

# Ist Deepseek besser als Chat-GPT? Überholt China die USA im KI-Wettlauf? Antworten auf die wichtigsten Fragen zum Hype

Mit einem neuen Chatbot gelingt der kleinen Forschungsgruppe Deepseek aus China ein unerwarteter Grosserfolg. Alles, was man darüber wissen sollte, um mitreden zu können.

Gioia da Silva, Philipp Wolf, Leonid Leiva Ariosa, Katrin Büchenbacher 27.01.2025, 16.54 Uhr

Illustration Simon Tanner / NZ

## Deepseek R1, was ist das?

Deepseek R1 ist die chinesische Antwort auf Chat-GPT o1: ein hochmoderner KI-Chatbot. Das Modell simuliert anders als typische Sprachmodelle eine Kette von Gedankengängen nach menschlichem Vorbild. Deepseek kann sowohl im Browser als auch als App genutzt werden. Am Montag war die App in den USA auf Platz 1 der Download-Charts von Apple, vor Chat-GPT. In der Schweiz und Deutschland war sie am Montagabend auf Platz 1 beziehungsweise 3. Mit dem überraschenden Erfolg hat Deepseek es geschafft, Schockwellen durch das Silicon Valley zu schicken.

Denn bezüglich der Leistungen und der Geschwindigkeit kann es R1 mit Chat-GPT o1 aufnehmen. Deepseek R1 kann beispielsweise Matheaufgaben lösen, Daten analysieren, Code programmieren, bei der Hochzeitsplanung helfen, Internetrecherchen durchführen.

Gegenüber Chat-GPT gibt es jedoch drei entscheidende Unterschiede: Deepseek ist nahezu kostenlos, hat seinen Software-Quellcode offengelegt und unterliegt dem chinesischen Recht und somit der Zensur und der Staatspropaganda.

## Funktioniert Deepseek ohne Anmeldung? Wie kann ich das Tool ausprobieren?

Unter [deepseekv3.com/chat](https://deepseekv3.com/chat) kann man den Chatbot laut Deepseek ohne Anmeldung ausprobieren. Im Test der NZZ funktioniert dies allerdings nicht. Alternativ kann man das Programm herunterladen und bei sich auf dem Gerät laufen lassen. Einfacher ist es, das Tool nach der Registration in der Web-Version zu benutzen. Anmelden kann man sich unter [chat.deepseek.com/](https://chat.deepseek.com/) oder auf der Deepseek-App.

Sowohl für die Version des Chatbots vor der Registrationsschranke als auch für jene dahinter muss man die [Nutzungsbedingungen](#) akzeptieren. Damit gibt man dem Unternehmen weitgehende Befugnisse für den Umgang mit persönlichen Daten wie Interessen oder dem genauen Standort des Geräts, die auch mit Tracking-Cookies von anderen Websites gesammelt werden. Weiter erhebt die Firma sogar die Geschwindigkeit und das Muster, mit denen eine Person auf der Tastatur tippt.

Laut den Nutzungsbedingungen dürfen explizit personenbezogene Daten gesammelt und an Drittparteien weitergegeben werden. Zudem kann bei chinesischen Firmen nie ausgeschlossen werden, dass sie von der Regierung zur Datenherausgabe gezwungen werden.

Umso attraktiver dürfte es sein, das Programm auf den eigenen Rechner herunterzuladen und selbst zu betreiben. Dann sollten keine Daten abfließen. «Gerade analysieren Hunderte Softwareentwickler den Open-Source-Code, um zu sehen, ob das tatsächlich so ist», sagt Peter Fintl, China-Experte und Leiter Technologie und Innovation beim Beratungsunternehmen Capgemini. Bisher wurde nicht über offensichtliche Sicherheitslücken oder Backdoors berichtet.

## **Deepseek auf Deutsch: Welche Sprachen kann das Tool?**

Deepseek passt seine Antwort automatisch der Sprache der Anfrage an. Bei der Anmeldung sieht es zwar so aus, als wäre das Tool auf die englische Sprache eingestellt. Doch alles, was man tun muss, um eine deutsche Antwort zu erhalten, ist, eine Frage auf Deutsch einzutippen.

In der Onlineversion kann man zudem explizit zwischen zehn Sprachen auswählen, darunter Deutsch, Englisch, Französisch, Italienisch und Spanisch.

