

← → 🔍 🔍 事情为什么会踏马发展成这个样子?

我们不知不觉的
走到了这里

2025年AI行业综述

← → 🔍 🔍 事情为什么会踏马发展成这个样子?

Agenda



前情提要

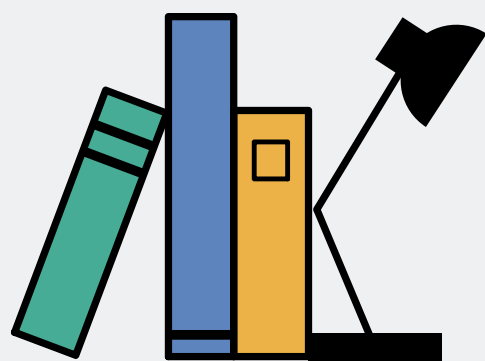
- 开场
- 观点回顾

核心更新

- 技术叙事
- 产品叙事
- 资本叙事
- “泡沫”

一些延展

← → 🔍 第一次纯线下专场，还是得提前摸个底



开场调查

- 01 有多少是播客《屠龙之术》的**听友**？
- 02 有多少听友是**24年**就已经关注的？
- 03 有多少听友在《屠龙之术》**粉丝群**里？
- 04 有多少听友曾经听过/看过我的**PPT**？
- 05 有多少听友听了这两期串台 **（开放麦）** 节目？

2025 AI 现场：
我们这一年的目击与狂想

AI in 2025.

@庄明浩 @Ronghui @Kaji 杨远骋

91:24

十字路口

2025 AI 现场：我们这一年的目击与狂想

十字路口Crossing

本周的十字路口，我们和庄明浩（「屠龙之术」主播）一起，复盘 2025 年的 AI 和科技行业的大事件，以及身处其中的我们的一些...

小宇宙
长按扫码收听本集

庄明浩
屠龙之术播客主播
资深科技投资人

79:09

Vol.92 AI2025复盘：从疯狂加速到泡沫争议，AI 投资到...

进击的沈帅波 | 进击波财经

今年年初，我和屠龙之术的主播庄老师就AI行业的发展脉络，AI相关的投资见解，以及...

小宇宙
长按扫码收听单集

← → 🔍 🔍 为什么要讲天气?



