

GPT-6 Sol and Luna

PRIMARY

Publisher	OpenAI
Published	2026-09-22
Retrieved	2026-09-23
Cite key	openai2026gpt
URL	https://openai.com/index/introducing-gpt-6-sol-and-luna/

Assessment. The vendor's own launch post. Carries the pricing, the DeepSWE 1.1 and Agents' Last Exam figures with the reasoning-effort setting labelled, the five alignment evaluation categories, and the caveat that those evaluations 'do not measure failure rates in typical use'. Used for everything attributed to OpenAI directly.

Claims drawn from this source. CLAIM-2026-09-23-01, CLAIM-2026-09-23-02, CLAIM-2026-09-23-03, CLAIM-2026-09-23-07

GPT-6 Sol and Luna — vendor announcement (extracted 2026-09-23)

Two new models in the GPT-6 family offering "frontier intelligence" at different cost-efficiency tiers, using "similar methods as GPT-6 Astra" while prioritising affordability over maximum capability. Introduced after GPT-6 Astra "earlier this month".

Benchmarks as published

AutomationBench 1.0.6 (business workflows across 47 tools)

- GPT-6 Sol (xhigh): 33.2%, \$0.27 per task
- Claude Opus 5 (max): 26.9%, \$2.99 per task
- GPT-6 Luna (high): 5.4 percentage points over its predecessor at 58% lower cost

Agents' Last Exam V1

- GPT-6 Sol (max): 56.4%, 60% lower cost than Claude Opus 5

FrontierCode 1.1

- GPT-6 Sol matches "Claude Fable 5.1 xhigh at much lower cost"

DeepSWE v1.1 (software engineering)

- GPT-6 Sol (max): 68.8%, within 1.1 points of Claude Fable 5 (69.9%) at 80% lower cost
- GPT-6 Luna (max): 66.6%, 93% cheaper than Opus 5, 96% cheaper than Fable 5

OSWorld 2.0 offline (computer use)

- GPT-6 Sol (xhigh): 60.5%, ~80% cheaper than Claude Opus 5 (60.3%)
- GPT-6 Luna (max): exceeds GPT-5.6 Sol at one-tenth the cost

Pricing (per 1M tokens)

- GPT-6 Sol: \$2 input / \$10 output — 50% cheaper

- GPT-6 Luna: \$0.10 input / \$0.50 output — 50% cheaper

Availability

"Available in ChatGPT Work and Codex starting today for all Plus, Pro, Business, Enterprise, and Edu users." Free and Go users can access GPT-6 Luna in the desktop app. API ids: gpt-6-sol, gpt-6-luna. Rollout in ChatGPT is gradual through the day. No regional or EU restriction is stated.

Alignment

The models "build on the alignment work introduced with Astra, our most aligned model to date." Five alignment evaluation categories are presented: coding deception, broken search, reviewer bypass, warning circumvention, unauthorized interaction.

Caveat, verbatim in substance: "The evaluations below deliberately test challenging situations and do not measure failure rates in typical use."

A system card is linked externally to the deployment safety site. No statement is made about training data or compute.