## **Gibt es eine Deepseek-Aktie? Wie kann ich in die Firma dahinter investieren?**

Deepseek ist ein KI-Forschungslabor, kein börsenkotiertes Unternehmen. Gegenüber dem chinesischen Tech-Portal 36Kr sagte der Gründer Liang, er habe Deepseek aus rein wissenschaftlichem Interesse gegründet. Wirtschaftlich ergebe es keinen Sinn, so etwas zu gründen, die Kapitalrendite sei bei Grundlagenforschung sehr gering.

## **Wer ist der Gründer von Deepseek?**

Die Macher hinter Deepseek werden in China als Nationalhelden gefeiert: Sie haben es geschafft, Chat-GPT zumindest zeitweise vom Thron zu stossen – und das, obwohl die USA mit Exportkontrollen sicherstellen, dass China keine mächtigen Hochleistungschips erhält. Kurz nach dem Launch von Deepseek R1 wurde dessen Gründer von Chinas Ministerpräsidenten Li Qiang zu einem Austausch eingeladen.

Der Gründer des Forschungslabors, das ist Liang Wenfeng, ein 40-jähriger IT-Ingenieur, der als Hedge-Fund-Manager Milliardär wurde. Deepseek hat er als Nebenprojekt lanciert. 2021 begann er, mit dem Vermögen, das er sich aufgebaut hatte, Tausende Chips von der amerikanischen Firma Nvidia zu kaufen, wie die [«Financial Times» einen Geschäftspartner von Liang zitiert](#). Zwei Jahre später gründete Liang Deepseek.

In einem seiner seltenen [Interviews](#) sagte er im Juli vergangenen Jahres der chinesischen Investment-Onlinezeitschrift «Waves», sein Ziel sei es, künstliche allgemeine Intelligenz zu entwickeln, das heisst ein autonomes KI-System, das menschliche Intelligenz erreicht. Liang sagte: «Unser Ausgangspunkt ist nicht, schnell Geld zu verdienen, sondern an die technologische Spitze zu gelangen und die Entwicklung des gesamten Ökosystems voranzutreiben.»

## **Ist Deepseek R1 besser als 01 von Open AI?**

Deepseek hat gleichzeitig mit der Veröffentlichung des R1-Modells eine [wissenschaftliche Studie](#) publiziert. Diese geht detailliert auf den Prozess ein, wie R1 trainiert wurde und wie das Modell in verschiedenen standardisierten Tests abschneidet.

Gemäss der Studie schneidet R1 beim Lösen mathematischer Aufgaben oder bei Programmiertests ähnlich oder leicht besser ab als das beste Modell von Open AI, o1. Die Aussagekraft dieser standardisierten Tests ist aber umstritten. Die Ergebnisse müssen deshalb noch von unabhängigen Experten überprüft werden.

Auf Programmierforen kursieren auch Berichte, dass R1 bessere «Gedankengänge» produziere, wenn es Antworten gebe. O1 verwendet die gleiche «Chain of thought»-Technik, um die einzelnen Schritte beim Lösen einer Aufgabe aufzuführen und sich notfalls selbst zu korrigieren. Auf diesen Ansatz hofft Open AI den Erfolg von o1 und vom Nachfolgemodell o3 zu gründen. Aber offenbar kann Deepseek das auch ziemlich gut.

Andere Nutzer zeigen sich beeindruckt von der Qualität der Ergebnisse, wenn sie R1 zum Schreiben kreativer Texte auffordern. Andere bemerken, dass R1 weniger dazu neige, «Fakten» frei zu erfinden – das von anderen Chatbots bekannte «Halluzinieren» trete also seltener auf.

Aber die grössten Vorteile von R1 liegen in der Effizienz, mit der Deepseek das Modell trainiert hat. Das gibt KI-Forschern und Unternehmen mit weniger Ressourcen als Open AI und Co. Hoffnung.

Deepseek hat diese Effizienz den eigenen Angaben nach mithilfe ausgeklügelter Tricks erreicht. Sie betreffen einerseits die Algorithmen, mit denen R1 trainiert wurde. Diese rezyklieren auf innovative Weise die Techniken des «verstärkenden Lernens», mit denen eine KI selbständig lernt. Ähnliche Techniken setzte vor vielen Jahren zum Beispiel Google Deep Mind ein, um Alpha Go zu entwickeln – die KI, die im fernöstlichen Spiel Go den besten menschlichen Spielern überlegen war.

Deepseek verwendete auch eine niedrigere Präzision bei einzelnen Rechenoperationen und konnte so den Rechenaufwand senken – ein Ansatz, der in der Vergangenheit dazu beigetragen hat, die automatische Bilderkennung effizienter zu machen.

