orb.

2025 state of AI agent pricing

Introduction

The pricing of AI agent products is still in its early stages. Companies are actively testing flat rates, usage-based fees, credit systems, and hybrid models to find the right balance between value and cost.

Clear patterns are starting to emerge. With more examples to study, teams can now make more informed decisions instead of relying on guesswork.

By analyzing the pricing strategies of 66 companies offering AI agent products, this report reveals how companies are structuring their pricing, where they're encountering friction, and which approaches are proving most resilient. Whether your AI agent is the core product or one piece of a broader platform, these insights can help inform your pricing strategy.

[Learn more about our methodology and dataset.](#)

orb.

Contents

orb.

Common pricing models

While AI agent pricing is still evolving and there are more nuances to pricing models, a handful of core models have emerged as the most common in the market. These categories are how we segmented and analyzed the 66 companies in this report. Each model reflects a different approach to balancing value delivered with costs incurred. This is particularly important when it comes to AI products, where every interaction carries a real compute cost.

Subscription

Subscription pricing is a flat recurring fee for access to the product. The fee is typically charged monthly or annually and includes unlimited users and usage.

Subscriptions offer simplicity, but they carry risk in AI. The variable costs of AI lead to margin compression since heavy users aren't offset by pricing.

Per user or seat

This pricing model refers to a recurring fee charged for each user or seat. User- or seat-based pricing has its roots in SaaS. With high fixed costs and near-zero marginal costs, charging per user results in high gross margins and scalable revenue.

While familiar and easy to forecast, per-seat pricing often struggles to reflect actual usage or value delivered, especially when AI features are used unevenly across a team.

Usage-based

Charges vary based on how much the product is used. This can be measured in terms of API calls, tokens, image generations, or other units of consumption.

Usage-based pricing aligns well with the real-time costs of AI, and it scales cleanly with customer value. However, it introduces unpredictability for both buyers and sellers, which is why many companies combine it with other models.

Outcome-based

Pricing is tied to specific results, such as the number of leads generated, tickets resolved, or conversations completed.

This model directly links pricing to business value, making it easier for buyers to justify the cost. It's emerging in categories where outcomes are clearly measurable and where the AI agent replaces a human task.

orb.

Freemium or free trial

Freemium models offer ongoing access to a limited version of the product at no cost, with upgrades available for increased usage or advanced features. Free trials offer full or partial access to the product for a limited time, after which users must convert to a paid plan.

These models reduce friction at the top of the funnel, especially in consumer or developer-facing products. The challenge is converting users at a sustainable rate without over-subsidizing expensive usage.

Tiered

Tiered pricing offers multiple pricing packages at increasing price points. Most commonly, this includes two or three tiers, but some companies offer more to capture a broader range of customer needs and willingness to pay.

Tiered pricing is also used to differentiate offerings by features, model quality, support level, or included usage. This strikes a balance between predictability and flexibility.

Add-ons

Add-ons are flat-fee enhancements to the core product, sold separately. These may include advanced features, premium models, higher-speed processing, or expanded support.

Add-ons allow companies to monetize power users without forcing everyone into a higher base tier. They also create upsell opportunities without increasing the complexity of the base plan.

Hybrid

Hybrid pricing models combine multiple types of pricing into one offering. With many different types of pricing available, there's an incredible amount of ways to structure a hybrid model. Most commonly, hybrid pricing combines pricing that generates predictable recurring revenue (e.g., subscription, user/seat-based) with pricing that captures value and covers costs (e.g., usage-based, outcome-based).

The goal of hybrid pricing is to get the best of both worlds: predictability and simplicity plus fairness and scalability.

