

GPT 5.4 vs Luna Pro

Ich habe mich informiert – die Info ist inzwischen echt: Es gibt tatsächlich eine "Luna"-Variante als Teil der neuen **GPT-5.6**-Familie (Sol, Terra, Luna), erschienen ab **26. Juni 2026** mit "Luna Pro" als Reasoning-Variante Anfang Juli aufgetaucht. Hier der Vergleich zu **GPT-5.4**:

Kontext: Was ist was?

Modell	Rolle
GPT-5.4	Vorheriges Flaggschiff-Modell (Release: März 2026)
GPT-5.6 Luna / Luna Pro	Das schnellste, günstigste Modell der neuen 5.6-Familie (neben Sol = Top-Tier, Terra = Mid-Tier). "Pro" ist die Variante mit <code>reasoning.mode = pro</code> für höhere Qualität bei gleichem Modell

Direkter Vergleich: GPT-5.4 vs. GPT-5.6 Luna

Preis

- GPT-5.4: \$2,50 / \$15,00 (Input/Output pro 1M Tokens)
- GPT-5.6 Luna: \$1,00 / \$6,00 → **~2,5× günstiger**

Kontextfenster

- GPT-5.4: 1,0 Mio. Tokens Input
- GPT-5.6 Luna: 1,05 Mio. Tokens Input (marginal größer)

Benchmarks – gemischtes Bild:

Bereich	Gewinner
GPQA, MMMU-Pro, Toolathlon	GPT-5.4 (leicht besser)
BrowseComp, FrontierMath, GDPval-AA, SWE-Bench Pro	Luna (deutlich besser)
Graphwalks BFS >128k	Luna klar vorn (51,2 % vs. 21,4 %)

Bereich	Gewinner
Agent's Last Exam	Luna übertrifft leicht GPT-5.4-Niveau

Geschwindigkeit

- GPT-5.4 hat niedrigere Latenz (schnellere erste Antwort)
- Luna hat höheren Durchsatz (schnellere Generierung, mehr Zeichen/Sek.)

Wichtige Einordnung von OpenAI selbst

Laut OpenAIs eigener Positionierung liegt **Luna auf dem Niveau von GPT-5.5** bei mehreren Tests – trotz des Preises des günstigsten Modells der Familie. Es ist explizit **nicht** als "besser als alles bisherige" gedacht, sondern für schnelle, alltägliche Aufgaben (Zusammenfassungen, Entwürfe, Routineautomatisierung), während **Sol** und **Terra** die anspruchsvolleren/komplexeren Use Cases abdecken.

Fazit

- **Rohleistung (Benchmarks):** Luna schneidet in der Mehrheit der vergleichbaren Benchmarks besser ab als GPT-5.4, besonders bei Agentic-/Tool-Use-/Coding-Aufgaben (SWE-Bench Pro, FrontierMath).
- **Kosten-Nutzen:** Luna ist deutlich günstiger und damit für produktionsnahe Skalierung attraktiver.
- **Reine Wissens-/Reasoning-Benchmarks (GPQA):** GPT-5.4 hält knapp die Nase vorn, aber der Unterschied ist minimal.
- **"Luna Pro"** (mit High-Reasoning-Modus) dürfte diese Werte durch mehr Reasoning-Zeit nochmal steigern, ist aber teurer/langsamer als Standard-Luna.

Kurz gesagt: GPT-5.6 Luna ist trotz Positionierung als "günstiges Einstiegermodell" in vielen Benchmarks besser als das ältere Flaggschiff GPT-5.4 – bei gleichzeitig geringeren Kosten. Wer aber die absolute Referenzleistung sucht, sollte GPT-5.6 **Sol** oder **Terra** statt Luna vergleichen.