf Trainingskosten für OpenAIs GPT-4 und Googles Gemini Ultra von circa 78 Millionen US-Dollar beziehungsweise 191 Millionen US-Dollar. Die Trainingskosten für das ursprüngliche Transformer-Modell, das die Architektur einführte, die praktisch jedem modernen LLM zugrunde liegt, beliefen sich 2017 auf rund 900 US-Dollar. →C

Dabei handelt es sich wohlgerne nur um die Kosten für gelungene Trainingsläufe. Die einhundert gescheiterten Experimente, die einem erfolgreichen Ablauf in der Regel vorausgehen, sind in dieser Rechnung noch gar nicht eingepreist. Auf diese Weise entsteht schnell ein gigantischer Kapitalbedarf, obwohl unklar ist, ob der Nutzen proportional zu den Investitionen steigt. Die Entwicklung der nächsten Generation dürfte sogar noch teurer werden.

**Die Entwicklungskosten für die Super-LLMs schießen in die Höhe, obwohl unklar ist, ob der Nutzen proportional zu den Investitionen steigt.**

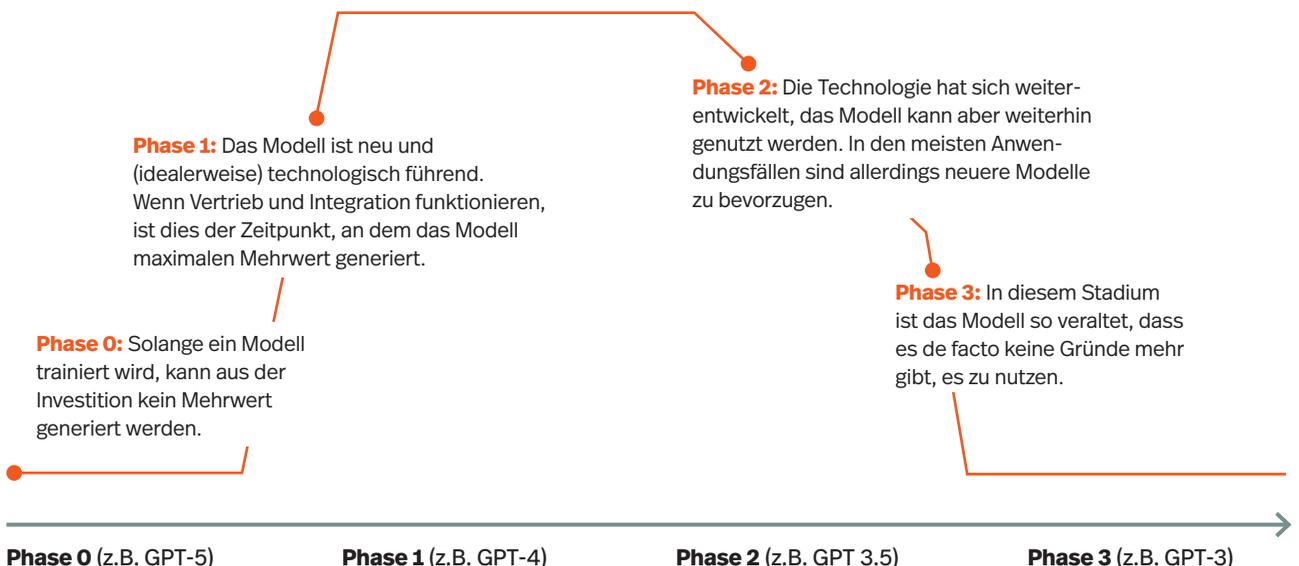
Problematisch sind die immensen Trainingskosten auch aus einem weiteren Grund: Jedes KI-Sprachmodell verfügt nur über eine relativ kurze Halbwertszeit, die zuletzt rund ein Jahr betrug. Danach stiftet das LLM kaum noch einen wirtschaftlichen Mehrwert. Der Grund: Jedes Nachfolgemodell, das auf einer neuen Softwarearchitektur basiert, muss auch von Grund auf neu trainiert werden. Das unterscheidet Gen-KI-Produkte ganz wesentlich von vorherigen Softwareinnovationen, die über Jahre oder gar Jahrzehnte hinweg das Grundfundament für alle folgenden Produktpassungen bildeten.

Auch die zum Training notwendige Hardware unterliegt einem relativ kurzen Lebenszyklus. Im Extremfall bleibt so von den hundert Millionen investierten Euro in Chips, Infrastruktur und den laufenden Betrieb auf mittlere Sicht nicht viel übrig – abgesehen von Reputationsgewinn, Kundenstamm und Nutzerdaten (bei Modellen ohne Datensouveränität). →D

Die sehr hohen Kosten für das Training und den Betrieb der Supersprachmodelle stellen aktuell eine große Anwendungsbarriere in der Praxis dar. Es ist jedoch zu be-