Want a deeper breakdown of these models and how to apply them? Download our Pricing AI Agents eBook for detailed strategies, examples, and frameworks.

orb.
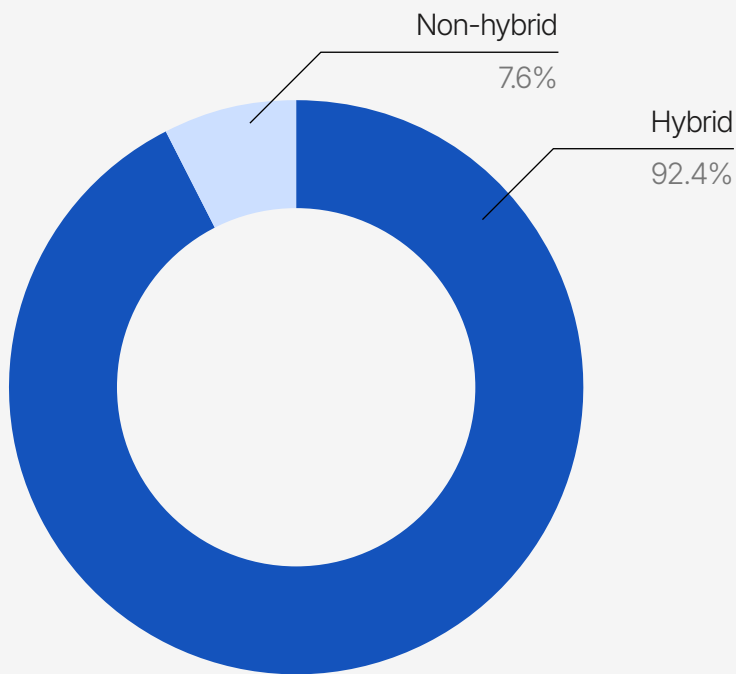
Findings and insights

Hybrid models reign supreme

The dominant trend in AI agent pricing is clear: hybrid models are winning. An over-whelming **92.4% of companies in our dataset use a hybrid approach**, combining multiple pricing components into a single strategy.

Companies using hybrid vs. non-hybrid pricing

Non-hybrid
7.6%

Hybrid
92.4%

This isn't a coincidence. Hybrid models offer a blend of predictability and flexibility. They generate recurring revenue through subscriptions or seats, while also capturing variable value through usage, outcomes, or add-on services. For companies selling AI agents, where infrastructure costs scale with usage, this flexibility is crucial for protecting margins and aligning prices with the value delivered.

orb.

Most commonly used components in hybrid pricing structures
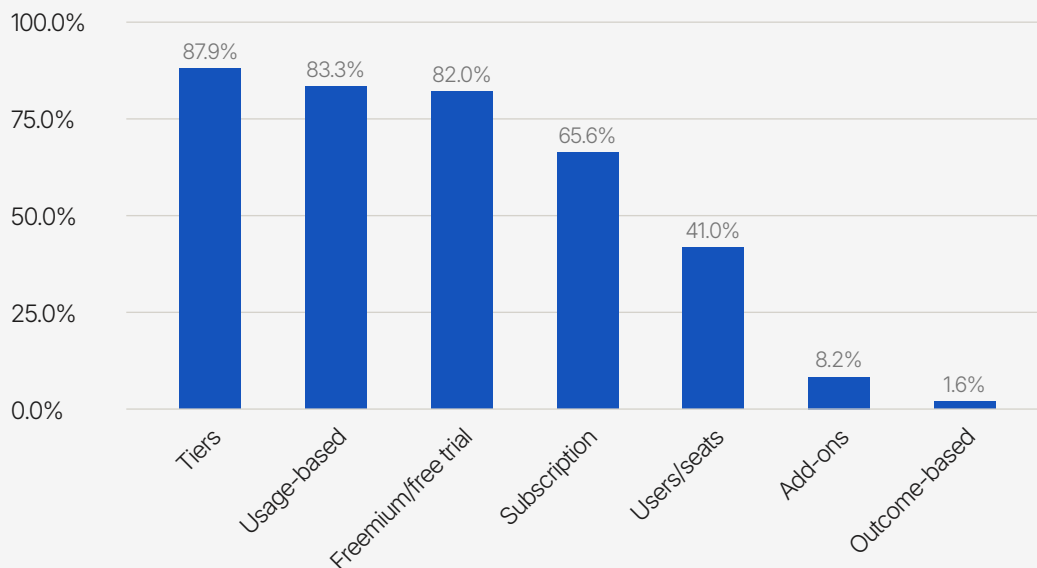
Hybrid models are much more than popular; they're multi-dimensional. Most companies in this category utilize four pricing components, mixing and matching them depending on product design, go-to-market strategy, and customer segmentation.