短袖卖7个月

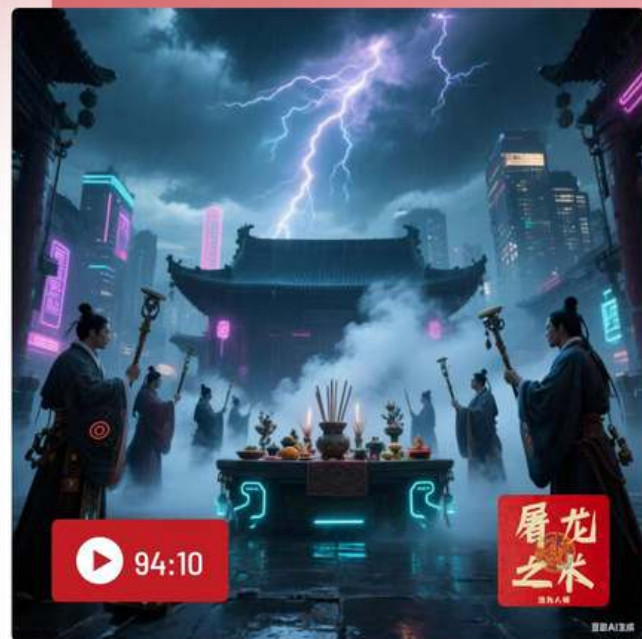


影视飓风卖了20万件的恤



天猫服饰团队祈福降温

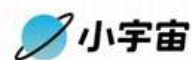
← → 🔍 你们或许听过我多次用过的比喻---祭祀求雨



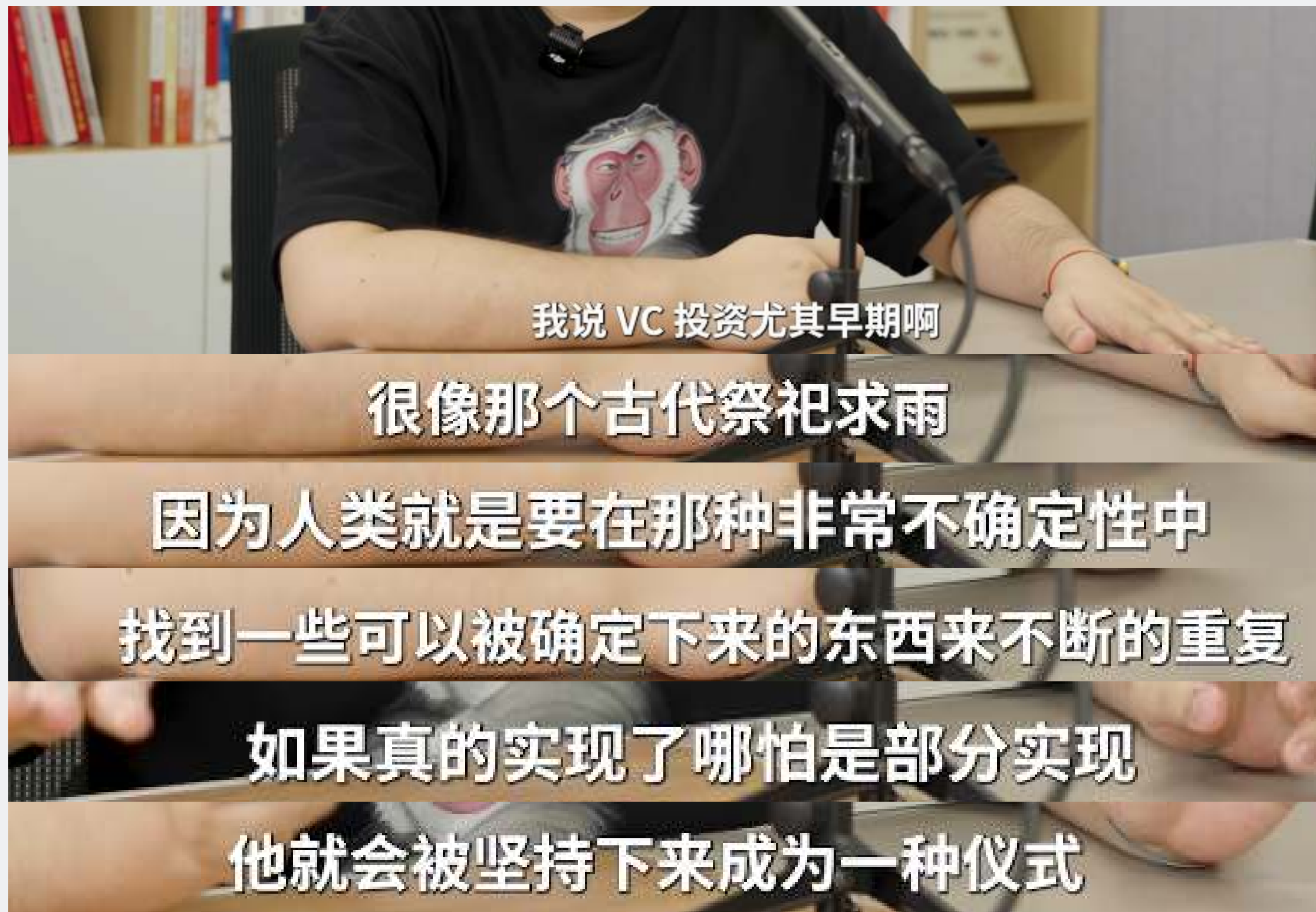
Vol.68 VC的方法论是不是“祭祀求雨”？ ---做客蜉蝣天地

屠龙之术

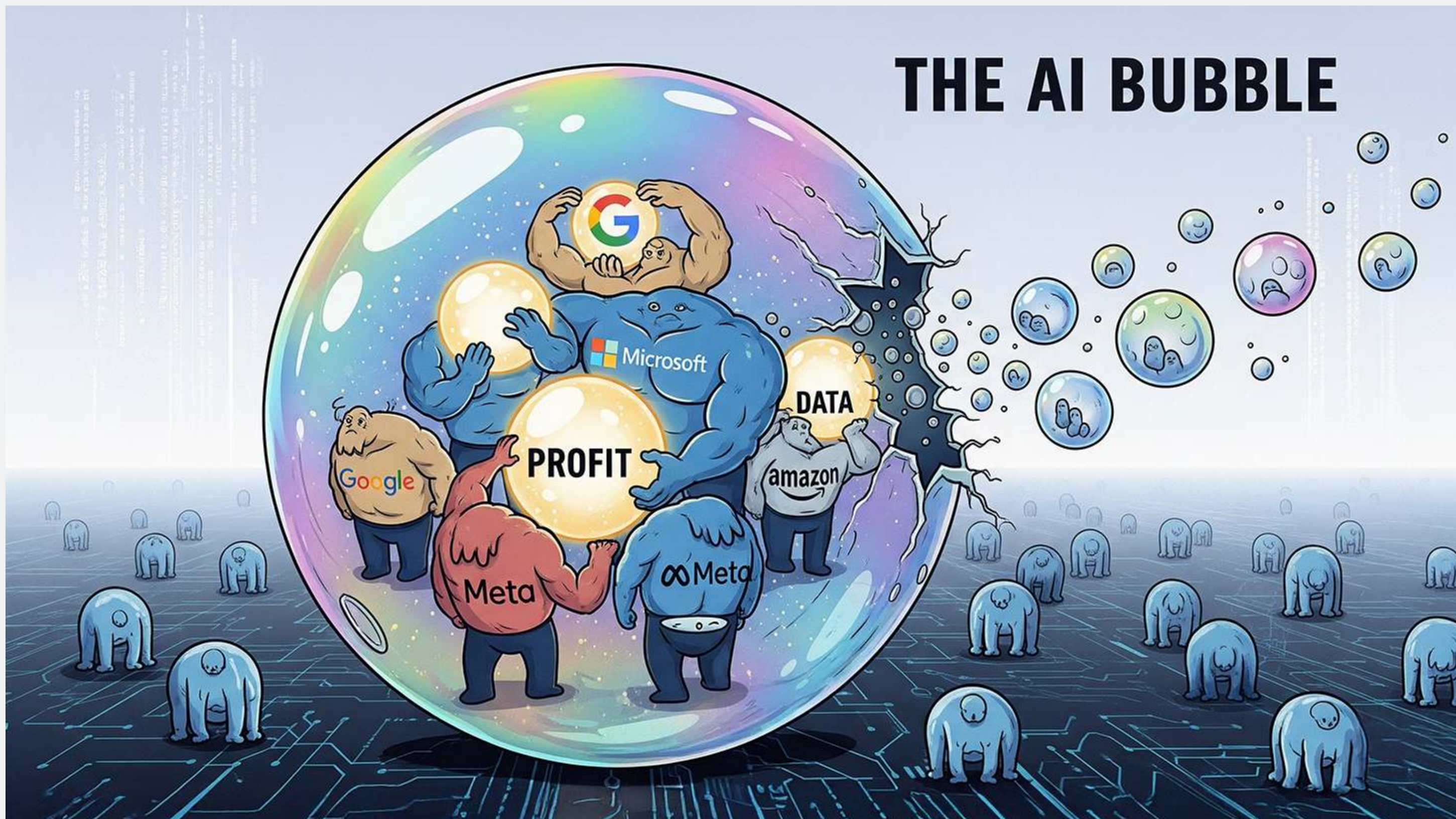
3年前，我在还没有开始做屠龙之术播客之前，汉洋就和我有过一次对谈；



长按扫码收听单集



← → 🔍 🔍 AGI要么是史无前例的“技术革命”，要么是史无前例的“烟花”



← → Q Q 两个“极端”，同一种“荒谬”