Andererseits hat Deepseek aber auch die Auslastung seiner Rechner optimiert und dadurch mehr aus seinen beschränkten Ressourcen herausgeholt. Diese Innovationen sind jetzt für jedermann verfügbar und dürften in Zukunft anderen KI-Entwicklern zugutekommen. In Sachen Reputation ist das ein grosser Gewinn für Deepseek.

## **Wie offen sind die Modelle von Deepseek tatsächlich?**

Deepseek hat das R1-Modell mit einer Lizenz veröffentlicht, die es Nutzern erlaubt, das Modell gratis herunterzuladen und beliebig für ihre eigenen Zwecke anzupassen. Die Resonanz auf diese Entscheidung spricht für sich: R1 ist auf der Open-Source-KI-Plattform Hugging Face derzeit das meistheruntergeladene KI-Modell mit fast 150 000 Downloads in nur sieben Tagen seit der Veröffentlichung.

R1 steht jedoch nicht uneingeschränkt offen. Denn Deepseek hat die Datensätze nicht offengelegt, mit denen das Modell trainiert wurde. Trotzdem verblüfft die Transparenz des Tools und lässt Unternehmen wie Open AI, die ihre Modelle geheim halten und von Nutzern zum Teil saftige Abogebühren kassieren, ziemlich alt aussehen.

Deepseek dachte auch an die Bedürfnisse verschiedener Nutzergruppen und erleichterte ihnen den Zugang zum neuen Modell. Das R1-Basismodell verfügt über 671 Milliarden Parameter – ein Mass für die Komplexität und Leistungsfähigkeit des Modells. Aufgrund dieser Grösse erfordert es viel

Speicherplatz und sehr schnelle Rechner. Dieses Topmodell werden also nur Unternehmen oder Hochschulen mit genügend Ressourcen und Rechenleistung ausserhalb der chinesischen Cloud benutzen können.

Deepseek veröffentlichte neben dem grossen Basismodell sechs kleinere Varianten, sogenannte «Distills», die Nutzer mit beschränkter Rechenleistung auf ihren eigenen Rechnern laufen lassen können. Die kleineren von diesen Varianten laufen sogar auf einem gewöhnlichen Laptop. Diese, wie auch das Basismodell, können ohne Verbindung zum Internet benutzt werden. Die Daten bleiben also bei einer Nutzung auf dem eigenen Computer gesichert.

## **Was bedeutet die neue Konkurrenz für Open AI?**

Für Open AI bedeutet die Veröffentlichung von Deepseeks Modellen, V3 im Dezember und R1 im Januar, einen ernsthaften Wettbewerb. Die Leistung der Modelle von Deepseek ist jenen von Open AI ebenbürtig. Deepseek berechnet seinen Nutzern in der Cloud aber nur fünf Prozent vom Chat-GPT-Preis. Wer das Modell herunterlädt und auf eigenen Computern laufen lässt, nutzt die chinesische KI kostenlos.

Open AI steht somit unter Zugzwang. Die chinesischen Modelle zeigen, dass der Vorsprung von Open AI langfristig nicht sicher ist. Die Kapitalgeber dürften sich fragen, wieso sie weitere Milliarden in die amerikanische Firma pumpen sollten.

Nutzer werden sich ihrerseits überlegen, ob sie bereit sind, mehr für Chat-GPT zu bezahlen. Sie dürften aber vor den Datenschutzbestimmungen der chinesischen Cloud zurückschrecken. Der Datenschutz spricht derzeit noch für Open AI. Aber Open AI muss mehr als Datenschutz liefern, wenn es Investoren und Nutzer langfristig binden will.

Open AI reagierte schnell auf die Veröffentlichung von R1. Am Donnerstag kündigte der CEO Sam Altman auf X an, dass o3 mini auf Chat-GPT frei verfügbar sein werde. Die kleine Version des neuesten Modells o3 steht allerdings nur in begrenztem Umfang zur Verfügung. Zahlende Kunden erhielten täglich hundert o3-mini-Abfragen, sagte Altman in einem späteren Post. o3 mini ist für relativ einfache Aufgaben konzipiert. Dafür verbraucht es weniger Ressourcen. Das belastet die Open-AI-Server weniger und reduziert die Wahrscheinlichkeit, dass sich Nutzer über langsame Antworten oder sogar über den Absturz der App ärgern.