## D Kurze Halbwertszeit

Generierter Mehrwert durch Investitionen in KI-Sprachmodelle



Quelle: Innovate Europe Foundation

2  
Die Dominanz von Big Tech  
und der falsche Fokus auf das  
eine Supersprachmodell

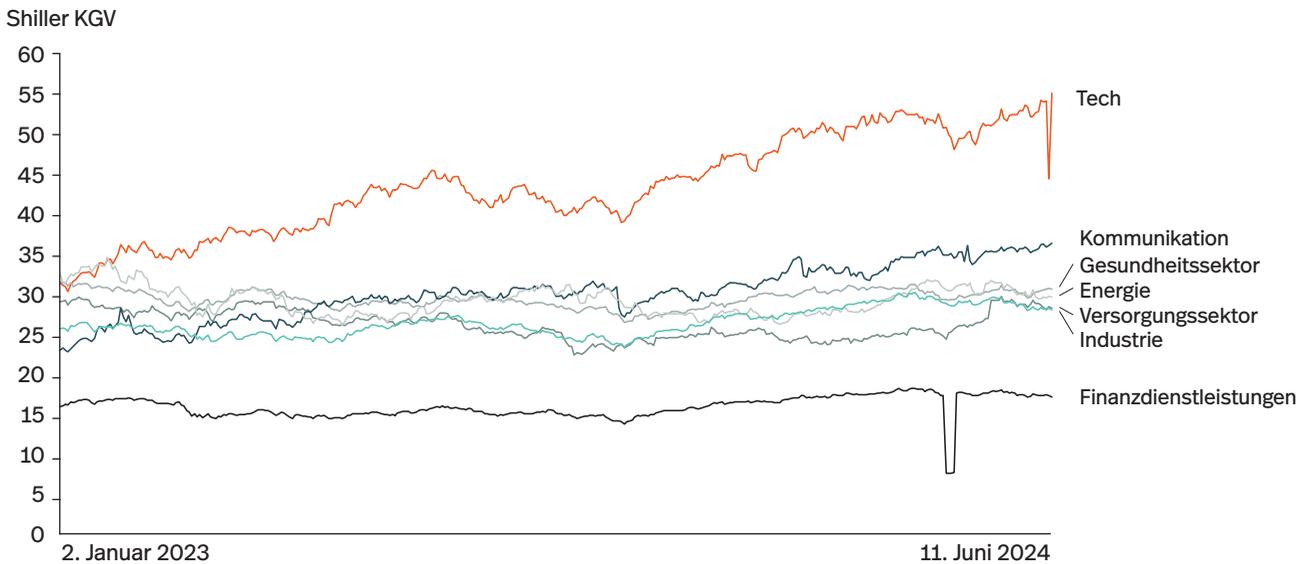
rücksichtigen, dass die Zahlungsbereitschaft auf der Nachfragerseite noch begrenzt ist. Die großen Tech-Firmen geben keine Auskunft über den Monetarisierungserfolg ihrer Modelle. Allerdings deuten Schätzungen darauf hin, dass die Erwartungen bisher nicht erfüllt wurden.

Die drei dominanten Cloud-Anbieter Alphabet, Amazon und Microsoft profitieren dennoch bereits jetzt von der Gen-KI-Revolution. Sie kontrollieren die Infrastruktur, die sowohl für das Training als auch für den Einsatz der

Sprachmodelle benötigt wird. Kein anderes Unternehmen kann in vergleichbarem Umfang die benötigte Rechenleistung zur Verfügung stellen. Diese Vormachtstellung im Cloud-Bereich erklärt zusätzlich zu deren Dominanz im Bereich der KI-Modelle, warum der Marktwert der drei großen Anbieter seit November 2022 um 2,5 Billionen US-Dollar gestiegen ist. Auch der Kapitalmarkt im Allgemeinen sieht das größte Wertschöpfungspotenzial gerade nicht in der Industrie, sondern bei den Tech-Giganten, wie ein Blick auf die Kursentwicklungen in den zurückliegenden Monaten zeigt. →E

