NEWS

Grok 4.7 veröffentlicht: Benchmarks, Preise und erste Reaktionen

xAI hat Grok 4.7 freigegeben. Der Preis bleibt bei 2 und 6 Dollar je Million Token, die Benchmarks fallen je nach Analyst sehr verschieden aus. Was belegt ist, was xAI behauptet und was die ersten Nutzer berichten.

Stand: 21.09.2026 · 5 Min. Lesezeit

xAI hat Grok 4.7 veröffentlicht. Das Modell ist seit dem 21. September 2026 über die API des Anbieters erreichbar, dazu in Cursor und im hauseigenen Werkzeug Grok Build. Elon Musk hatte den Start Anfang September für den 11. oder 12. des Monats angekündigt. Gehalten hat der Termin nicht, und das ist bei dieser Modellreihe inzwischen die Regel.

Am Preis hat sich nichts geändert. Grok 4.7 kostet weiterhin 2 Dollar je Million Eingabe-Token und 6 Dollar je Million Ausgabe-Token. Damit liegt xAI dort, wo sonst chinesische Anbieter liegen, und deutlich unter den westlichen Spitzenmodellen. Ob das ein Schnäppchen ist, hängt allerdings davon ab, wie viele Token das Modell für dieselbe Aufgabe verbraucht. Genau an diesem Punkt setzt die erste Kritik an.

Was xAI verspricht

Nach Angaben des Herstellers steckt hinter Grok 4.7 ein neues, größeres Basismodell. Trainiert worden ist es mit einem längeren Reinforcement-Learning-Durchlauf auf einer schwierigeren Aufgabenmischung, ausgerichtet auf Probleme, an denen ein Modell mehrere Stunden arbeitet. Dazu kommt eine verbesserte Selbstprüfung und ein natives Verständnis für die Schnittstelle des Grok-Bots.

Zur Größe des Modells schweigt xAI. Musk hat Anfang September von 2,1 Billionen Parametern gesprochen, rund 40 Prozent mehr als beim Vorgänger. Bestätigt ist diese Zahl nicht, ein Model Card mit Architekturangaben gibt es nicht.

Belegt sind dagegen die Eckdaten aus der Entwicklerdokumentation. Das Kontextfenster umfasst 500.000 Token und ist damit genauso groß wie bei Grok 4.6. Der Wissensstand reicht bis Mai 2026, die Kennung in der Schnittstelle lautet `grok-4.7`.

