

TOTAL MONTHLY AI SPEND

\$47,325

Seats + usage + infra

PAID-SEAT UTILIZATION

208 / 323

64% of paid seats active

RECOVERABLE WASTE

\$6,361

Open findings, per month

MONTHLY SPEND BY CATEGORY

CATEGORY	MONTHLY	SHARE
Seats	\$8,824	19%
Usage	\$35,501	75%
Other	\$3,000	6%
Infra	\$0	0%

TOP FINDINGS

OpenAI API spend spiked the week of 2026-07-06

COUNTEDCRITICAL

\$8,850/mo

Teams using Cursor heavily merge 16% more PRs

ESTIMATEDCRITICAL

\$7,153/mo

1,873 USD/mo recoverable by batching Nightly evals

COUNTEDMEDIUM

\$1,873/mo

ChatGPT renews 15 Oct 2026 with 62% of paid seats idle

COUNTEDMEDIUM

\$1,860/mo

62 unused ChatGPT seats

COUNTEDMEDIUM

\$1,860/mo

RETURN ON AI SPEND

-0.87× – -0.81× net ROI

Point estimate -0.84x · net monthly value \$7,496 on \$47,325 AI cost · J-curve 9 mo

$$\text{prs_month} = \text{DiD_prs_per_eng_week} \times (\text{workdays_month} \div 5) \times \text{engineers} \times \text{adoption_extrapolation};$$

$$\text{net_monthly_value} = \text{prs_month} \times \text{hours_per_pr_baseline} \times \text{loaded_cost_per_eng_hour} \times \text{quality_adjustment} \times (1 - \text{verification_tax_pct});$$

$$\text{netRoi} = (\text{net_monthly_value} - \text{total_ai_cost}) \div \text{total_ai_cost}$$

ASSUMPTIONS & SENSITIVITY (±20% PER DRIVER)

ASSUMPTION	VALUE	LOW ROI	HIGH ROI
Share of the standard price saved by running work on the batch tier (both vendors publish 50%)	0.5 ratio	-0.84x	-0.84x
Ceiling a batched request may take to return — work that cannot wait this long is not eligible	24 hours	-0.84x	-0.84x
Attribution keys a model needs before a best-vs-worst cache comparison is allowed	2 keys	-0.84x	-0.84x
Price of a cached-read input token as a multiple of the base input rate	0.1 ratio	-0.84x	-0.84x
Times a cached prefix is read back before it expires — measured from your own data where possible, this value is the fallback	2 reads	-0.84x	-0.84x
Cache-hit rate the modeled recovery raises a low-hit-rate model to	0.7 ratio	-0.84x	-0.84x
Price of a cache-write input token as a multiple of the base input rate (1-hour TTL)	2 ratio	-0.84x	-0.84x
Price of a cache-write input token as a multiple of the base input rate (5-minute TTL)	1.25 ratio	-0.84x	-0.84x
AI credits included per Copilot Business seat per month, pooled at the billing entity	1900 credits	-0.84x	-0.84x
Billed value of one GitHub AI credit	1 cents	-0.84x	-0.84x
AI credits included per Copilot Enterprise seat per month, pooled at the billing entity	3900 credits	-0.84x	-0.84x
Modeled share of frontier-class tokens assumed movable to mini-class models	0.25 ratio	-0.84x	-0.84x
Frontier-class token share above this trips the model-mix finding	0.5 ratio	-0.84x	-0.84x
Baseline engineering hours per merged PR	6 hours	-0.87x	-0.81x
A seat is unused after this many days with no activity	60 days	-0.84x	-0.84x
Fully-loaded engineer cost per hour	110 usd	-0.87x	-0.81x
Modeled mini-class cost as a ratio of frontier-class cost	0.2 ratio	-0.84x	-0.84x

Two coding seats each below this many active days per month count as redundant	5 days	-0.84×	-0.84×
Cohort deltas above this share are treated as self-selection, not tool effect — no dollars are modeled on them	0.5 ratio	-0.84×	-0.84×
A cohort outcome delta at or above this share counts as meaningfully positive	0.1 ratio	-0.84×	-0.84×
A seat priced above public list by more than this share counts as price drift	0.01 ratio	-0.84×	-0.84×
Discount for quality cost of AI-assisted work	0.85 ratio	-0.87×	-0.81×
A cost-of-going-faster measure worsening beyond this share while usage rises trips a finding	0.2 ratio	-0.84×	-0.84×
A renewal becomes a right-sizing opportunity when utilized seats sit below this share	0.7 ratio	-0.84×	-0.84×
A contract renewing within this many days is in the negotiation window	60 days	-0.84×	-0.84×
Trailing weeks that form the spend-spike baseline	8 weeks	-0.84×	-0.84×
Weekly spend this many standard deviations over baseline trips a spike	2.5 sigma	-0.84×	-0.84×
Monthly spend an API key or workspace must reach before the token rules evaluate it	25000 cents	-0.84×	-0.84×
A seat is underused below this many active days per month	3 days	-0.84×	-0.84×
Share of gains eaten by review/verification	0.15 ratio	-0.85×	-0.84×
Working days per month	21 days	-0.87×	-0.81×

APPENDIX — EVIDENCE & PROVENANCE

Every number above is reproducible: the source of record, the exact formula, and the inputs per finding. Thresholds are the org's editable assumptions; row-level drill-through is available on every finding in-product.

OpenAI API spend spiked the week of 2026-07-06

COUNTED

Source: Spend facts from card, invoice, and usage-API sources

Formula: " =weeklyCents trailing-8-week mean

zScore: 34.17 · weekIso: 2026-07-06 · weeklyCents: 1178199 · baselineWeeks: 8 · sigmaThreshold: 2.5 · baselineMeanCents: 293240

Teams using Cursor heavily merge 16% more PRs

ESTIMATED

Source: Team usage intensity + delivery telemetry (PRs merged)

Formula: estimated = deltaPct × lowBandMonthlyPRs × hours_per_pr_baseline × loaded_cost_per_eng_hour × quality_adjustment × (1 verification_tax_pct)

cap: 0.5 · caveat: self-selection: teams chose their tools — correlation, not proof · metric: prs_merged · deltaPct: 0.1596 · toolSlug: cursor · threshold: 0.1 · lowBandMean: 18.8 · highBandMean: 21.8

1,873 USD/mo recoverable by batching Nightly evals

COUNTED

Source: Billed usage-API dollars tagged with the vendor's own service tier, plus the latency-tolerant flag an admin set on the key

Formula: recoverable = synchronous-tier spend × batch_discount_share

label: Nightly evals · caveat: recovery assumes the whole key's work tolerates the turnaround you asserted · vendor: OpenAI · toolSlug: openai_api · attributionKey: proj_nightly_evals · batchDiscountShare: 0.5 · batchTurnaroundHours: 24 · synchronousSpendCents: 374633

ChatGPT renews 15 Oct 2026 with 62% of paid seats idle

COUNTED

Source: Contract terms (settings) + seat roster + active-day telemetry

Formula: trueDown = (1 activeShare) × monthly seat cost at latest snapshot

renewsOn: 2026-10-15 · toolSlug: chatgpt · totalSeats: 100 · activeSeats: 38 · activeShare: 0.38 · daysToRenewal: 45 · seatSnapshotAsOf: 2026-08-31 · utilizationBasis: usage_active_days_30d

62 unused ChatGPT seats

COUNTED

Source: Vendor seat roster (seat_facts snapshot) via connector sync or CSV import

Formula: recoverable = \$monthlyCostCents over inactive active seats

inactiveDays: 60 · monthlyCostCents: 186000 · affectedSeatCount: 62