CURLY'S CONSULTING / RESEARCH PAPER 04

How to price AI

Ammanuel "Tom" Selameab, CPP

Certified Pricing Professional | Curly's Consulting

September 2026

A research-led method for choosing the offer, charging unit, price level, and commercial controls.

Abstract

Pricing AI requires a decision about the offer, the charging unit, the price level, and the risk each party accepts. Inference cost supplies an economic constraint, but it does not establish customer value or willingness to pay. Competitor prices supply context, but comparisons break when one supplier charges per attempt, another per successful result, and a third includes baseline AI inside a subscription. This paper presents a research-led method for answering four questions: which workflow creates enough value to fund a purchase; which commercial structure buyers understand and accept; what price and allowance produce attractive contribution; and how the business will verify success and control the bill. Public examples from HubSpot, Intercom, and GitHub demonstrate that changing the counted event can matter as much as changing the headline rate. An illustrative operating model then shows how success rate, review work, and usage affect both buyer value and supplier contribution. The conclusion is that AI pricing should be commissioned as an evidence problem. Buyer interviews, realistic offer-choice research, cost telemetry, and a paid validation step provide the inputs required for a defensible recommendation. A rate card alone cannot answer those questions.

1 Define the paid problem before defining the AI plan

"AI assistant" describes a technology category, not a purchase. "Resolve a defined class of support requests using approved customer data" describes a workflow with a user, a result, and a boundary. The second description makes research into value, adoption, supervision, and budget possible.

Start with a commercial hypothesis: a named buyer experiences a consequential problem, a credible alternative is costly or inadequate, and the proposed workflow creates a benefit large enough to justify change. Specify the user, economic buyer, approver, and implementation owner. These roles may have different preferences and constraints.

Research the customer's current process before demonstrating the AI product. Ask about a recent instance, frequency, handling time, failure consequences, workaround, and expenditure. A respondent who admires a demonstration may still have no purchase trigger or usable budget. The study should distinguish curiosity from an economically meaningful buying occasion.

Define the alternative honestly. It may be existing staff, a general-purpose AI service, an incumbent product, outsourcing, or choosing not to do the work. Value is incremental to that alternative. Comparing

the product with an implausibly expensive manual process can manufacture apparent headroom that disappears in an actual sale.

2 Measure value after adoption, review, and errors

A useful starting model for a repetitive workflow is:

$$V_{\text{gross}} = Nastw, \tag{1}$$

$$V_{\text{net}} = V_{\text{gross}} - C_{\text{review}} - C_{\text{error}} - C_{\text{change}}. \tag{2}$$

Here N is eligible tasks, a adoption, s the proportion of attempted tasks producing an accepted result, t hours avoided per accepted result, and w the relevant hourly value. The cost terms capture review, expected error consequences, and the customer's change effort. Define every input in the actual workflow.

Time value is not automatically cash saving. The customer may retain the same staff and use released capacity for other work. That can be valuable, but the claim should then concern capacity or additional output, rather than a booked reduction in labor expense. Revenue benefits need their own mechanism and evidence.

Avoid double counting. Higher throughput and less handling time may describe the same released capacity. A lower error rate can create additional value, but only if the errors and their consequences are defined and the improvement is established against a credible baseline.

Quality and assurance belong in the economic proposition. NIST's Generative AI Profile identifies risks including confabulation, privacy, and information security.[1] The commercial implication is that review and controls can be required to make the workflow purchasable. Their costs should not disappear from the model merely because they are absent from a product demonstration.

3 Choose which part of the product the customer buys

AI can be included in a core subscription, differentiate a premium package, form a separate add-on, or support a usage or outcome charge. These are alternatives to test, not stages every company must pass through.

Architecture	Buyer proposition	Research question
Included AI	Better core workflow at the package price	Does inclusion improve choice, retention, or realized core price enough to cover cost?
Premium package	A complete higher-value workflow	Which segment will upgrade rather than accept a cheaper substitute?
Separate add-on	A distinct capability for a subset of buyers	Is there an identifiable use case, budget, and incremental willingness to pay?
Usage charge	Payment varies with an observable activity	Does the bill follow value, and does charging suppress useful adoption?
Outcome charge	Payment follows a defined accepted result	Can success be verified, attributed, and explained without dispute?
Base plus allowance	Predictable access with scalable additional use	Which included quantity, commitment, and overage rule fit the buyer and cost distribution?

Model the effect on the existing product. An AI add-on can reduce required seats, replace paid services, or shift demand into another package. Its own revenue is not necessarily the net monetization effect. The relevant comparison is total contribution under the full offer menu.

Likewise, inclusion is a commercial strategy rather than evidence that the capability has no value. The company may monetize it through broader adoption or premium selection. Those mechanisms need evidence; they should not be assumed because a familiar company has chosen to bundle AI.