E Aktuelles KI-Wertschöpfungspotenzial aus Anlegersicht

Entwicklung des Shiller KGV Index seit 2023 für ausgewählte Sektoren



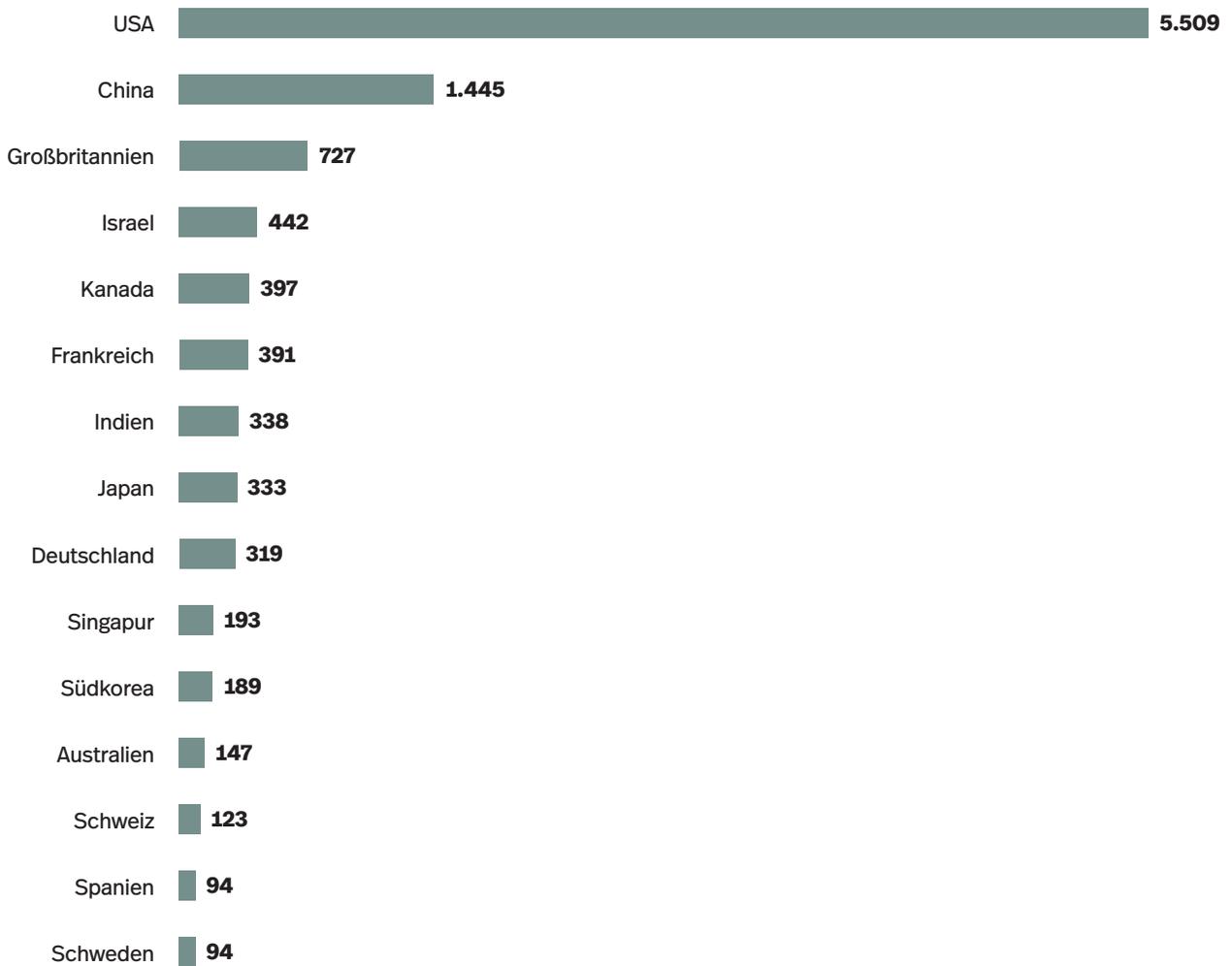
Quelle: GuruFocus

---

## F Unangefochtener Spitzenreiter USA

---

Anzahl der frisch finanzierten KI-Firmen nach Ländern, 2013–2023



---

Quelle: Stanford University

Angesichts der Vormachtstellung US-amerikanischer Tech-Unternehmen überrascht es wenig, dass sich auch am Kapitalisierungsgefälle zwischen den USA und Europa nichts geändert hat. Im Gegenteil: Das Ungleichgewicht nimmt weiter zu. Auch 2023 führte die USA das Ranking der Privatinvestitionen in generative KI mit großem Abstand an. Laut einer Analyse der Stanford University fiel das Investitionsvolumen dort zuletzt rund 30 mal höher aus als in der Europäischen Union und Großbritannien zusammen. Gleiches gilt für die Anzahl der neu gegründeten KI-Firmen. Auch hier liegt ein beträchtlicher Abstand zwischen den USA und dem Rest der Welt. Umso öfter drängt sich deshalb die altbekannte Frage auf: Wie kann es Europa schaffen, wieder zu den USA aufzuschließen? Aber ist dieser Fokus wirklich noch der richtige? Oder geht es nicht vielmehr darum, wie die EU ihre Position im internationalen Wettbewerb um Talente und Ressourcen stärken kann, ohne sich dabei einem Benchmark-Wettrennen zu unterwerfen? →F

Die technologische Führerschaft der US-Unternehmen ist unbestritten. Der Blick auf die aktuellen Investitionsvolumen legt sogar nahe, dass die Dominanz in den nächsten Jahren weiter steigen dürfte. Dennoch besteht kein Anlass für übertriebenen Pessimismus. Parallel zu den mächtigen Allroundmodellen haben auch die großen Tech-Firmen verstärkt damit begonnen, an kleineren und anwendungsorientierteren Modellen zu arbeiten. Wie im nächsten Kapitel dargelegt, ist der Durchbruch in der Praxis maßgeblich von der Einbettung dieser (kontextbezogenen) Spezialmodelle in den Unternehmensalltag abhängig.

**Der Durchbruch  
in der Praxis hängt  
maßgeblich von  
der Einbettung  
kontextbezogener  
Spezialmodelle  
in den Unter-  
nehmensalltag ab.**



3

**DER PRAXISTEST:  
WARUM DIE BESTE  
TECHNOLOGIE OHNE  
DOMÄNENWISSEN  
INS LEERE LÄUFT**

Die Erfahrung aus früheren Umbruchzeiten zeigt, dass es dauern kann, bis das Potenzial einer neuen Technologie auf die ganze Wirtschaft übergreift. Die durchschnittliche Firma und der oder die durchschnittliche Angestellte benötigen Zeit, um bestimmte Lern- und Erfahrungsprozesse zu durchlaufen. Es müssen Experten ausgebildet werden, die in der Lage sind, neue Technologien anzupassen, zu integrieren und die Wertschöpfungsprozesse neu definieren zu können. Die Vergangenheit hat gezeigt, dass die Durchdringung dieser Entwicklungen einige Zeit in Anspruch nehmen kann. Vor dem Beginn der PC-Ära verstrich mindestens eine Dekade, bis sich Computer im großen Umfang im Büroalltag durchsetzen konnten. Auch bei der Nutzung umfangreicher Gen-KI-Anwendungen wird es noch einige Zeit dauern, bis sich die hohen Produktivitätserwartungen tatsächlich in der Praxis manifestieren.

Richtig ist: Es gibt bereits vielversprechende Use Cases und beinahe täglich werden es mehr. Vergleicht man die Wirtschaftssektoren allerdings miteinander, so sind es laut einer Erhebung des amerikanischen Census Bureau vor allem Unternehmen aus der Informationsindustrie, die derzeit überdurchschnittlich stark mit generativer KI experimentieren. Demgegenüber stehen Unternehmen aus dem Verarbeitenden Gewerbe, in denen die Implementierung eher schleppend anläuft. Am Wertschöpfungspotenzial der KI-Sprachmodelle zweifeln aber die wenigsten.

Dass die Unternehmen generative KI nicht schon in viel größerem Umfang in ihre Abläufe integrieren, liegt an einer ganzen Reihe handfester Hürden. Am schwersten wiegt dabei das fehlende Fachwissen in den Anwender-

unternehmen, wie aus einer Umfrage des IT-Konzerns IBM hervorgeht. Auf Platz zwei und drei rangieren eng mit diesem Problem verwandte Themen: Entweder ist die Datenkomplexität in den Unternehmen zu hoch, oder aber die Projekte lassen sich nicht adäquat in interne Prozesse integrieren und skalieren. →G

Das fehlende Fachwissen und die Schwierigkeiten bei der Prozesseinbettung erklären auch eine weitere Schieflage: Aktuell sind größere Unternehmen mit 250 und mehr Angestellten unter den Early Adoptern deutlich überrepräsentiert. Sie verfügen noch am ehesten über die Ressourcen, KI-Experten an sich zu binden und die nötigen Investitionen in Prozesse und Infrastruktur zu tätigen, beziehungsweise das notwendige Kapital, um solche Expertise als Dienstleistung einzukaufen. Kleineren Firmen fällt dies deutlich schwerer.

