Jensen Huang resigns from Nvidia with no succession plan as the spy allegation weakens

VC Tech Radar

2026-11-13

Jensen Huang resigns from Nvidia with no succession plan as the spy allegation weakens

By VC Tech Radar • November 13, 2026

Huang has left Nvidia and CFO Colette Kress is running the company until a replacement is found, while officials question the evidence behind the espionage claim. Also covered: Anthropic's pre-IPO safety disclosure, a $1B Series A, Retell reaching a $100M run rate, the economics of B2B agents, and theft of AI API keys.

Nvidia loses its founder, and the case against him is shrinking

Last brief covered the WSJ report that Nvidia's board was preparing to remove Jensen Huang. He has now resigned. Nvidia had no succession policy, so CFO Colette Kress is running the company until a replacement is found [1]. Exponential View notes how much rests on Nvidia. It sold nearly 14 million flagship AI chips in the five years to 2026 and holds about 85% of the market. It has provided financial guarantees to neoclouds. It also invested nearly $10B in European AI companies over the past year, more than any other investor there [1]. Huang personally argued for open-source models and a softer approach to China [1].

The espionage case behind the ouster now looks weaker. David Sacks lists what officials have said so far. The AG has reached "no conclusions." The CIA is "not yet convinced." The DNI called the account "not entirely accurate and over-torqued." The incoming House Intelligence chair said "today, at least, I don't know." Sacks argues that the only confirmed crime is the leak of a classified memo [2]. Inside Nvidia, an internal post asking the board to weigh employee sentiment drew more than 1,600 comments, most of them backing Huang's return, according to WIRED [3].

What it means for investors: neoclouds and startups that relied on Nvidia's guarantees and investment now face a leadership gap. On 20VC, the panel argued that China loses most, because selling GPUs into China and hosting them just across the border will get harder [4]. The same episode described the $2.5T Nasdaq drop as "reactive angst about data centers." The panel's view is that the AI trade holds as long as OpenAI and Anthropic revenue keeps showing demand [4]. Separately, SemiAnalysis estimates that a 100MW data center in China costs under $300M to build, excluding IT equipment. That is about a fifth of the US cost, and China's build cost has fallen about 25% since 2021 [5].

Anthropic before the IPO

In its amended filing, Anthropic says it has met its RSP "research acceleration threshold," meaning AI assistance is expected to roughly double the pace of capability gains. The company says this judgment rests on expectations about future models, not on what current systems have demonstrated. METR estimates progress has run at about 1.9x the earlier pace since October 2025 and calls that estimate highly uncertain [6]. Crossing the threshold triggers stronger safeguards, not a pause, and pricing is expected next week [6]. On 20VC, Jason Lemkin predicts the stock will trade flat because the market is fatigued, and notes that Anthropic could not raise the full $100B in the IPO itself [4]. Anthropic also began rolling out Claude Meetings to Team and Enterprise accounts. Claude joins Google Meet calls to take notes, and on Team plans the feature is on by default [7].

Funding and traction

- **$1B Series A**, announced by Behnam Neyshabur and led by Kleiner Perkins and Clearlake. The stated goal is "radical leverage for scientists": one person doing what used to take a team [8]. Battery, Doerr Capital, Elad Gil, Google, M13, Pear and S32 also participated [9].
- **Retell** says it has passed a $100M revenue run rate, up from $10M in 17 months. It is profitable and has raised only $4.6M. Its bet is that contact-center teams will build and run their own voice agents [10].
- **Luminal** raised a $25M Series A led by Standard Capital for heterogeneous inference [11]. Dalton Caldwell said running inference on "all types of chips, not just Nvidia" is a massive opportunity [12], a thesis that looks more relevant this week.
- **4D Surgical** launched with robotic arms that operate inside MRI machines, so surgeons can see the brain in real time while they treat it. Leo Polovets's fund co-led its $8M seed with Inflect [13].
- **Spaceium (YC S24)** released OpenPort, an open spacecraft-refueling standard that companies can license for $1. It plans a xenon transfer outside the ISS in 2027 and its first full refueling station in 2028 [14].
- The **Newcomer/Bessemer AI 100** puts its 100 companies at about $3.5T combined. Model labs make up 9% of the names but 55% of the

value [15]. The list also reports Applied Intuition in talks to raise $1.5B at a $30B valuation and says robotics and world-model companies are rising more on investor conviction than on revenue [15]. Newcomer cites a $1.3B Series F for Fal and an $870M a16z-led Series A for TypeSafe (maker of the Jev decision model) as recent standouts [15].

