

# Analysing the AI Frenzy: Paths to Profitability and OpenAI's Record Valuation

#### Introduction

In 2019, OpenAI's CEO Sam Altman famously declared that the company did not have a clear plan on how to make money, but rather it planned to build Artificial General Intelligence first and then ask it how to generate a return. Nowadays, OpenAI stands at the centre of the most important technological and financial race of the decade. OpenAI has led an industry-wide surge that has drawn hundreds of billions in venture capital and pushed hardware demand to unprecedented levels. Yet, behind the record valuations and rapid technological adoption, fundamental questions remain. Can generative AI companies, whose models consume enormous resources and whose user bases are mostly non-paying, ever achieve sustainable profitability?

### The current landscape of generative AI and LLMs

Since the launch of ChatGPT in late 2022, the generative AI market has grown at an extraordinary pace. ChatGPT, built on OpenAI's GPT series of Large Language Models, has surpassed 500m monthly users in only two years, working as a catalyst for the surge in venture funding and competition in the sector. LLMs are advanced AI systems trained on vast datasets to understand and generate human language. They function as statistical prediction machines, learning patterns in text to anticipate the next word in a sequence. Their versatility allows for a vast range of applications, such as text and code generation, summarization, sentiment analysis, and reasoning. However, LLMs face key limitations, as they can produce inaccurate or misleading information and often reflect biases from their training data. The landscape of generative AI firms has nonetheless attracted record funding in the recent period, and following OpenAI's early lead has seen established tech giants and high-growth start-ups join the fierce competition for leadership in LLMs.

Anthropic has emerged as OpenAI's closest competitor. Founded in 2021, the company has built its Claude series around principles of safety and enterprise reliability, differentiating itself through a 'Constitutional AI' framework designed to align model outputs with ethical standards. This positioning has attracted major strategic backers, including Google [NASDAQ: GOOGL] and Amazon [NASDAQ: AMZN], and a vast range of corporate clients. Anthropic's focus on the B2B market has translated into higher revenue per user and longer-term contracts, with run-rate revenue climbing from roughly \$1bn at the start of 2025 to over \$5bn by August. The company's \$13bn Series F round, led by ICONIQ Capital, valued it at \$183bn. However, profitability is still far, and the firm continues to burn significant capital on computing resources and R&D.

Other significant players in the LLMs space are currently represented by Google's Gemini, Meta's [NASDAQ: META] Llama series, and xAI's Grok model. Google has leveraged its vast ecosystem to scale AI distribution faster than any of its rivals. Its Gemini AI app surpassed 400m monthly active users in 2025, while AI-powered overviews in Google Search now reach over 1.5bn people each month. At the same time, xAI, founded by Elon Musk and merged with X in early 2025 with a combined valuation of \$113bn, introduced the practice of using social media data for model training infrastructure. Musk's intent will then plausibly be to leverage public data from X to enhance his Grok chatbot and build a vertically integrated AI platform. Meta is instead pursuing a distinct path with its Llama series. Integrated across Facebook, Instagram, and WhatsApp, Meta's AI products now reach over 1bn users, positioning the company to monetize generative models within its social media ecosystem. Llama, downloaded more than 400m times, has become the most popular quasi—open-source model, available freely to developers although without full transparency of its underlying code. Additionally, Meta is investing heavily in a "superintelligence" research lab to compete directly with OpenAI and Google.



At the start of 2025, however, this optimistic outlook for LLMs, characterised by sustained investment, was briefly threatened by the launch of DeepSeek, a hyper-efficient Chinese AI model. Its strong performance at a fraction of the cost prompted fears that tech companies and Silicon Valley start-ups were overspending on data centres and expensive hardware to power their models. Long-date venture capitalist Mary Meeker highlighted that while US players like OpenAI and Google continue to lead, their training costs are escalating sharply. Meanwhile, cheaper domestic and overseas competitors are emerging with smaller, specialized models tailored to specific tasks. This trend threatens to erode pricing power and disrupt the "one-size-fits-all" LLM architecture that has defined the industry so far. Adding to this pressure is the rise of open-source AI models. Unlike proprietary systems, open models make their code publicly available (or at least in Meta's case, only some parameters the model uses to generate outputs), allowing developers to modify and deploy them freely. They eliminate licensing costs, but they still require significant infrastructure investment.