Der Blick auf die Praxis zeigt außerdem, dass ein größeres und mächtigeres KI-Sprachmodell für sich genommen nicht die Anwenderhürden beseitigen würde. Im Gegenteil: Das alleinige Vertrauen auf Super-LLMs könnte den Graben zwischen Technologiepartnern auf der einen Seite und der Industrie auf der anderen Seite sogar noch vertiefen. Nicht nur das Training der großen KI-Sprachmodelle ist nämlich teuer, auch der laufende Betrieb erfordert enorme Rechenkapazitäten. Viele Anwender schreckt das ab.

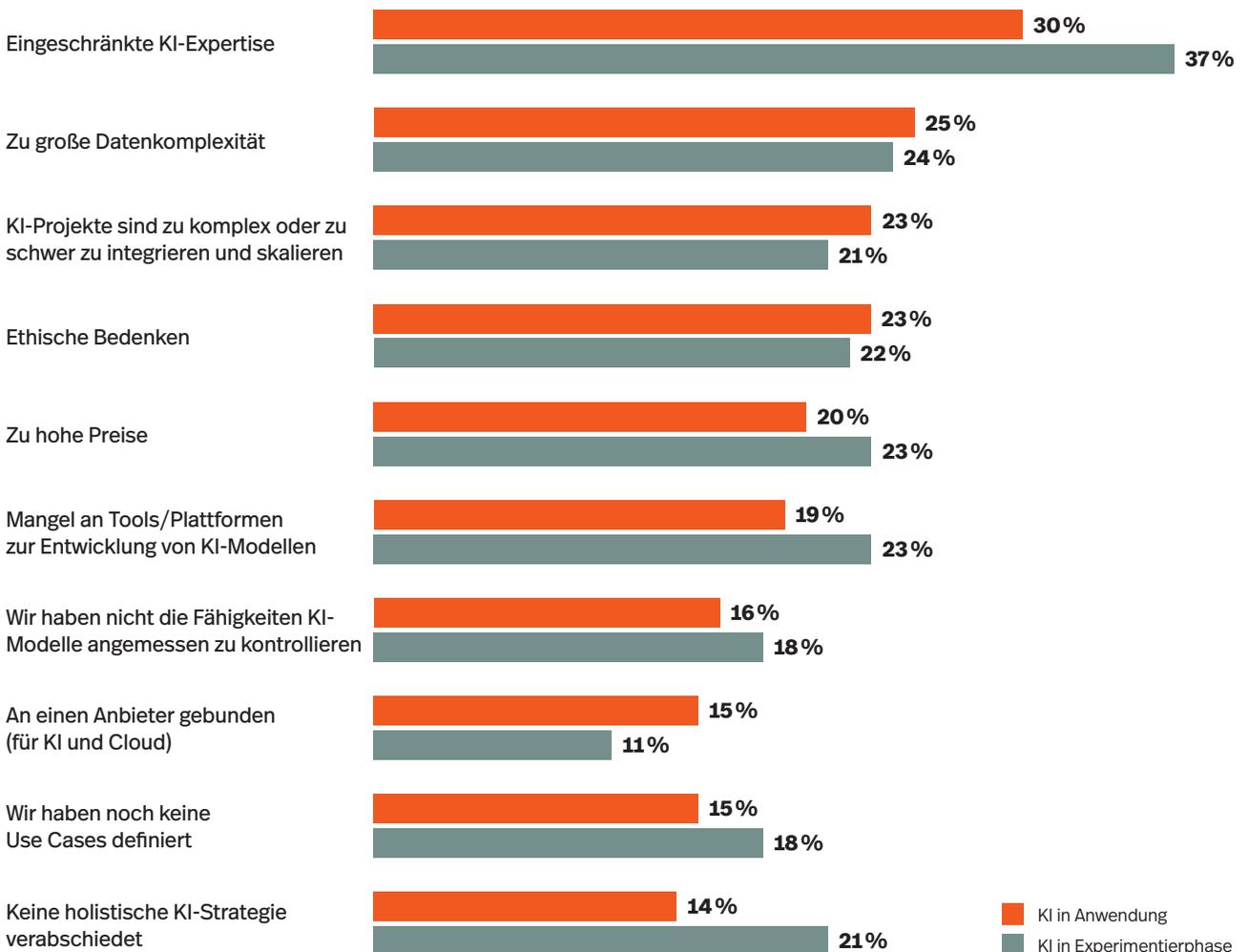
Dabei ist Größe allein ohnehin nicht das einzige Kriterium. Analysen zeigen, dass Modelle mit einer kleineren Anzahl an Parametern durchaus an die logischen Fähigkeiten größerer Modelle heranreichen und sie in entscheidenden Bereichen sogar übertreffen können.