## **Erfahrungen mit Deepseek R1: Wie setzt das Tool die chinesische Zensur um?**

Deepseek funktioniert ähnlich wie Chat-GPT. Man tippt einen Prompt ein und erhält innerhalb weniger Sekunden eine Antwort. Im Test der NZZ funktionierte das Tool für unpolitische Fragen einwandfrei. Es generierte eine Programmiervorlage für einen Social-Media-Bot, Struktur und Text für eine Präsentation über Schweizer Ostereier, Empfehlungen für Beziehungskrisen.

Fragen nach Fakten, die in China von der Zensur betroffen sind, zeigen eine Auffälligkeit: Erst beginnt das Tool sichtbar eine Antwort zu verfassen, doch wenige Sekunden später löscht Deepseek die Antwort und schreibt einen Einzeiler: «Sorry, das übersteigt meinen Anwendungsbereich. Lass uns über etwas anderes sprechen.»

Diese Antwort kommt bei politischen Fragen zu Taiwan, bei Fragen zur uigurischen Minderheit in Xinjiang, zu Xi Jinping, über das Tiananmen-Massaker oder auch zu den Covid-Protesten in China im Jahr 2022.

## **Chinesische Internetzensur live: Deepseek im Experiment der NZZ, wenn man fragt, was am 4. Juni 1989 in Peking geschehen sei. Antwort: das Tiananmen-Massaker.**

Deepseek

Dass die Antwort erst formuliert, dann aber gelöscht wird, zeigt, dass die Zensur erst nachgelagert greift. Vermutlich prüft ein Zensurprogramm sämtliche Antworten von Deepseek auf Aussagen, die in China als heikel eingestuft werden. Der Tech-Blog «Ars Technica» [berichtet](#), dass die Zensur ausschliesslich in der Onlineversion des Tools auftrete. Sie lasse sich umgehen, indem man das Programm herunterlade und selbst betreibe.

## **Überholt China die USA nun im KI-Wettlauf? Und was heisst das für Europa?**

Der globale Erfolg von Deepseek ist ein Weckruf, nicht nur für die Branche, sondern auch für die Politik. Die USA galten als führend in der KI, und China hinkte hinterher. Deepseek fordert die amerikanische Dominanz heraus. Kai-Fu Lee, der ehemalige Chef von Google in China, stellt in seinem Buch «AI Superpowers. China, Silicon Valley and the New Global Order» die These auf, dass die Durchbrüche in der KI-Forschung aus den USA kämen, aber bei der Umsetzung sei China besser und schneller.

Deepseek gilt als ernstzunehmender Konkurrent, weil er es geschafft hat, ohne die neusten Hochleistungschips ein mit Chat-GPT vergleichbares Modell zu bauen. Ausserdem liegt der Fokus noch auf der Forschung und Entwicklung statt auf der Kommerzialisierung. Das stellt die Geschäftsmodelle der etablierten Player wie Open AI und Anthropic infrage. Der Open-Source-Ansatz von Deepseek lädt Entwickler und Firmen auf der ganzen Welt dazu ein, noch bessere KI zu entwickeln. Es könnte also sein, dass der nächste Erfolg aus den USA, aus China oder gar aus Indien oder Europa kommt.

«Für Europa ist der Erfolg des KI-Chatbots eine gute Botschaft», sagt der Technologieexperte Fintl und rät davon ab, den Chatbot kategorisch abzulehnen, weil er aus China stamme. «Konkurrenz belebt das Geschäft. Das Rennen ist auch in Europa noch nicht entschieden.»

## **Ist Deepseek der Beweis dafür, dass die amerikanischen Hightech-Sanktionen nicht funktionieren?**

Der Deepseek-Gründer Liang hatte sich schon Tausende Nvidia-Chips gesichert, bevor die USA deren Export beschränkten. Als die Exportkontrollen ab Oktober 2022 in Kraft traten, war Deepseek gezwungen, Nvidia-Chips zu kaufen, die im Vergleich nur halb so leistungsfähig waren. Man könnte argumentieren, dass die Sanktionen ihr Ziel erfüllt hätten. China wurde vom Zugang zu Hochleistungschips abgeschnitten.

Gleichzeitig zeigt Deepseek, wie sich chinesische Forscher, Unternehmen und Institutionen an amerikanische Exportkontrollen angepasst haben. Zur Entwicklung seines Modells brauchte Deepseek nach eigenen Angaben lediglich 2000 solcher Nvidia-Chips mit gedrosselter Rechenleistung. Deepseek bestätigt den Standpunkt zahlreicher Experten: Die amerikanischen Sanktionen führten lediglich dazu, dass chinesische Unternehmen innovativer und effizienter würden.