Lastly, when assessing the current AI landscape, one should question whether LLMs represent the right path for the future development of AI. While they have defined the current AI cycle, many experts argue that LLMs are approaching their limits. Models such as OpenAI's GPT series have remarkable linguistic fluency but remain pattern recognition systems that predict words rather than understand meaning. This reliance on statistical inference leads to persistent flaws such as hallucinations, bias, and a lack of genuine reasoning. Signs of diminishing returns from ever-larger models raise doubts about maintaining the trend of rapid improvements observed over the last two years. As a result, researchers are turning to hybrid approaches such as neurosymbolic AI, which combines the pattern recognition of neural networks (as in LLMs) with the structured reasoning of symbolic logic. Alternatively, other options being considered are Mixture of Experts models that activate only the required parts of a model to enhance efficiency and reduce processing costs. Ultimately, the valuation and long-term success of today's leading AI companies will depend on their ability to adapt to these technological shifts and to the evolving foundations of artificial intelligence itself, adding yet another source of uncertainty for investors, even in leading firms like OpenAI.

## The case of OpenAI: few paying customers and high fixed costs

OpenAI, with one of the fastest revenue growth trajectories in software history, reached an ARR of roughly \$12bn in July 2025. The user base has also expanded rapidly, increasing from around 500m weekly active users in March to approximately 800m today. However, less than 5% of these users pay for a subscription, with the majority of revenue generated by enterprise clients. When the premium version was first introduced in 2023, ChatGPT counted roughly 100m users but while global adoption of its services has grown at an estimated 165% CAGR over the past two years, the number of individual subscribers has not increased at the same rate. The freemium model has been crucial in driving mass adoption, data collection, and scale advantages, but it has also exposed the limits of OpenAI's monetization strategy. With no advertising-driven revenue stream, OpenAI must rely on premium services and enterprise partnerships to balance rapid user growth with financial sustainability. The company's revenues are currently derived from three main sources. Consumer subscriptions, including ChatGPT Plus and Teams, account for roughly 55% to 60% of total income. Enterprise solutions, such as ChatGPT Enterprise and tailored AI integrations, contribute an additional 25% to 30%. The API and developer platform segment represents 15% to 20% but is the fastest-growing line of business, driven by expanding usage of GPT-API by enterprises.

On the cost side, OpenAI's structure is capital-intensive. In the first half of 2025, capital expenditures significantly rose as both R&D spending and stock-based compensation increased by 50% to reach \$2.5bn. Analysts project that, by 2030, AI firms will collectively need to deploy around \$500bn per year in capital expenditures to meet the anticipated demand. At current growth rates, however, the sector is expected to fall short of the roughly \$2tn in annual revenue required to sustain that level of investment. Operating expenses are also significant, driven largely



by energy consumption and GPU usage necessary to maintain and expand operations. To mitigate these costs, OpenAI has entered partnerships with Nvidia [NASDAQ: NVDA], AMD [NASDAQ: AMD], and Broadcom [NASDAQ: AVGO] to internalize part of its chip supply and improve hardware efficiency.

### Possible paths to profitability for OpenAI and LLMs firms

The analysis of OpenAI's economics in the previous section highlights the firm's challenge in targeting profitability. The combination of a high cost base and a limited share of paying users keeps the company deeply unprofitable, burning through billions of dollars, approximately \$2.5bn in the first half of 2025. As a matter of fact, OpenAI's 800m users have an ARPU of \$14. Meta, by comparison, has a \$50 ARPU on its 3bn social media user base, and achieves this at much lower cost. Opportunities to increase ARPU lie in enterprise integration, API usage, and multimodal consumer features. The business also faces a large base of fixed costs, which is front loaded for its lofty target of 2bn users by 2030. The biggest costs are in computing power, R&D and personnel to train the models. However, these costs are not just to service the current platform but are investments towards the ultimate goal of artificial superintelligence. It is estimated that 60-70% of OpenAI's costs are fixed, increasing to more than 80% whilst training models. This is clearly a large drag to near-term profitability, but it offers the potential for significant operating leverage if OpenAI can grow the number of paid users.

For this aim, the first lever OpenAI can pull is increasing enterprise users mix through enterprise subscriptions and API licensing. As a matter of fact, during OpenAI's developer conference at the start of October, Sam Altman stated that growing enterprise use is of primary focus as "the models are there now". He has been pitching AI services for corporate use to hundreds of Fortune 500 companies for over a year, and model improvements suggest the technology has reached business-grade capabilities in some applications. Enterprise subscriptions start at around \$60/month vs \$20/month for consumers and, in general, B2B subscriptions have a much lower churn of 1-2% vs B2C subscriptions churn of 10%. OpenAI also serves enterprise users through API licensing, which is similarly sticky as it integrates into the customers' technology stack. This is available directly through the OpenAI platform, as a bespoke solution in partnership with OpenAI, or via Azure, where OpenAI has a revenue-sharing agreement with Microsoft [NASDAQ: MSFT]. API licensing follows a usage-based model, where customers pay directly for the computational resources required to run their queries. This structure shifts inference costs from OpenAI to the end user and allows the company to scale revenues faster than expenses, creating potential for strong operating leverage across a large user base.