4 Research the denominator of the price

Public changes illustrate why a rate is incomplete without the billable definition. HubSpot's April 2026 change moved Customer Agent from $1 per conversation to $0.50 per resolved conversation, while its prospecting agent moved from recurring contact enrollment to payment for a recommended lead.[2] Narrowing the event and changing the rate are separate economic actions.

Intercom's March 2026 Fin announcement moved from fully autonomous resolutions toward configured outcomes, including successful Procedures ending in a planned human handoff.[3] A broader definition can increase billable events without a higher nominal rate. A handoff that is intentionally completed is also different from a failed attempt or an unplanned escalation.

GitHub's April 2026 announcement scheduled Copilot's move from premium requests to token-based AI Credits for June 1, with exceptions for existing annual plans.[4] That ties consumption more closely to model and workload intensity. It does not, by itself, establish that the customer can predict the bill or values every token equally.

For your product, test whether buyers can explain the unit, reproduce the count, estimate expenditure, and influence the activity without defeating the workflow's purpose. A charge for retries or internal tool calls may follow vendor cost while looking like a charge for inefficiency to the buyer. A completed-task unit may be easier to value but harder to verify.

5 Specify success before selling outcome pricing

Write a billable-event definition before quoting an outcome rate. Include the task, acceptance rule, measurement period, exclusions, duplicate handling, correction policy, and treatment of reopened work. Clarify whether human review is compatible with the promised outcome.

Separate a product-controlled event from an external business result. A qualified lead is not a closed-won sale. A generated answer is not necessarily a resolved problem. A processed document is not necessarily an accepted or useful document. The contract should not imply certainty about results that depend on unobserved customer actions.

Test the definition with both buyer and vendor operating teams. Give each party the same sample records and ask them to calculate the bill. Disagreement reveals a measurement problem that cannot be fixed by selecting a more attractive price. Include exceptions and failure cases rather than only a successful demonstration.

Research trust as well as comprehension. Buyers may understand a vendor-controlled success measure and still doubt its fairness. They may prefer a different unit, a capped bill, or a commitment with clearer controls. Those are commercial preferences to measure, not objections to dismiss.

6 Build supplier economics from the usage distribution

Record cost at the workflow and account level. Relevant drivers can include model routing, context, output length, retrieval, tool calls, retries, infrastructure, assurance, and support. The average API charge is only one element of delivery.

For a bill function $B(q)$, total relevant delivery cost $D(q, s)$, and incremental account-level expense S, contribution is:

$$K = B(q) - D(q, s) - S.$$

The success rate s matters when unsuccessful attempts consume cost but do not produce revenue. Usage q matters when flat fees expose the supplier to heavy consumption. Both may change when the tariff changes, so applying a new price to the old usage distribution is only a mechanical scenario.

Inspect low-volume accounts, typical use, high consumption, and unusual cost profiles. A plan can be attractive at the median while losing money on customers whose legitimate use is expensive. An allowance, commitment, different workflow scope, or model-routing change may be more appropriate than an unlimited promise.

A target contribution margin sets a constraint, not willingness to pay. If relevant cost is C and the target margin is m, the required price under a simple fixed-fee model is $P \geq C/(1 - m)$. If credible buyer demand does not support that price, the business must change cost, scope, segment, or the economic target. A preferred margin does not make the customer value appear.

7 Worked example: buyer value and supplier margin can diverge

Illustrative assumptions; not vendor prices or a client result. A customer has 12,000 eligible monthly tasks. Half are attempted through AI and 80 percent of attempts produce an accepted result. Each accepted result avoids two minutes of work valued at $45 per hour. Gross capacity value is:

$$12{,}000 \times 0.50 \times 0.80 \times \frac{2}{60} \times \$45 = \$7{,}200.$$

Every attempted task requires 0.4 minutes of review, worth $1,800 in total. Allow another $600 for expected error and change effort. Net capacity value before the product fee is $4,800 per month. Whether that becomes cash saving or additional useful output requires evidence from the buyer.

For the supplier, assume delivery cost of $0.07 per attempt and $300 of incremental account-level infrastructure and support. At 6,000 attempts, total relevant cost is $720. A $1,200 flat fee creates $480 of contribution, or 40 percent before development and shared operating expense. A 70 percent target margin would require $2,400 at the same cost, subject to demand.

Now consider charging $0.25 per accepted result. The normal case still produces a $1,200 bill, but the supplier absorbs more performance risk:

Scenario	Attempts	Acceptance	Bill	Contribution
Normal workflow	6,000	80%	$1,200	$480
Lower acceptance	6,000	50%	$750	$30
Higher usage	18,000	80%	$3,600	$2,040

Each scenario uses $0.07 cost per attempt, $300 additional account expense, and $0.25 per accepted result. Shared expense is excluded. Acceptance and usage are assumptions, not empirical estimates.