Preise und Verfügbarkeit

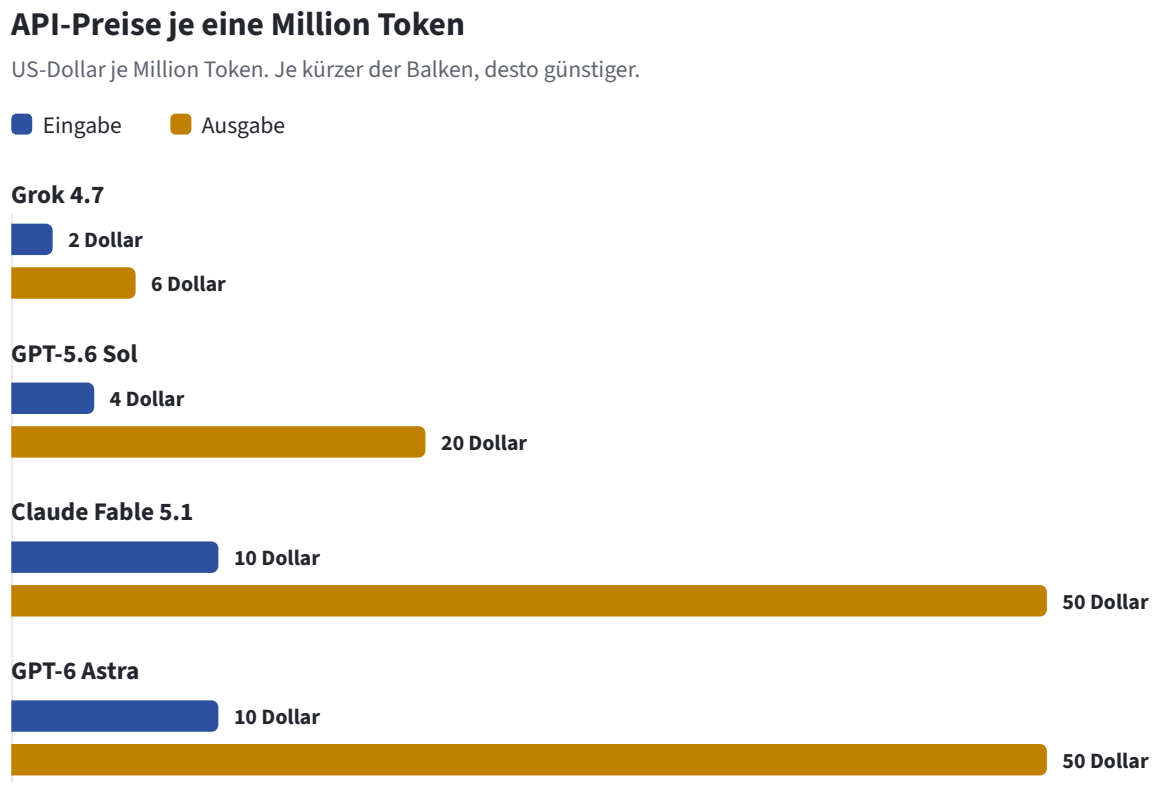
Der Einstiegspreis gilt nicht für jede Anfrage. Ab 200.000 Token im Prompt verdoppelt xAI den Eingabepreis auf 4 Dollar und den Ausgabepreis auf 12 Dollar, und zwar für die gesamte Anfrage, nicht nur für den den Teil,

AI NEWS pro - KI-News ohne Hype! · © 2026 SOFTCLICK Labs · Stand des Beitrags: 21.09.2026

der die Schwelle übersteigt. Zwischengespeicherte Eingaben kosten 50 Cent je Million, oberhalb der Schwelle einen Dollar.

> **Zwei Dollar sind der Mindestpreis, nicht der Normalfall**
>
> Wer mit 500.000 Token Kontext arbeitet, zahlt nie den beworbenen Satz. Eine einzige Anfrage über der Schwelle wird komplett zum doppelten Tarif abgerechnet. Bei Agenten, die den Kontext über viele Schritte anwachsen lassen, ist das der Normalfall und nicht die Ausnahme.

Daneben steht eine schnellere Variante, die auf besserer Infrastruktur läuft und das Doppelte kostet. Sie ist nur über Cursor und Grok Build zu haben, nicht über die offene Schnittstelle.



Quelle: Anbieter-Preisseiten, Stand: 21. September 2026 · Grafik: AI NEWS pro

Modell	Eingabe	Ausgabe
Grok 4.7	2 Dollar	6 Dollar
GPT-5.6 Sol	4 Dollar	20 Dollar

AI NEWS pro - KI-News ohne Hype! · © 2026 SOFTCLICK Labs · Stand des Beitrags: 21.09.2026