The most common hybrid combination (39.3%) was:

Subscription + usage-based + freemium or free trial + tiers

The second most common (20.7%) was:

Users/seats + usage-based + freemium or free trial + tiers

This level of complexity enables companies to serve both high- and low-volume users, reduce onboarding friction, and scale revenue in tandem with usage. But it also introduces operational challenges.

orb.

Managing the complexities of hybrid pricing

With so many components that can make up a hybrid pricing structure, billing quickly becomes complex, and there are significant consequences to not having the right billing infrastructure in place. Customers may churn or you may lose out on revenue due to inaccurate billing or usage tracking. Hybrid pricing can also cause internal friction by making it difficult to understand how invoices are calculated or what drives revenue.

Implementing and managing hybrid pricing requires flexible, agile, and accurate billing infrastructure like Orb. With Orb, companies have the control and visibility needed to run sophisticated hybrid pricing at scale, without sacrificing billing accuracy or time to market.

orb.

Example of hybrid pricing

Windsurf has several AI agent products: an AI editor, AI plugin, and AI code reviewer. It offers different tiers to target both individuals and companies, and these tiers are some of the more complex hybrid pricing models in our dataset.

Pricing plans start with a free, subscription, or per user component. Then, usage limits are layered on, with each plan including a limited number of AI credits. Most plans include the ability to buy additional credits, except for the Free tier. This is probably because the company wants to encourage Free users to upgrade to Pro.

Interestingly, the Teams plan is the only one that has a feature add-on. This add-on for single sign-on is priced per user. As the only feature that is available a la carte, this indicates it may be critical enough for companies that they're willing to pay extra for it.

orb.

SaaS pricing has its place, but only when paired with usage

Traditional SaaS pricing — charging per user or offering a flat-rate subscription — doesn't work when applied to the economics of AI. In SaaS, additional usage doesn't significantly increase cost. In AI, every query, generation, or interaction carries real compute expense. That makes flat-rate pricing risky on its own.

Still, SaaS pricing hasn't disappeared. Instead, it's being combined with other models, especially usage-based pricing. **Among companies using subscription or user/seat-based pricing, 85.2% also included usage-based pricing** in their models. This reflects a broader shift: companies are preserving the simplicity and predictability of SaaS while layering in variable pricing to align revenue with real-time costs.

Between the two types of SaaS pricing, **subscriptions (62.1%) were more commonly used than user/seat-based pricing (39.4%)**. This may be due to combining traditional SaaS pricing with usage-based pricing. By charging for usage, it makes less sense to also charge per user or seat.

Both pricing models attempt to capture more revenue as a product is used more. However, usage-based pricing is more successful because it reflects the fact that usage varies across users. Pricing per user or seat results in companies overpaying for low-usage users or vendors losing out on potential revenue from power users.

Done right, SaaS pricing models can still play a role in a successful monetization strategy for AI agents. However, without usage-based components, they'll expose companies to unpredictable margins and loss of revenue.

The takeaway: SaaS pricing isn't dead, but it needs to be paired with usage-based pricing to create a sustainable revenue model for AI products.

orb.

Example of user/seat-based pricing

Fireflies.ai is an AI meeting assistant that's mostly targeted toward companies, but can also be used by individuals. Its pricing plans are tiered and centered around charging monthly per seat, with the ability to get started for free and usage limits.

By charging per seat, the company can capture more revenue with its target customer (businesses). Simultaneously, it places limits on usage of certain features and grants credits for AI features. This is probably to prevent its cost from skyrocketing with power users.

orb.

