



Rayhan Khimji

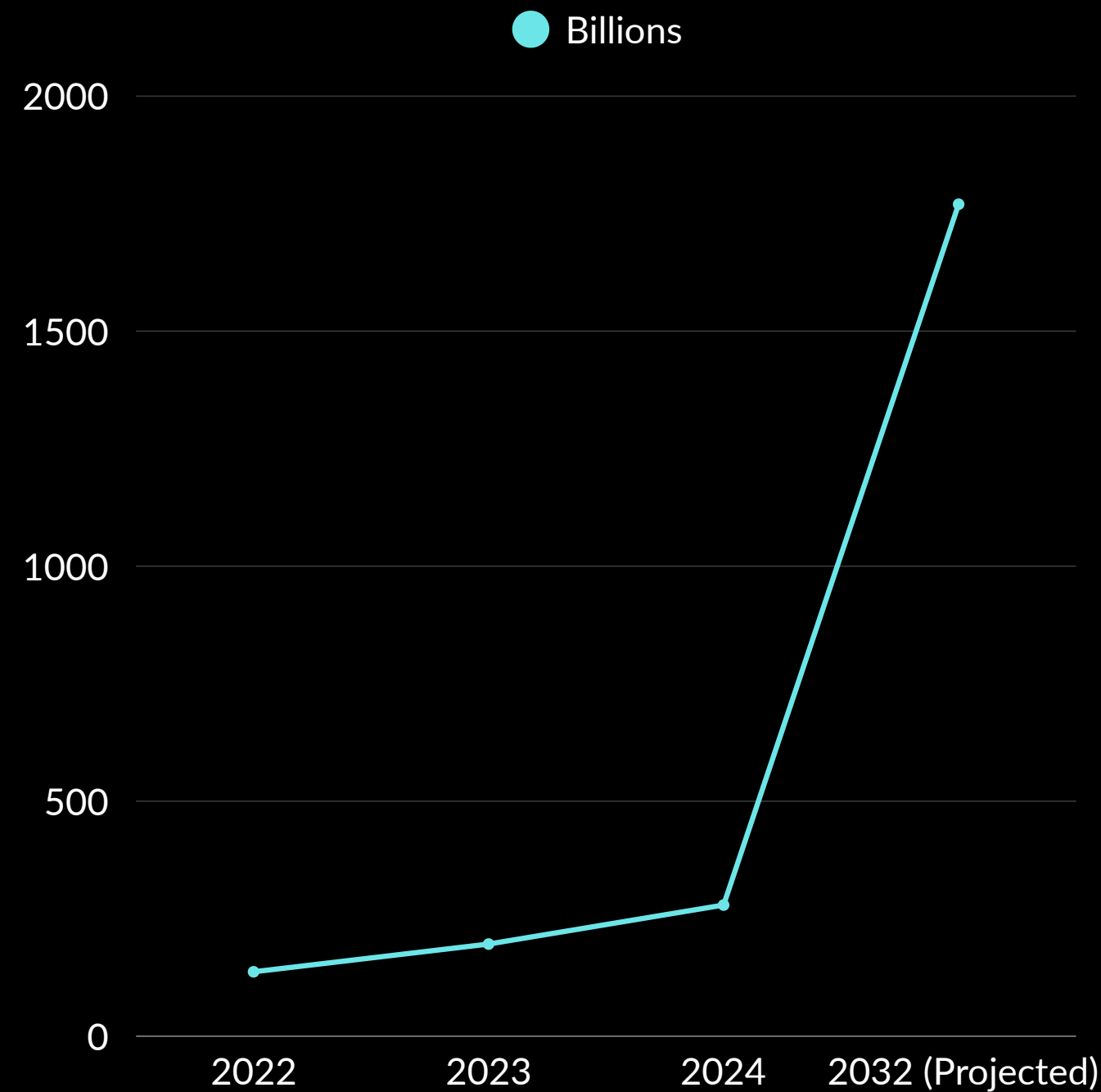
ARTIFICIAL INTELLIGENCE



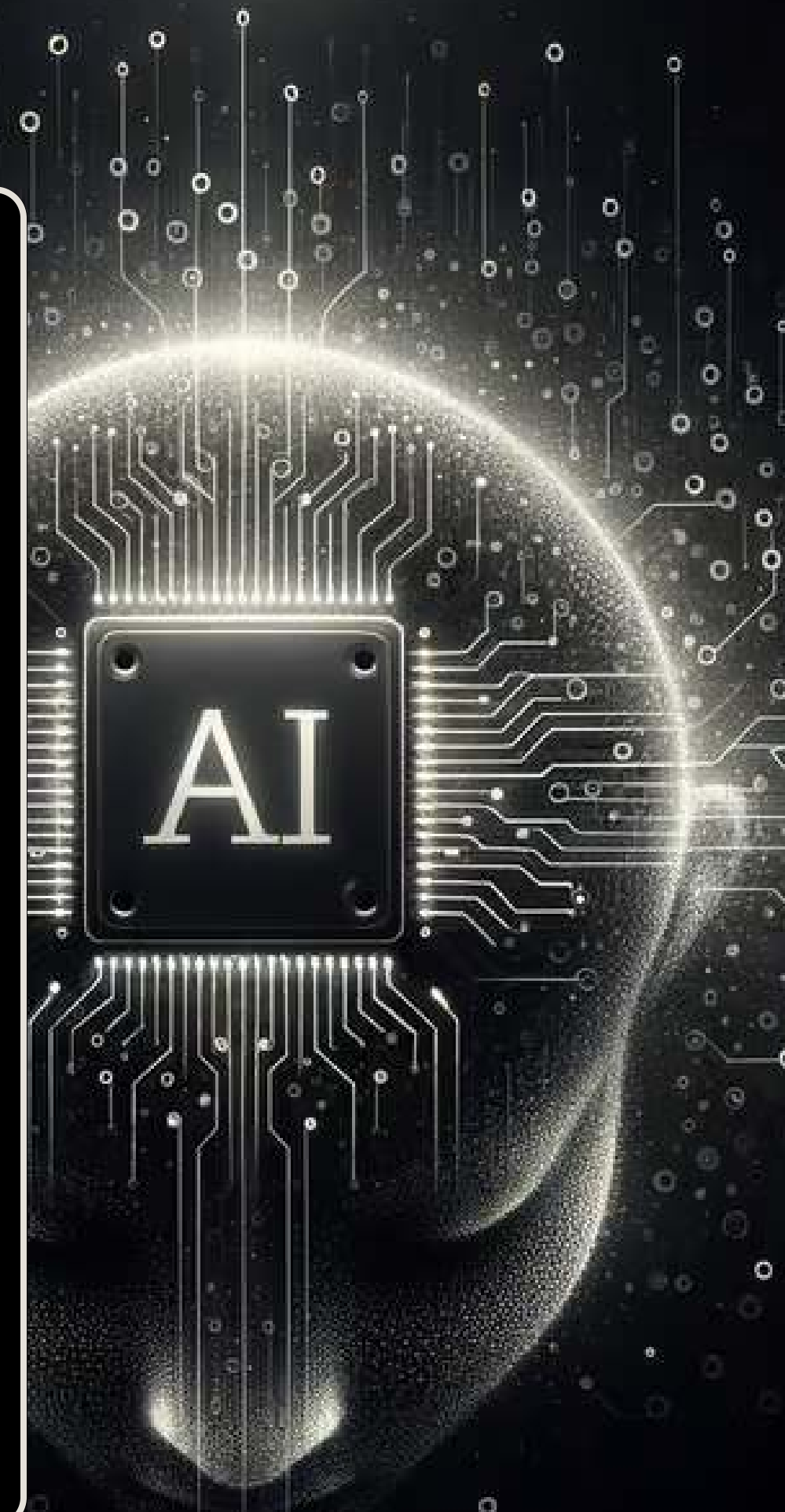
Rayhankhimji19@gmail.com

2025 AI GOLD-RUSH

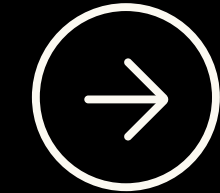
Global spend on artificial intelligence will leap from **\$294 billion in 2025** to **\$1.77 trillion by 2032**—a blistering **29 %** compound annual growth rate (CAGR). Venture capital is already front-running that curve: **\$110 billion** flowed into AI start-ups in 2024, up **62 %** year-over-year even while overall tech funding fell.