The outcome unit makes the buyer's bill depend on accepted work, but it does not make supplier risk vanish. A price recommendation must connect the buyer's choice with the success and cost distribution under that offer. The example therefore creates specific research questions rather than a universal answer about flat or outcome pricing.

8 Estimate a price range from realistic offer choices

Qualitative research should define the relevant buyer, alternatives, workflow, and language. Quantitative work should test concrete offers with scope, unit, expected bills, limits, and service responsibilities. Include the option to retain the current process or not purchase.

Choose the method according to the decision. A fixed-offer price test can compare a manageable range of rates. A choice-based study is more useful when price, package, unit, and allowance change together. Holdouts and diagnostic checks help assess model performance within the study, while actual purchase observations address external validity.

Willingness-to-pay research comparing elicitation methods with real purchase behavior reinforces the need to calibrate stated response.[5] Do not equate a survey's choice share with an unqualified forecast of paid adoption. Recruitment, implementation effort, procurement, and the purchase horizon can all differ from the research setting.

Optimize the business objective rather than conversion alone. The candidate with the most demand can produce less contribution. A higher price may attract a different segment with higher support requirements. Simulate the complete menu using segment-specific response and cost assumptions, and show which uncertainty changes the preferred offer.

9 Use paid validation to test a named uncertainty

A useful pilot has a question: will buyers pay at these terms, does accepted output support the value claim, are high-usage economics viable, or does the offer cannibalize the core product? Define the observation needed before recruiting participants.

Record purchase, realized price, activation, attempted and accepted work, review, cost, support, and continued use. Preserve the exact pilot offer. Intensive custom assistance, special guarantees, or deep discounts can make the pilot different from the standard product the company intends to sell.

Set expansion, revision, and stop criteria around economics, buyer comprehension, and product performance. The threshold is company-specific. A successful pilot does not need to prove every long-run outcome, but it should resolve the uncertainty required for the next commitment and identify what remains unobserved.

10 Conclusion: commission the evidence behind the AI price

The right AI price is part of a commercially coherent offer. It depends on a valuable workflow, a buyer with a reason and ability to pay, an understandable bill, a reliable measure of success, and a cost distribution the supplier can sustain. Competitor tariffs and inference costs cannot supply those buyer-specific answers.

Pricing research should determine which segment values which outcome, which package and unit buyers choose, how they evaluate expected and high bills, and what price range supports attractive contribution. Cost and product evidence should establish accepted output, failure handling, review, and operating requirements. Commission those inputs as one decision scope so the result is a recommendation the company can implement and validate.

Curly's Consulting's **Pricing Transformation** combines original buyer research and willingness-to-pay analysis with monetization design, size-of-prize economics, and execution support. For AI, the scope should follow the workflow and the unresolved commercial questions. If an internal team will conduct the work, the **Pricing Workshop** provides the tailored research roadmap, methods, and tools; the team remains responsible for collecting evidence and developing its recommendation.

Turn the AI pricing question into a research brief

Request a Pricing Transformation quote. For a team-led research effort, buy the Pricing Workshop.

References

[1] Autio, C., et al. (2024). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. NIST AI 600-1. https://doi.org/10.6028/NIST.AI.600-1.

[2] HubSpot. *HubSpot's Customer Agent and Prospecting Agent: Now you pay when the task is complete*. 2026-04-13. https://www.hubspot.com/company-news/hubspots-customer-agent-and-prospecting-agent-now-you-pay-when-the-task-is-complete.

[3] Intercom. *From resolutions to outcomes: Evolving how Fin delivers value*. 2026-03-12. https://www.intercom.com/blog/from-resolutions-to-outcomes-evolving-how-fin-delivers-value.

[4] GitHub. *GitHub Copilot is moving to usage-based billing*. 2026-04-27. https://github.blog/news-insights/github-copilot-is-moving-to-usage-based-billing/.

[5] Miller, K. M., Hofstetter, R., Krohmer, H., and Zhang, Z. J. (2011). How should consumers' willingness to pay be measured? An empirical comparison of state-of-the-art approaches. *Journal of Marketing Research*, 48(1), 172–184. https://doi.org/10.1509/jmkr.48.1.172.

About the author. Ammanuel "Tom" Selameab, CPP, is a Certified Pricing Professional and founder of Curly's Consulting. He helps private equity investors and portfolio-company leaders answer pricing and monetization questions through primary buyer research, commercial analysis, and financial modeling.

Curly's products | tom@curlysconsulting.com

© 2026 Curly's Consulting. All rights reserved.