Multiple pricing models expand user base and enable experimentation

A **subset of companies (12.1%) have adopted multiple distinct pricing models**. These aren't just hybrid structures, but different monetization strategies offered in parallel that often target different user types.

The **most common approach (62.5%) is to differentiate between per-user/seat pricing and flat-rate subscription**. Many AI agent products are useful for businesses and consumers, so this split reflects that audience segmentation: per-user pricing is aimed at businesses, while flat-rate subscriptions are designed for individuals. In both cases, these models are also layered with tiers, usage-based pricing, and freemium or free trial access.

This strategy provides companies with the flexibility to reach broader markets without imposing a one-size-fits-all pricing structure.

Offering different pricing structures also enables companies to try out new, unproven models while decreasing risk. We can see this when it comes to outcome-based pricing. Among companies in our dataset that use outcome-based pricing, two-thirds also offer at least one other pricing model. Outcome-based pricing holds promise for aligning price with business value, but it's still emerging.

Companies are currently working to define the right outcome metrics and what to charge, as outcomes need to be easy to track, clearly tied to delivered value, and defensible when customers ask, "Why does this cost what it does?" Offering alternative models alongside outcomes gives companies a safety net. It lowers the risk if outcome-based pricing underdelivers on revenue or if the AI agent doesn't deliver the expected results.

For companies still refining their monetization strategy, offering multiple models is about revenue diversification. It's also a path to faster learning and broader adoption.

orb.

Example of multiple pricing models


The AI code editor Cursor is useful to both individual developers and companies, and this is reflected in its tiered pricing. The company offers a Hobby and Pro plan that's geared toward individuals, along with a Business plan for teams. These plans are made up of a combination of a free trial, usage limits, and either a flat-rate subscription or user/seat-based pricing.

orb.

Few charge for outcomes, presenting a large opportunity

Outcome-based pricing is the least common model (4.5%) in our dataset. Of those, only one company relies on it exclusively. The rest offer outcome-based pricing alongside at least one other model, allowing customers to choose the structure that best fits their needs.

Companies using outcome-based pricing

Offers outcome-based pricing and at least one other model
66.7%

Only outcome-based
33.3%
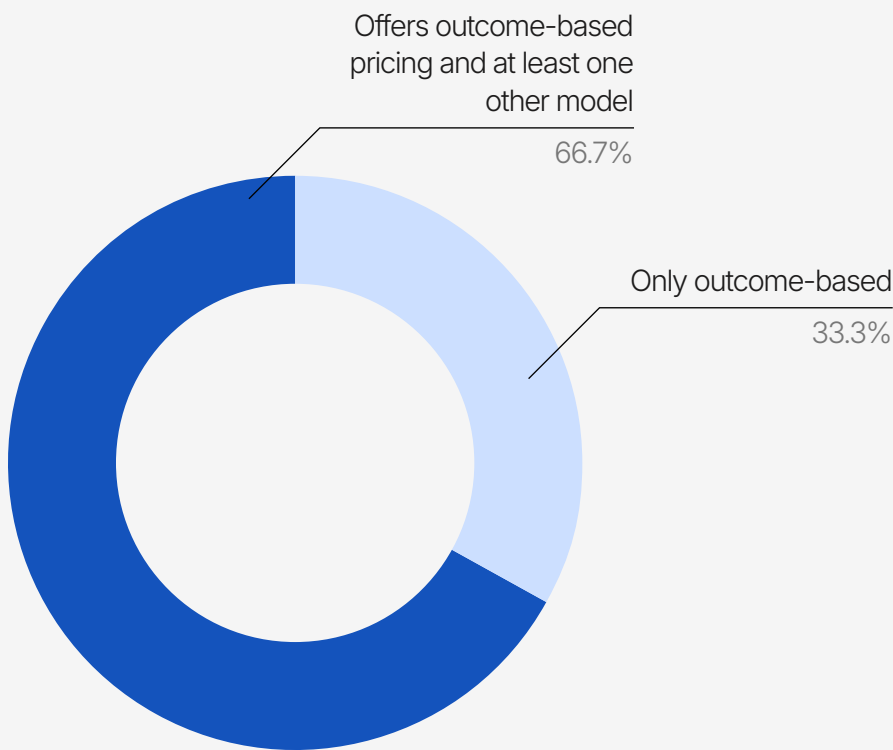
This signals that outcome-based pricing is still in its experimental phase. It also shows that a very small number of companies have figured it out and are confident enough to solely rely on it. When outcome-based pricing works, it works well: outcome-based pricing directly ties revenue to customer success, offering a powerful way to differentiate your product and maximize return on investment (ROI).