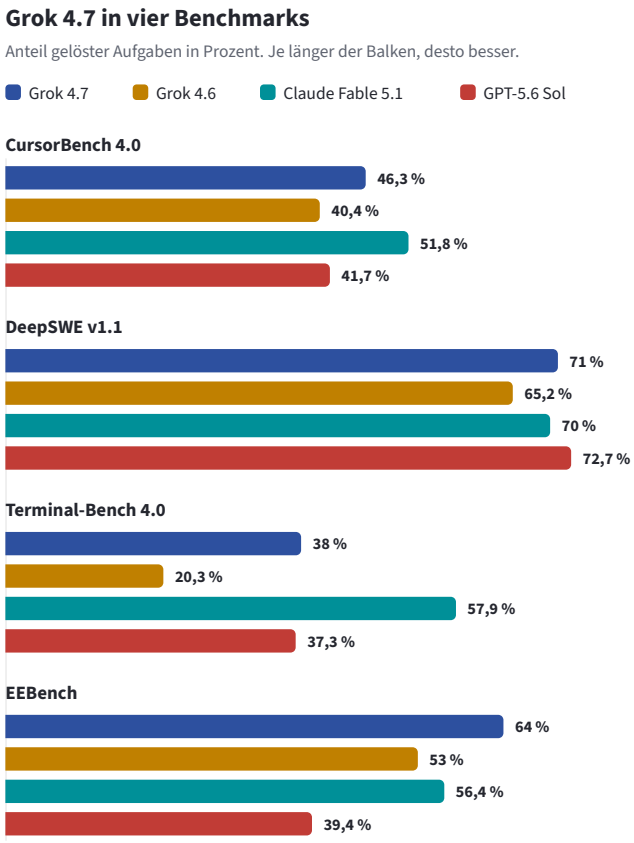
Modell	Eingabe	Ausgabe
Claude Fable 5.1	10 Dollar	50 Dollar
GPT-6 Astra	10 Dollar	50 Dollar

Listenpreise je eine Million Token, Stand 21. September 2026. Bei Grok 4.7 gilt der Satz bis 200.000 Prompt-Token, bei GPT-6 Astra bis 272.000.

Seit dem Starttag läuft das Modell auch in GitHub Copilot, für die Tarife Pro, Pro+, Max, Business und Enterprise. Die Freischaltung erfolgt schrittweise, Administratoren können sie über die Modellrichtlinien unterbinden.

Die Benchmarks im Detail

xAI hat zum Start eigene Messwerte veröffentlicht. Bemerkenswert daran ist weniger, wo Grok 4.7 vorn liegt, sondern wie oft es das nicht tut.



Quelle: xAI (Herstellerangaben), Stand: 21. September 2026 · Grafik: AI NEWS pro

AI NEWS pro - KI-News ohne Hype! · © 2026 SOFTCLICK Labs · Stand des Beitrags: 21.09.2026

Benchmark	Grok 4.7	Grok 4.6	Claude Fable 5.1	GPT-5.6 Sol
CursorBench 4.0	46,3 %	40,4 %	51,8 %	41,7 %
DeepSWE v1.1	71,0 %	65,2 %	70,0 %	72,7 %
Terminal-Bench 4.0	38,0 %	20,3 %	57,9 %	37,3 %
EEBench	64,0 %	53,0 %	56,4 %	39,4 %
HealthBench Professional	56,7 %	48,5 %	62,1 %	60,5 %
Harvey Legal Agent	19,6 %	15,8 %	6,7 %	2,5 %

Alle Werte laut xAI, Stand 21. September 2026. Den Aufwandsmodus weist der Hersteller nur bei DeepSWE aus, dort lief Grok 4.7 auf "high effort".

Gegenüber dem eigenen Vorgänger legt das Modell überall zu, beim Terminal-Bench sogar um 17,7 Punkte. Gegenüber Claude Fable 5.1 geht der Vergleich drei zu drei aus, und die Verteilung ist aufschlussreich.

Das Anthropic-Modell führt bei den Coding- und Agentenaufgaben, beim Terminal-Bench mit fast 20 Punkten Abstand. Grok 4.7 führt dort, wo Fachwissen abgefragt wird: bei EEBench, einem Test aus der Elektrotechnik, und beim juristischen Agenten-Benchmark von Harvey.

Musk hatte angekündigt, Trainingsdaten aus der Raketen- und Satellitenentwicklung von SpaceX zu nutzen. Der Vorsprung bei EEBench passt dazu, ein klarer Beweis für einen Zusammenhang ist er hingegen nicht.

Wenn zwei Analysten verschieden messen

Die interessanteste Zahl des Tages steht nicht in der Ankündigung. Der Benchmark-Dienst Artificial Analysis hat Grok 4.7 im Modus "xhigh" durch seinen Intelligence Index geschickt und kommt auf 46 Punkte. Claude Fable 5.1 und GPT-6 liegen dort bei je 53.