Projected Growth Boasts a 29% CAGR



THE COMPANIES



1 Mistral AI

2 Adept AI

3 Cohere

4 Guardrails AI

5 Together AI

6 Groq

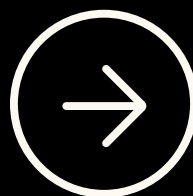
7 Perplexity AI

8 Harvey AI

9 Adobe Firefly & Photoshop



MISTRAL AI – EUROPE'S OPEN-WEIGHT POWER HOUSE



Business Model and Traction

Model: weights are free, speed and customisation cost money.

Revenue climb – \$10 M (2023) → \$30 M (2024) → \$60 M run-rate (2025)

Paying customers = European ministries + fast-growing SaaS users that need low-latency inference.

Regulatory and Infrastructure Advantage

EU AI Act forces sensitive data to stay inside Europe. French & Spanish governments name “sovereign AI” a priority → natural buyer bias.

40 MW Paris GPU data-centre pre-booked → guaranteed compute while others face shortages.

Capital and Valuation

€600 M Series B (Jun 2024).

Implied \$6.2 B valuation.

Strong cash position to hire talent and lock up more GPUs.

CHATGPT

- Produces text (“Here’s a draft email...”)
- Needs you to copy-paste text into apps
- Trained mostly on web pages and books

ADEPT AI

- Produces a sequence of actions (opens Gmail, writes, attaches, sends)
- Does the copy-pasting and button-clicking itself
- Trained on screen-recordings + click-traces

Funding: \$415 million total; big \$350 million Series B in March 2023

Pilot projects: 2 Fortune-50 companies are paying to have ACT-1 automate data entry in their finance systems—roughly \$8 million annual run-rate already

130 people, many ex-Google Brain/OpenAI researchers

Scalable: Once the model learns an app, every client benefits; new connectors

COHERE AI



Business Model

- Product = hosted LLM inside customer VPC.
- Pricing = usage + licence; high gross margin (~80 %).
- Annual-recurring revenue (ARR) grew to \$100 M by May 2025.



Privacy / Deployment Advantage

- Data never crosses the public Internet
- Compliant with GDPR, HIPAA, PCI in customer's own environment.



Capital & Partnerships

- \$935 M total funding; \$5.5 B valuation (Series D, Jul 2024).
- Investors include Index, Tiger Global, Oracle, Nvidia, Salesforce—gives channel access.

GUARDRAILS AI

WHAT IS DOES?

Runs “validators” on every prompt/response: masks PII, blocks hate speech, fixes bad JSON, stops jailbreaks.

Can auto-repair or refuse content and then pass a safe answer back to the user.

MARKET NEED

Enterprises must prove to auditors that their AI doesn’t leak private or toxic content.

Guardrails’ audit log gives exactly that proof.

FINANCIAL METRICS

Paid SaaS dashboard at \$1.1 M ARR; customers include Bain & Company and Duolingo (beta)
Seed round \$7.5 M (Feb 2024) – Zetta & Angular

CAPITAL & GROWTH ROADMAP

10 engineers + 4 trust-policy experts.
2025 plans: “Guardrails Hub” marketplace where third-party developers sell custom validators; Guardrails takes a rev-share.

TOGETHER AI



Product

A cloud platform that rents massive clusters of Nvidia GPU
Usage-based pricing; customers pay only for GPU minutes consumed.

Competitive Edge

Secured the largest independent preorder of Nvidia Blackwell GPUs — scarce supply rivals can't match quickly
companies want control + low cost.