Many companies are interested in outcome-based pricing, as The 2025 State of B2B Monetization reported a 20% increase in companies that will use outcome-based pricing as their primary model by 2028.

orb.

To achieve this though, companies face three barriers:

1. **Picking the wrong outcome metric** — one that's hard to measure, not compelling to the customer, or doesn't scale with value.

2. **Failing to deliver the promised outcome**, especially if customer success is influenced by external factors.

3. **Lacking the infrastructure to track outcomes with confidence**.

This last point is a big blocker that can be overcome with the right billing infrastructure. Unlike usage-based pricing, which can be tied to easily captured events (e.g., token counts, queries), outcome-based pricing demands precise, real-time visibility into whether a result has been achieved.

That's where a billing platform like Orb is essential. Orb's raw event architecture lets companies track and price against the exact outcomes they care about, whether that's a lead generated, a conversation resolved, or a report delivered. Orb also decouples event data from pricing metrics and enables anyone to change pricing using SQL or a visual editor, making it easy to iterate. If your initial metric doesn't work, you can swap it out or refine it without overhauling your entire billing system or requiring engineering help.

For companies willing to navigate the complexity, outcome-based pricing presents a significant whitespace. The ones who figure it out early will have a meaningful edge.

orb.

Example of outcome-based pricing

Intercom's Fin AI Agent is included in Intercom's core product. Companies pay per resolution for Fin, as well as per user for access to Intercom's helpdesk. Fin is also available as a standalone product that is priced per resolution. For both options, a free trial is provided.

orb.

Methodology

This report is based on a pricing model analysis of 66 companies that either:

- Offer an AI agent as their primary product,

- Have a core product that includes an AI agent feature or add-on,

- Or they offer platforms for building AI agents.

Companies that offer APIs were excluded.

To ensure relevance and statistical confidence, we based our market sizing on an estimated 17,500 AI companies in the U.S., which results in a 10% margin of error.

Pricing model components analyzed

Pricing component	Definition
Subscription	Flat recurring fee for access to the product
Per user or seat	Recurring fee charged per individual user or seat
Usage-based	Pricing varies based on volume of consumption (e.g., API calls, generations)
Outcome-based	Charges tied to specific outcomes achieved (e.g., leads generated, tickets resolved).
Freemium or free trial	Free access to limited functionality or time-bound access to the full product
Tiered	Multiple pricing packages at different price points
Add-ons	Flat-fee enhancements or upgrades offered in addition to a base plan

orb.

If a company used more than one of these pricing components, we classified them as using a hybrid model.

We also made a distinction between companies that offer multiple distinct pricing models (e.g., subscription for individuals and seat-based pricing for teams) versus those that offer a single pricing model with multiple components.

In the case of companies with multiple models, each offering was broken down individually to reflect the components it included.

Maximize revenue from your AI agent with Orb

Whether you're experimenting with usage-based billing, rolling out a new hybrid model, or exploring outcome-based pricing, Orb gives you the infrastructure to move fast without sacrificing accuracy.

With Orb, you can:

• Track usage and outcomes in real time

• Test and launch new pricing models without engineering lift

• Ensure every invoice reflects your value, down to the last event

With AI rapidly advancing, your billing system needs to be just as dynamic as your product. Let's make sure you're pricing for growth and getting paid for the value your AI agent creates, even as the market evolves.

orb.

[Learn more about Orb](#)
to see how we can help you implement the right monetization strategy for your AI agent product.

orb.