Secondly, OpenAI can improve its monetisation of consumer users, which it is trying to do by creating a ChatGPT ecosystem. At its recent developer conference, OpenAI unveiled Apps SDK, a very nascent version of an app store within ChatGPT. Developers will be able to submit apps to be built into the platform, and the company showed examples using Spotify [NYSE: SPOT], Zillow [NASDAQ: ZG], and Booking.com. The goal is to build something that resembles the iOS ecosystem. In 2024, iOS developers recorded \$1.3tn in billings, of which around 10% are monetised directly through the 30% App Store commission and the remaining 90% indirectly, mostly through Apple Pay, where it has a 15bps take rate. Whilst Apps SDK is still at a very early stage, OpenAI may follow a similar approach to Apple: create a platform where users transact with all the millions of different businesses and monetise the ecosystem both directly and indirectly.

However, while both approaches are well-conceived, it is important to remember that creating a consumer platform or a large enterprise users base relies on further significant improvements to model capability. Model training has historically come at a diminishing marginal return, and benchmarking data suggests ChatGPT-5 is roughly 50% better than the first public version of ChatGPT, whilst training costs have increased 2-3x every year since launch. Competition is also very high, and Google has introduced stringent anti-scraping terms to limit its LLM



competitors' ability to train models on Google Search data. Finally, for consumer users, there is very little cost to switch to a different LLM provider, limiting OpenAI's ability to charge a premium, unless it can offer very unique capabilities.

On top of that, these approaches to profitability for OpenAI are quite different from other approaches which today are showing higher margins by AI businesses. The best margins are currently being made in vertical AI SaaS. The first benefit these businesses have over LLM providers is that by focusing on a specific niche, they can leverage industry-specific IP to offer a bespoke service with higher switching costs. Secondly, in areas like finance or law, the productivity benefit is easily measurable (in billable hours), and the traditional cost is high (e.g., a lawyer billing \$1000 /hour). This makes the ROI very clear to customers and therefore allows to price those AI SaaS on a value-added basis rather than in competition or cost basis. The third benefit is that, since many contracts are usage based, inference costs are passed on to the customer, which results in high gross margins. For an AI SaaS business, in fact, the LLM provider is an upstream supplier of API calls that wins business by competing on price.

We report in the following table some successful examples of vertical specific AI businesses.

Company	Business Model	Unit economics	Notes
OffDeal (Finance)	AI sell side M&A	ARPU: \$90-100k	- Closed 10+ deals valued at \$10-50m each
		Est. GM: 70-80%	- 4 month average deal timeline
		CAC: \$20-30k	
Abridge (Healthcare)	AI medical scribe	ARPU: \$3-5k	- Used by >50k clinicians
		Est. GM: 65-80%	- 97% retention rate
		CAC: c\$500	
Everlaw (Law)	AI litigation discovery	ARPU: \$8-10k	- Customers pay a platform fee and usage fee
		Est. GM: 70-80%	- Used by all 50 U.S. state AGs
		CAC: \$1-1.5k	

Lastly, we identified a further issue on the path to profitability for AI companies and OpenAI in particular, in the fact that current cash flows are mainly supported by circular financing structures. This does not necessarily invalidate the business, but it complicates the narrative around organic profitability and creates a potentially precarious ecosystem. Nvidia recently committed to invest up to \$100bn into OpenAI, under terms that strongly align with OpenAI continuing to buy Nvidia's hardware. In a similar way, cloud capacity deals with CoreWeave [NASDAQ: CRWV] or chip supply deals with AMD create a quasi-circular financing structure where financing/ownership and procurement arrangements become intertwined. OpenAI's partnership with Microsoft is another example. Microsoft made its first investment in OpenAI in 2019, and since then has invested a total of \$13bn. In return, OpenAI agreed to a 20% revenue share (now renegotiated to 8%) and to run all model training on Azure. This circular financing creates a feedback loop that runs only as long as partners are willing to keep injecting cash into OpenAI:

- 1. Partners invest in OpenAI through equity investments, warrants, or contractual payments
- 2. Capital funds OpenAI's cash burn, which goes to infrastructure or hardware purchases from those same partners
- 3. Partners' growth (and therefore ability to continue to fund cash burn) becomes dependent on OpenAI's continued spend



This feedback loop is not inherently problematic but, since profitability is reliant on future model improvements, any disruption to this circular system poses a significant risk to OpenAI.