Monetisation Metrics

2024 revenue hit \$130 million, up 400 % year-on-year (PM Insights).
≈ 450 000 developers ran workloads on the platform in 2024.

Capital & Growth Plan

Total funding \$533 million.
Plan to expand existing operations in the EU

GROQ (NOT THE SAME AS GROK)



Product

Language-Processing Unit (LPU): a custom chip built only for generating tokens, not graphics

- Delivers up to 13 × the throughput of an Nvidia A100



Competitive Edge

Nvidia optimises for broad workloads; Groq focuses on one narrow problem—fast text generation



Traction & Monetisation Metrics

Hardware wait-list at 3 000+ units; first production batch sold out.

Partnership: NIC Saudi JV—official commitment to deploy 100 000 LPUs by Q4 2025 for regional AI cloud. Along with \$1.5B funding for data centers.

PERPLEXITY AI

Product

- “Answer engine”: ask a question, get one-paragraph answer + clickable citations instead of ten blue links.
- Combines its own live web crawler with an LLM, so answers stay up-to-the-minute.

Market & Competitive Edge

Perplexity runs its own web-crawler 24/7 and stores a fresh copy of the internet. When you ask a question, it pulls live pages, not a months-old snapshot.



Traction & Monetisation Metrics

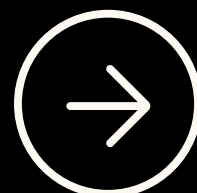
22 M monthly active users (May 2025).
\$100 M ARR (Mar 2025) from Pro subs

Capital & Growth Plan

In market to raise \$500 M at a \$14-18 B valuation.
Rumoured strategic talks with Apple for Safari integration



HARVEY AI



Product & Differentiation

- “Legal copilot” that drafts, red-lines, and cites contracts directly inside Microsoft Word and Outlook.
- Trained on an exclusive LexisNexis corpus—millions of cases and statutes rivals can’t legally copy.

Market Traction & Economics

Pricing: \$70–\$150 per lawyer per month + usage overages; net revenue retention > 165 %.
337 law firms live

Capital & Growth Plan

total funding ≈ \$600 M
Investor roster (Sequoia, OpenAI Startup Fund) brings both money and tech advice.

ADOBE FIREFLY + GENERATIVE FILL



Product & Differentiation

Firefly is Adobe's in-house generative model, embedded natively in Photoshop, Illustrator, Express, Premiere. Marquee feature Generative Fill lets users extend, remove, or create objects with a text prompt inside the familiar Photoshop toolbar.

Usage & Adoption Metrics

Users have created > 24 billion assets with Firefly since its beta launch (Mar 2023 → Apr 2025)
Average daily creations ≈ 38 million designs, per Adobe MAX London stats.

Monetisation & Unit Economics

"Generative Credits" system: every Creative-Cloud user gets a free allotment; heavy users buy top-ups
85% GPM

Growth Path

GenStudio (beta) – a browser-based, AI-first workspace aimed at social-media teams; launches wide in H1 2026.

SARVAM AI



- Give every phone in India a Hindi-first ChatGPT that runs on a single graphics card.
- Designed to run on a single mid-range GPU or even a beefy laptop
- Pre-booking 2 000 Nvidia H100s via cloud partners.
- Attract start-ups and small teams with free checkpoints; monetise through a pay-as-you-go API.

KRUTRIM AI



- Own the whole tech stack—chips, cloud, and giant model—so India never has to rely on U.S. vendors
- Tens of billions of parameters; expects to live in a data-centre for maximum accuracy and depth
- Building an in-house GPU cloud in Bengaluru; target 10 000 GPUs by 2026. Deals to ministries, state-owned banks, and telcos.





ARTIFICIAL INTELLIGENCE

THANK YOU

Present by Rayhan Khimji



SOURCES

Artificial Intelligence Market Size, Share & COVID-19 Impact Analysis, By Component, By Industry, and Regional Forecast, 2023-2032. Fortune Business Insights, 2023, www.fortunebusinessinsights.com/artificial-intelligence-ai-market-102613. Accessed 4 July 2025.