---

## G Fehlendes Fachwissen

---

Die größten Hindernisse für eine erfolgreiche Gen-KI-Anwendung in der Praxis



Stand: November 2023

---

Quelle: IBM

Generell versteht man unter Parametern numerische Werte, die während des Trainings erlernt werden. Sie bestimmen, wie das jeweilige Modell die Eingabedaten interpretiert und zu welchen Prognosen es kommt. Aber auch hier gilt: Mehr ist nicht gleich besser. Die Qualität der Datensätze spielt eine wichtige Rolle. Wenn zum Beispiel repetitive oder fachfremde Daten auf geschickte Art und Weise sorgfältig herausgefiltert werden, können kleinere Modelle durchaus eine ähnliche Perfor-

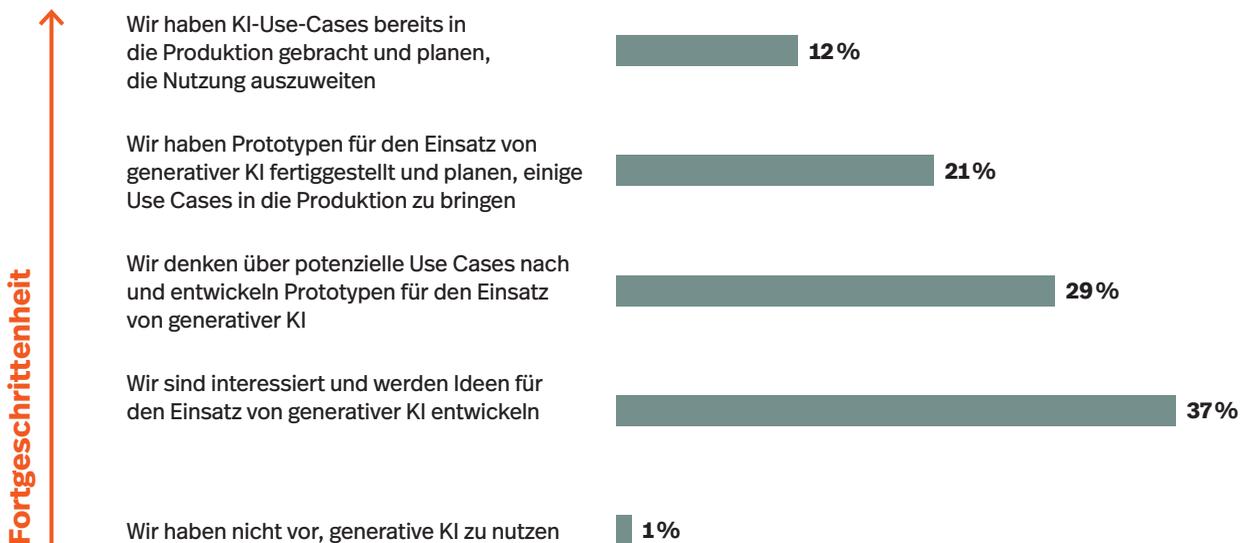
mance erzielen. Sprachmodelle, die mit sorgfältig kuratierten Datensets trainiert werden, sind nicht nur kostengünstiger, sondern haben noch einen weiteren Vorteil. Ein Kunde kann sie unter Umständen mit weniger Rechenleistung und energieeffizienter „on premise“ betreiben, also in der eigenen Hardwareumgebung, anstatt Informationen in die Cloud zu schicken. Gerade für datenschutzbewusste Anwender dürfte dies entscheidend sein. →H

---

## H Zögerlicher Start

---

Aktueller Stand in Sachen Gen-KI-Anwendung über alle Branchen hinweg



---

Quelle: Coleman Parkes Research

Bisherige Use Cases belegen außerdem, dass Unternehmen gut damit beraten sind, ihr KI-Sprachmodell anhand spezifischer Anwendungserfordernisse auszuwählen. Es kommt beispielsweise vor, dass ein Modell in den allgemeinen Benchmarkings deutlich besser abschneidet als ein Konkurrent, in der für die spezifische Anwendung relevanten Kategorie jedoch schlechtere Ergebnisse liefert. In diesem Fall wäre die Entscheidung für das nominell „bessere Modell“ die falsche, weil das gewählte Modell für den praktischen Anwendungsfall weniger geeignet ist. Ein gutes Praxisverständnis und eine umfassende Evaluation von Benchmarkings entscheidet deshalb über Erfolg oder Misserfolg in der Praxis.

Dazu passt, dass nach Ansicht von Forschern der Stanford University bisherige Benchmarks, mit denen die Leistungsfähigkeit eines Modells in verschiedenen Anwendungsbereichen gemessen wird, ohnehin an ihre Grenzen gestoßen sind. Zur Leistungsbewertung werden sie immer seltener herangezogen. Unter Umständen wurden die Modelle auch bereits mit speziellen Fragenkatalogen auf den jeweiligen Benchmarktest hin trainiert, was die Aussagekraft der Ergebnisse schmälert. Es hilft daher, sich die für den jeweiligen Use Case relevanten Bereiche genau anzusehen und nicht sofort auf ein Allroundmodell zu setzen. Darunter könnten zum Beispiel Modelle mit einem Schwerpunkt auf „Befolgung von Bildanweisungen“, „moralischem Denken“ oder „Planung“ fallen. Letztendlich obliegt es dem Anwender zu definieren, in welchem Bereich welche Anforderungen notwendig sind, um das für diesen Kontext beste Modell zu identifizieren, und zusammen mit dem entsprechenden Technologiepartner ein proprietäres Feintuning vornehmen.

Die Auswahl des Sprachmodells ist, wie bereits erwähnt, ohnehin nicht der einzige und alles entscheidende Erfolgsfaktor. Einen immer größeren Stellenwert nimmt das individuelle Feintuning je Fachbereich und Unternehmen ein. Auf diese Weise kann jeder Anwender am Ende sein eigenes oder mehrere LLMs gleichzeitig operieren, auf die ausschließlich die eigenen Mitarbeiter Zugriff haben. Dadurch ließe sich die volle Souveränität sicherstellen und der Abzug von Domänenwissen über Interaktionsdaten verhindern. Die bisherigen Use-Case-Erfahrungen legen nahe, dass genau eine solche Kombination aus Technologie – gepaart mit Domänenwissen und dem Verständnis für unternehmensspezifische Prozesse – den größten Mehrwert generiert. In der Praxis kann dies wie in den folgenden drei Use Cases aussehen:

## CASE 1

# Wie ein virtueller KI-Assistent die Produktions- profitabilität tatsächlich steigern kann

Die sich radikal wandelnde Industrielandschaft erfordert eine kontinuierliche Optimierung des Produktionsprozesses, um Ausfallzeiten zu reduzieren und Kosten zu senken. Der Einsatz von KI-getriebenen virtuellen Assistenten kann dabei eine hilfreiche Unterstützung bieten. Mit ihnen lassen sich Routineaufgaben automatisieren und durch Echtzeit-Datenabgleiche bessere Entscheidungen treffen. Damit ein solcher virtueller Assistent aber tatsächlich die Produktionsprofitabilität in der Praxis steigern kann, muss er passgenau auf die jeweiligen Abläufe abgestimmt sein. Fabrikangestellte haben zum Beispiel oft schlicht nicht die Zeit, Anfragen manuell in ein Computer-Interface einzugeben. Darum ist es von Vorteil, wenn das Personal direkt mit dem Chatbot interagieren kann.