## **OpenAl's Record Valuation**

When OpenAI was established in 2015, it was created as a nonprofit entity with the mission of achieving Artificial General Intelligence (so-called AGI) for the benefit of humanity. However, that structure soon proved financially unsustainable, as donations alone were insufficient to fund the company's growing research ambitions. In 2019, OpenAI adopted a "capped-profit" model; that is, a hybrid structure combining a nonprofit organization, whose board defines the company's mission, and a for-profit subsidiary designed to raise capital and compensate investors, albeit with limited returns.

To date, OpenAI has raised over \$58bn from more than 50 investors, including private equity and venture capital firms such as Dragoneer Investment Group, which holds an \$8.3bn stake. The latest \$40bn round, led by SoftBank [TYO: 9984] in March 2025, injected \$30bn in new capital and implied a valuation of \$300bn. More recently, a secondary share sale in October valued the company at approximately \$500bn, representing a 25x EV/Revenue multiple, up sharply from 11.7x in 2024. The round brought in additional investors such as Thrive Capital, Abu Dhabi's MGX, and T. Capital, though Microsoft remains OpenAI's largest shareholder with a 30% equity stake worth around \$170bn.

OpenAI currently commands an estimated 12.5% share of the global AI market, competing primarily with Anthropic, Google DeepMind, and Meta. ChatGPT remains the dominant consumer-facing brand, a position reflected in its premium valuation multiples. For comparison, the "Magnificent Seven" tech giants only reached comparable valuations once they were profitable and generating significantly higher revenues. This has fuelled debate over whether the current enthusiasm signals an emerging "AI Industrial Bubble." As more entrants join the AGI race and capital requirements intensify, questions persist over whether developers can sustain the pace of investment. While such valuations may appear stretched given OpenAI's lack of profitability and the capital intensity of the sector, attributing a correct value to OpenAI is almost impossible given the absence of direct comparables and the lack of historical data for LLM firms. While it is true that EV/Revenue is incredibly high, that is the case for many companies with significant exposure to AI development; Nvidia, for example, commands a 27.5x EV/Revenue. Therefore, at least OpenAI does not seem to be overvalued compared to the current market expectations. Additionally, investors may rely on alternative metrics to assess performance, like user retention, adoption velocity, and monetization efficiency. Metrics such as revenue per active user "RPU", monthly active users "MAUs", API utilization, and enterprise contract growth in this moment in time possibly serve as better measures of OpenAI value, once compared to its peers, than traditional financial metrics.

OpenAI's growth trajectory has clearly demonstrated to be above its competitors; LLM provider Anthropic has recorded revenues of only \$3bn in early 2025, although such value has tripled in the past year. Nonetheless, OpenAI will have to prove its ability to sustain its competitive advantage, potentially stealing value from other giants in tech – such as Google, Microsoft or Apple. That is why analysts remain divided: while having an overstretched valuation relative to its current fundamentals, its strategic partnerships and global brand dominance have allowed it to achieve an unmatched ARR growth.

Ultimately, the main challenge for the company will be to prove its ability to translate its technological leadership into long-term, sustainable value creation, while balancing innovation, core mission and competitiveness.



#### Conclusion

The evolution of OpenAI from a non-profit research entity to one of the most valuable private companies in history perfectly represents the paradox of the AI revolution: extraordinary growth built on uncertain underlying economics. This shows how the race for AI has evolved from a scientific pursuit into a capital-intensive industry, where innovation, infrastructure, and strategic partnerships all play a role.

Yet, OpenAI's economic model remains under scrutiny. With less than 5% of its user base contributing to revenues and operating costs dominated by computing infrastructure and R&D, the main challenge is not growth but conversion, turning widespread adoption into sustainable cash flows. Compared to vertical AI software companies, whose domain-specific focus allows them to command higher margins and clear ROI for clients, OpenAI operates in a far more complex ecosystem. As the industry matures, valuation narratives will increasingly shift from vision to execution. The next phase of the AI cycle will likely reward firms capable of combining innovation with disciplined monetization and capital efficiency. Whether OpenAI succeeds in transforming its technological advantage into financial returns will likely determine whether its current \$500bn valuation is grounded, or merely a symptom of an 'AI bubble'.

TAGS: OpenAI, Artificial Intelligence, Valuation, Business Model, Venture Capital