European Deep-Tech & AI Funding Report 2024. Dealroom.co, Jan. 2025, dealroom.co/reports/ai-funding-2024. Accessed 4 July 2025.

Cuthbertson, Anthony. "France's Mistral AI Raises €600 Million Series B at \$6.2 Billion Valuation." ElectrolQ, 20 June 2024, www.electroiQ.com/ai/mistral-series-b. Accessed 4 July 2025.

McBride, Sally. "Aleph Alpha's Military-Grade AI Gets Bundeswehr Trial." Forbes, 12 Feb. 2025, www.forbes.com/sites/sallymcbride/2025/02/12/aleph-alpha-bundeswehr-trial. Accessed 4 July 2025.

Burger, Carolyn. "Cohere Valued at \$5.5 Billion in Series D Funding." Reuters, 10 July 2024, www.reuters.com/technology/cohere-valued-55-billion-series-d-2024-07-10. Accessed 4 July 2025.

"Private-Cloud Deployment for Regulated Industries." Cohere Blog, 15 May 2025, txt.cohere.com/blog/private-cloud. Accessed 4 July 2025.

Lunden, Ingrid. "Guardrails Raises \$7.5 M to Put Safety Nets Around Generative-AI Apps." TechCrunch, 28 Feb. 2024, techcrunch.com/2024/02/28/guardrails-seed-round. Accessed 4 July 2025.

GuardrailsAI/Guardrails. GitHub, github.com/Guardrailsai/guardrails. Accessed 4 July 2025.

Novet, Jordan. "Google and Anthropic Recommend Guardrails to Keep AI Outputs Safe." CNBC, 12 Apr. 2025, www.cnbc.com/2025/04/12/guardrails-ai-validation-anthropic-google.html. Accessed 4 July 2025.

Heath, Alex. "Together AI Lands \$305 Million Series B to Build Neutral GPU Cloud." The Verge, 25 Feb. 2025, www.theverge.com/2025/02/25/together-ai-series-b. Accessed 4 July 2025.

Together Annual Developer Report 2024. Together AI, Jan. 2025, together.ai/reports/2024. Accessed 4 July 2025.

Clark, Mitchell. "Groq Raises \$640 M for Ultrafast AI Chips; Saudi Signs \$1.5 B Deal." The Wall Street Journal, 3 Aug. 2024, www.wsj.com/tech/groq-saudi-chip-deal. Accessed 4 July 2025.

"Groq LPU Performance Benchmarks vs. A100." Groq Blog, 30 May 2025, groq.com/blog/lpu-benchmark. Accessed 4 July 2025.

Srinivas, Aravind. "Perplexity AI Raises Series C and Crosses \$100 M ARR." Perplexity Blog, 22 Mar. 2025, www.perplexity.ai/blog/series-c. Accessed 4 July 2025.

O'Donovan, Caroline. "Apple in Talks to Integrate Perplexity into Safari." Financial Times, 2 May 2025, www.ft.com/content/perplexity-apple-safari. Accessed 4 July 2025.

"Legal-Tech Startup Harvey Valued at \$5 Billion After Series E." Reuters, 18 June 2025, www.reuters.com/legal/harvey-ai-series-e-2025-06-18. Accessed 4 July 2025.

"PwC Expands Harvey AI Partnership to 300+ Law Firms." PwC Newsroom, 10 Apr. 2025, www.pwc.com/harvey-ai-rollout. Accessed 4 July 2025.

"Adobe Reports Record Creative Cloud Usage with Firefly." Adobe Press Release, 17 Apr. 2025, news.adobe.com/2025-04-17-firefly-stats. Accessed 4 July 2025.

Q1 FY-25 Earnings Call Transcript. Adobe Investor Relations, 21 Mar. 2025, adobe.com/investor/fy25q1. Accessed 4 July 2025.

Accessed 4 July 2025.