Darüber hinaus sind in vielen Fällen internationale Experten in den Produktionsprozess involviert. Die Möglichkeit, mit dem KI-Assistenten in verschiedenen Sprachen zu kommunizieren, ist somit ein entscheidender Faktor, auch wenn die Betriebsanleitung zum Beispiel nur auf Deutsch oder Englisch verfügbar ist. Gleiches gilt für die Kommunikation via Bilder: Sie ist im vorliegenden Use Case von größerer Bedeutung als andere Hauptkategorien, die in traditionellen Benchmarktests geprüft werden. Funktioniert sie, kann ein Fabrik spezialist zum Beispiel die Standposition eines Roboters fotografieren und den KI-Assistenten fragen, ob sie sicher ist. Die Effizienzgewinne und die Reduzierung von Ausfallzeiten durch ein solches eng mit dem Produktionsprozess verzahntes KI-Spezialmodell liegen auf der Hand.

## CASE 2

# Echter Mehrwert im Kundenservice statt kurzfristiger „Effizienzgewinne“

Der Einsatz von generativer KI zur Optimierung des Kundenservices gehört zu den Parade-Use-Cases, die bereits seit längerer Zeit erprobt werden. Die Idee dahinter: Mitarbeiter im Kundendienst können durch den Zugriff auf Echtzeit-Entscheidungshilfen, die auf dem Wissen der gesamten Fallhistorie einer Firma basieren, Fälle besser und effizienter lösen. Im Idealfall könnte ein solcher KI-Service-Agent bereits einen Großteil der Standardanfragen im Vorfeld beantworten. Dadurch stünden den Angestellten mehr Ressourcen zur Verfügung, um sich um besonders anspruchsvolle Fälle zu kümmern oder individuelle Unterstützung zu leisten.

In der Praxis erweist sich die Einbettung des Service-Agenten in vielen Fällen als deutlich komplexer. Das zeigt das Beispiel einer KI-unterstützten Pannenhilfe. Die Analyse der Kundendaten hat gezeigt, dass je nach Kundengruppe eine unterschiedliche Unfallab-

wicklung gewünscht wird. Während die Gen Z durchweg Selbsthilfetools bevorzugte, wollten andere ab einer bestimmten Etappe direkt mit einem Kundendienstmitarbeiter kommunizieren.

Die Herausforderung bestand folglich darin, den KI-Service-Agenten auf diese wichtige Entscheidung hin zu trainieren und keine kurzfristigen „Effizienzgewinne“ durch einen Dauereinsatz anzustreben. Erst nach dieser Optimierung konnte die Produktivität auch tatsächlich gesteigert und eine Kundenzufriedenheit von 98 Prozent erreicht werden.

### CASE 3

## Das enorme Potenzial der Mensch-Maschine-Zusammenarbeit im Compliance-Bereich

Compliance-Überprüfungen gehören zu den schwierigsten Use Cases für KI-Sprachmodelle, die es aktuell zu lösen gilt. Und das gleich aus mehreren Gründen: Zum einen müssen in aller Regel streng vertrauliche Verträge und Dokumente überprüft werden. Zum anderen sind die relevanten Inhalte hochkomplex und sehr fachspezifisch. Bei Fehlern droht außerdem ein enormes Reputationsrisiko. Einen Audit ganz an eine KI outzusourcen, ist vielen Unternehmen aus verständlichen Gründen zu riskant.

In der Praxis hat sich jedoch herausgestellt, dass eine Mensch-Maschine-Zusammenarbeit enormen Mehrwert stiften kann. Das zeigt unter anderem das Beispiel der neuen EU-Verordnung DORA (Digital Operational Resilience Act). Mit ihr soll ein einheitlicher Rahmen für das effektive und umfassende Management von Cybersicherheitsrisiken auf den Finanzmärkten etabliert werden. Konkret: Die betroffenen Unternehmen müssen belegen, dass selbst im Fall einer schwerwiegenden Betriebsunterbrechung keine Gefährdung der

Sicherheit der Informationssysteme droht. Bis vor Kurzem war das ein aufwendiger und vor allem kostenintensiver Prozess, bei dem pro Unternehmen in manchen Fällen Hunderte von Verträgen überprüft werden mussten. Erste Anwendungstests zeigen nun, dass sich die Compliance-Überprüfung mithilfe von KI um ein Vielfaches effizienter gestalten lässt. Dafür müssen aber eine Reihe von Bedingungen erfüllt sein: Benötigt wird zunächst ein Special-Purpose-Modell, das exakt auf das Erkennen der relevanten Vertragspassagen hin trainiert wurde. Herkömmliche Allzweckmodelle sind für diese Anforderungen nicht geeignet.

Von ebenso großer Bedeutung ist die Interaktion zwischen Fachkraft und KI. Das Modell muss auf alle Interpretationsspielräume und inhärente Unsicherheiten transparent hinweisen. Nur so kann der Compliance-Experte die letztinstanzliche Einordnung erfolgreich vornehmen, ohne dass Mehrarbeit entsteht oder Entscheidungen auf falscher Basis getroffen werden. Das heißt auch, Feedbackloops müssen so in den Workflow eingebettet werden, dass das Modell während der Nutzung kontinuierlich anhand des bereits vorhandenen Domänenwissens weiterlernt und feinetunt wird. Der Erfolg des Use Case hängt außerdem an der Verfügbarkeit von souveränen Angeboten. Viele Unternehmen wollen verständlicherweise die absolute Hoheit über vertrauliche Daten behalten. Im Zweifelsfall muss das KI-Sprachmodell aber auch aus dem eigenen Datacenter heraus betrieben werden können.