Deutlicher wird der Unterschied beim Terminal-Bench 4.0. xAI misst 38,0 Prozent, Artificial Analysis misst 26 Prozent. Das sind zwölf Punkte Abstand für dasselbe Modell im selben Benchmark. Unabhängig gemessen landet Grok 4.7 damit auf dem Niveau von DeepSeek V4.1 Flash, das auf 27 Prozent kommt, während GPT-6 Astra mit 60 und Claude Fable 5.1 mit 55 Prozent weit vorn liegen.

Auflösen lässt sich die Differenz von außen nicht. Beide Unternehmen veröffentlichen ihre Setups nicht vollständig, und ein Benchmark-Ergebnis hängt an Details wie Werkzeugumgebung, Zeitlimit und Aufwandsstufe. Wer die Zahlen zur Entscheidung braucht, misst am besten am eigenen Anwendungsfall.

AI NEWS pro - KI-News ohne Hype! · © 2026 SOFTCLICK Labs · Stand des Beitrags: 21.09.2026

Erste Reaktionen

Im Forum von Cursor haben die ersten Nutzer noch am Starttag ihre Eindrücke gesammelt, und die Tendenz ist verhalten. Ein Entwickler nennt das Modell schlicht ein "sidegrade" und rechnet vor, dass es fast doppelt so viele Token verbraucht wie Grok 4.6. Ein anderer beschreibt es als "super spendy" gemessen an dem, was dabei herauskommt.

Ein dritter Beitrag zielt auf die Vergleichsgrundlage. Danach ist Grok 4.7 im Aufwandsmodus "xhigh" gegen Grok 4.6 im Modus "high" angetreten. Beim Preis je Token stimmt der Vergleich dann, beim Preis je erledigter Aufgabe nicht mehr.

Genau das ist die offene Frage hinter dem Preisschild. Ein Modell, das für dieselbe Aufgabe doppelt so viele Token verbraucht, kostet am Ende dasselbe wie eines zum doppelten Tarif. Zahlen zur Token-Effizienz nennt xAI diesmal nicht, anders als beim Start von Grok 4.5 im Juli.

Quellen

1. xAI: Introducing Grok 4.7 · abgerufen 21.09.2026 · https://x.ai/news/grok-4-7
2. BigGo Finance: Musk Announces Grok 4.7 Launch in Ten Days, Touts 2.1 Trillion Parameters to Beat All Models · abgerufen 21.09.2026 · https://finance.biggo.com/news/cdeb763e-3e82-4f0b-82bd-4f473881bf08
3. xAI: Grok Models & Pricing · abgerufen 21.09.2026 · https://docs.x.ai/developers/models
4. xAI: API Pricing · abgerufen 21.09.2026 · https://docs.x.ai/developers/pricing
5. GitHub: Grok 4.7 is now available in GitHub Copilot · abgerufen 21.09.2026 · https://github.blog/changelog/2026-09-21-grok-4-7-is-now-available-in-github-copilot
6. Artificial Analysis: Grok 4.7 (xhigh) · abgerufen 21.09.2026 · https://artificialanalysis.ai/models/grok-4-7
7. The Decoder: xAI launches Grok 4.7 at bargain prices, but benchmarks reveal a wide gap to Claude and GPT-6 · abgerufen 21.09.2026 · https://the-decoder.com/xai-launches-grok-4-7-at-bargain-prices-but-benchmarks-reveal-a-wide-gap-to-claude-and-gpt-6/
8. Cursor-Forum: Share your Thoughts on Grok 4.7 · abgerufen 21.09.2026 · https://forum.cursor.com/t/share-your-thoughts-on-grok-4-7/172527

Themen: xAI · Grok · LLM · Benchmark · Release

AI NEWS pro - KI-News ohne Hype! · © 2026 SOFTCLICK Labs · Stand des Beitrags: 21.09.2026