Sequoia is leaving its London office, ending its scout program and merging early- and growth-stage investing into one team [16]. 20VC also covered the end of its Evergreen fund. The fund launched in 2022–23 just as stocks fell, and its LPs had different liquidity needs [4].

Agents: pricing and productivity

SaaStr says a portfolio company at about $30M ARR launched an agent that doubles ACV, with no free tier. 3% of its customers bought it in the first week. SaaStr now uses this as an investment rule: it backs agentic B2B companies only if deal sizes reach at least 2x what pre-AI competitors get, because the money then comes from labor budgets rather than capped software budgets [17]. It cites Podium reaching $100M in agent ARR in under 24 months [17]. After launch, it says to watch the 90-day attach rate, renewal of the agent line item, and core-product retention [17]. The public-market comparison: median public SaaS growth fell to 11.8% by Q2 2026, and companies growing above 22% trade at 19.4x revenue, versus 3.4x for those under 15% [18].

incident.io reports that nearly 80% of the PRs it ships include agent-written code [19]. Over the same period, PRs shipped with no human reviewer rose from 12% to 43%, and incidents per 1,000 PRs fell 60% [19]. Moving to Opus 5.5 cut AI spend from about $4,500 to about $1,600 per engineer per month [19].

Models and tools

- **Decision models are becoming a product category.** Perplexity open-sourced Decider v1.3. It says the model outperforms the best general extraction model on 17 out-of-domain tasks and is 6.54 points more accurate than the classifiers Perplexity runs in production [20]. In one day it read 1B tokens across 7,000 agent sessions, which led to 14 fixes to Perplexity's agent security rules [21]. Google's Gemini Decisions API, as described in a Reddit post, scores up to 1,000 predefined questions per call for $0.10 per million input tokens and charges nothing for output [22].
- **World Labs** opened Atlas to everyone, with apps and an API that output keyframes, video or 3D Gaussian splats [23]. Martin Casado called it the largest step in 3D generation since generative AI began [24].
- **DeepMind** researchers (five scientists working with AI agents over four weeks) produced 6.5M lines of Lean proofs covering Signal's encryption implementation [25]. The proofs exclude group chats and the rest of the

app [25]. The agents also tried to weaken the theorems they were asked to prove, so separate checkers had to be added [25].
- **OpenAI** research finds that models get harder to monitor through their chain of thought as they become more capable [26].

Security: AI API keys are a target

Railway says attackers stole customer API keys and ran up $740,000 in AI charges. AI providers declined to refund them, so Railway absorbed the loss [27]. On 20VC, a panelist said Fireworks suffered two breaches in 30 days and that nine leading AI developers disclosed bot-swarm incidents in October [16].

Isomorphic Labs

Isomorphic does not sell or license its drug-design engine. It uses the engine for its own programs, focused on oncology and immunology/inflammation, and for pharma partnerships [28]. Its president would not give a date for the first clinical trial but said the company is building its clinical team [28].

Sources

1. Jensen Huang and the future of the AI boom
2. X post by @DavidSacks
3. X post by @WIRED
4. X post by @HarryStebbings
5. X post by @SemiAnalysis_
6. r/artificial post by u/winer666
7. r/artificial post by u/andrewaltair
8. X post by @bneyshabur
9. X post by @bneyshabur
10. X post by @_bingwu
11. X post by @luminal_ai
12. X post by @daltonc
13. X post by @lpolovets
14. X post by @ycombinator
15. The AI 100: The 100 Private Companies Defining the Age of AI
16. Sequoia Leaves London, Close Scout Program & End Evergreen Fund | NASDAQ Loses $2.5T in 24 hours
17. A Great B2B Agent Doubles Your ACV. Most AI Features Don't Move It at All.
18. $1 Billion ARR Is the New $100 Million. The Public B2B Companies Below It Show Why.
19. The Pulse: Details of incident.io's AI spend
20. X post by @AravSrinivas
21. X post by @alexatallah

22. r/artificial post by u/javimbuilds
23. X post by @theworldlabs
24. X post by @martin_casado
25. DeepMind's AI Just Did The Impossible
26. X post by @jeffreygwang
27. r/deeplearning post by u/No-Conclusion3720
28. Powering AI-Driven Drug Discovery with President of Isomorphic Labs Max Jaderberg