# 4

## „STÄRKEN STÄRKEN“: WIE DER AUFBAU EINES UNABHÄNGIGEN ÖKOSYSTEMS GELINGEN KANN



Die bisherige Experimentierphase hat gezeigt, dass es nicht ausreicht, allein auf das Potenzial der Technologie zu setzen. Darauf zu vertrauen, dass sich Produktivitätssteigerungen automatisch einstellen, ist ein schlechter Ratgeber. Erfahrungsgemäß lassen sich die besten Ergebnisse immer dann erzielen, wenn die Auswahl und Anwendung des Sprachmodells von fach- und domänenspezifischen Überlegungen geleitet wird. Hier stehen die Technologiepartner, die Industrie und die Politik gleichermaßen in der Pflicht. Ihre Aufgabe ist es, genau diesen Dreiklang aus „Technology“, „Processes“ und „People“ zu ermöglichen – also dass die am besten geeignete Technologie unter Einbeziehung der Fachkräfte vor Ort perfekt an die jeweiligen Unternehmensprozesse angepasst wird. Dies kann nur innerhalb eines funktionierenden Ökosystems aus starken Technologiefirmen, innovationsstarken Industrieunternehmen und klugen Köpfen zum Erfolg führen.

Um solch ein Ökosystem auf den Weg zu bringen, müssen drei Grundbedingungen erfüllt sein: Industrieunternehmen müssen das Domänenwissen ihrer Mitarbeiter für die Anwendung KI-basierter Technologien nutzen. Für den Betrieb dieser Technologien bedarf es, zweitens, einer durch Wettbewerb, Vielfalt und Interoperabilität geprägten Dateninfrastruktur. Drittens erfordern Investitionen und Innovationen einen zeitgemäßen, attraktiven und wettbewerbsfähigen Kapitalmarkt.

## 1. Das Domänenwissen der europäischen Industrie nutzen

Generative KI kann nur für höhere Produktivität und damit mehr Wohlstand sorgen, wenn die neue Technologie von breiten Teilen der Wirtschaft eingesetzt wird. Entscheidend für die erfolgreiche Implementierung künstlicher Intelligenz ist nicht allein die Technologie, sondern ihre Integration in Unternehmensprozesse auf Basis des Domänenwissens der Mitarbeiter.

Dazu ist es einerseits wichtig, dass das Unternehmen und auch die Mitarbeiter sich der Bedeutung ihres Wissens bewusst sind. Das ist keine Selbstverständlichkeit, schließlich liegen die Kenntnisse vor allem sehr erfahrener Fachleute meist in Form impliziten Wissens vor. Dieses „stille Wissen“ ist nicht formalisiert und stark prozess- und handlungsgebunden. Jede erfolgreiche Nutzung KI-basierter Technologien ist auf dieses Wissen angewiesen. Deshalb müssen die Mitarbeiter in die Entwicklung der konkreten Anwendungsfälle generativer KI eingebunden werden.

Andererseits müssen die Mitarbeiter über ein ausreichendes Verständnis der eingesetzten KI-Technologien verfügen. Zum einen, damit sie aktiv bei der Entwicklung des konkreten Anwendungsfalls mitarbeiten können. Und zum anderen, um reibungslose Abläufe an der Schnittstelle zwischen Mensch und KI zu gewährleisten. Das heißt, dass sehr umfangreiches kontextspezifisches

Fachwissen kein Ersatz für Kenntnisse digitaler Technologien ist. Im Gegenteil, vor allem Mitarbeiter mit umfassendem Domänenwissen benötigen auch digitale Fähigkeiten, um den erfolgreichen Einsatz generativer KI zu ermöglichen. Deshalb sind Schulungen und Weiterbildungen zu Themen digitaler Technologie sowie generativer KI unerlässlich.

Diese Verknüpfung fachspezifischen Wissens mit digitalen Fähigkeiten und Technologien kann nur gelingen, wenn sie in die richtige Unternehmenskultur und die passende Organisationsstruktur eingebettet ist. Mit Blick auf die Unternehmenskultur muss insbesondere das lebenslange Lernen einen hohen Stellenwert einnehmen. Weiterbildung muss als Chance, nicht als Belastung kommuniziert und verstanden werden. Überdies müssen Fehler als Anreiz zu Verbesserungen gesehen werden. Die Organisationsstruktur muss sicherstellen, dass die Träger des Fachwissens nahtlos mit den KI-Experten zusammenarbeiten. Dazu bedarf es klarer Verantwortlichkeiten und Anreize.

## **2. Keine einseitigen Abhängigkeiten bei der Infrastruktur**

Europas zögerliches Agieren beim Ausbau der europäischen Cloud- und Rechenzentren rächt sich in Zeiten des Gen-KI-Booms doppelt. Sowohl die Hersteller als auch die Anwender von Sprachmodellen benötigen Zugang zu sehr viel Rechenleistung. Oft ist daher das Angebot der drei dominanten amerikanischen Cloud-Dienste attraktiv, da sie sowohl die Infrastruktur als auch die KI-Anwendungen innerhalb dieser Infrastruktur und der angebotenen Dienstleistungen anbieten können. Diese spielen ihre Marktmacht zunehmend aus – auch weil sie so auf absehbare Zeit ebenbürtigen europäischen Konkurrenten den Wettbewerb erschweren.

Die Politik steht deshalb vor einem schwierigen Balanceakt: Wenn sie den Marktzugang für US-amerikanische Cloud-Anbieter erschwert, drohen den europäischen Unternehmen Nachteile in Form geringerer Produktivität. Denn bis ein gleichwertiges europäisches Cloud-Ökosystem aufgebaut wäre, würden Jahre vergehen. Gleichzeitig muss dem wettbewerbshemmenden Verhalten einzelner Marktteilnehmer stärker der Riegel vorgeschoben werden.

Sogenannte „Lock-in-Strategien“ sind beispielsweise seit Jahren bekannt, wurden aber noch immer nicht zufriedenstellend unterbunden. Auch der Digital Markets

Act der EU dürfte hier nur ein erster Schritt sein, zumal er konsequent durchgesetzt werden muss. Einigen Cloud-Dienstleistern gelingt es nach wie vor, durch die Vertragsgestaltung, technische Barrieren oder hohe Transfergebühren den Wechsel zu einem anderen Anbieter zu verhindern. Die Politik und die Kartellbehörden müssen hier stärker intervenieren. Sie müssen verhindern, dass Cloud-Dienste für den Datenabfluss überbeuerte Gebühren fordern, der Datenzufluss aber kostenfrei angeboten wird. Auch Rabatttaktiken, die darauf abzielen, Unternehmen über längere Zeit an einen einzigen Anbieter zu binden, sollten eingeschränkt sowie strenge Interoperabilitätsvorschriften zur Pflicht gemacht werden.

Es sei zudem darauf verwiesen, dass leistungsfähige europäische Cloud-Firmen durchaus am Markt vertreten sind. Ein weiteres Wachstum ist jedoch unabdingbar, um im Wettbewerb bestehen zu können. Die Politik könnte das Wachstum europäischer Cloud-Firmen fördern, z.B. indem die öffentliche Hand deren Dienste stärker in Anspruch nimmt. Eine solche nachfrageorientierte Stärkung von souveränen europäischen Angeboten ist ein zentraler Hebel, der neben anderen Instrumenten wie der öffentlichen Förderung von GPU-Clustern viel stärker als bisher genutzt werden sollte.

### 3. Die Kapitalmarktlücke schließen

Das Wettrüsten im Gen-KI-Bereich hat sich massiv beschleunigt. Für europäische Firmen in der Wachstumsphase bedeutet dies, dass sie über eine entsprechende Kapitalausstattung verfügen müssen, um im Wettbewerb bestehen zu können. Schon vor dem aktuellen KI-Boom war es für europäische Firmen deutlich schwieriger als für amerikanische, das benötigte Kapital aufzutreiben.

Ohne einen Zuwachs an privatem Kapital kann die digitale Transformation der EU schlichtweg nicht finanziert werden. Derzeit ist allerdings ein grenzüberschreitender EU-Kapitalmarkt im Sinne eines grenzüberschreitenden Marktes für börsliches und außerbörsliches Kapital trotz der Fortschritte in jüngster Zeit noch nicht in Sicht. Zu groß sind die Hürden, etwa beim Insolvenz- und Börsenrecht, als dass ihre Harmonisierung derzeit Aussicht auf Erfolg hätte. Die Politik sollte sich davon nicht entmutigen lassen und den eingeschlagenen Weg der kleinen Verbesserungen kontinuierlich fortsetzen.

Ein erster Schritt wäre die Stärkung der institutionellen Anleger. Über einen Ausbau der ergänzenden kapitalgedeckten Altersvorsorge könnten Fonds zum Beispiel mehr Kapital einsammeln und im Gegenzug dazu verpflichtet werden, einen vertretbaren Anteil in den Venture-Bereich zu investieren. Auch eine zeitgemäße Anpassung der Anlageverordnung hätte ihre Vorteile.

Ziel sollte es sein, dass europäische Pensionskassen nach amerikanischem Vorbild und in demselben Umfang in risikoreichere Anlageformen investieren.

Trotz dieser enormen Herausforderungen sollte eine gute Nachricht nicht vergessen werden: Aufgrund des geringeren Wettbewerbs unter Investoren in Europa und der niedrigeren Kapitalzuflüsse lagen die Renditen im europäischen VC-Sektor zuletzt über denen in den USA. Auch private Anleger würden also von Investments in diesem Bereich profitieren. Dafür müssen aber zwei Bedingungen erfüllt sein: Erstens sollte die „financial literacy“ durch Bildungsmaßnahmen auf nationaler Ebene quer durch alle Altersgruppen hinweg verbessert und für eine kluge Portfolio-Diversifizierungsstrategie geworben werden. Zweitens sollten breitere, paneuropäische Finanzprodukte aufgelegt und insbesondere auch privaten Anlegern zugänglich gemacht werden. Die Vision dahinter: ein Ökosystem für europäische Investitionen, das seinen Namen auch verdient.

**Paneuropäische Finanzprodukte sind ein wichtiger Schritt hin zu einem Ökosystem für europäische Investitionen, das seinen Namen auch verdient.**

## **Herausgeber**

**Innovate Europe  
Foundation (IE.F)**  
Uhlandstraße 175  
10719 Berlin  
[www.ie.foundation](http://www.ie.foundation)

## **Autor**

**IE.F**  
**Clark Parsons**  
[c.parsons@ie.foundation](mailto:c.parsons@ie.foundation)

## **Kontakt**

**IE.F**  
**Clark Parsons**  
Geschäftsführer  
Innovate Europe Foundation (IE.F)  
[c.parsons@ie.foundation](mailto:c.parsons@ie.foundation)  
Tel.: +49 30 8877 429-400

## **Veröffentlichung**

Juli 2024

## **Haftungsausschluss**

Diese Studie dient ausschließlich der generellen Orientierung. Leser sollten Aktivitäten nicht ausschließlich auf Basis der Inhalte dieser Studie anstoßen, insbesondere nicht ohne vorherige professionelle und individuelle Beratung. Die IE.F ist nicht haftbar für Schäden, die aus Handlungen auf Basis dieser Studie entstehen.