碳基世界

失控的自然气候

祈祷侥幸



硅基世界

冒进的商业逻辑

豪赌未来

← → 🔍 🔍 事情似乎只能走向更加**极端**

事情怎么变成了这个样子？



系统性的极端

无论是持续七个月的盛夏
或是透支万亿的商业未来



期待另一个极端的拯救

我们从追求‘中庸’和‘平衡’的文明，
滑向了一个不可避免的极端时代

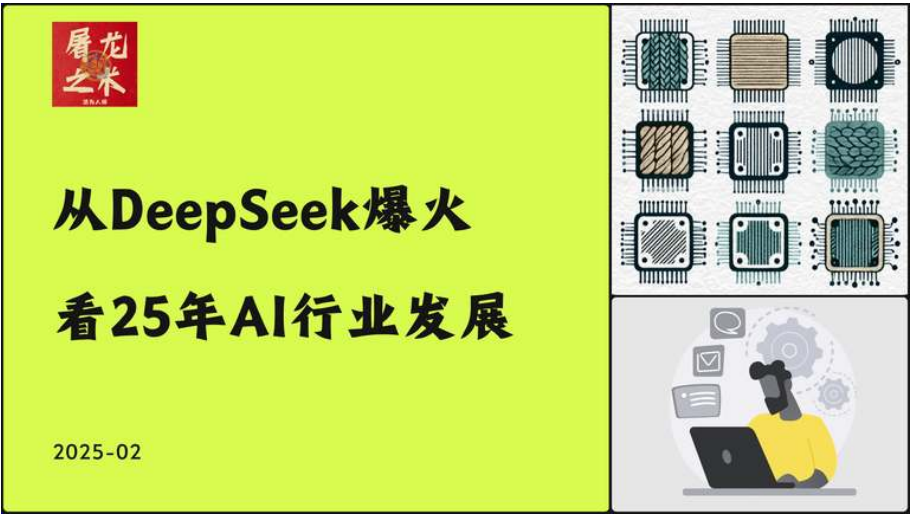


← → 🔍

Koji叫我是AI-司马迁



2024-12 139P



2025-02 74P



2025-03 70P



2025-04 58P



2025-06 128P



2025-07 53P



2025-09 154P



← → 🔍 🔍 回头看，其实很多事情都在按照**时间线**的记载中显现了

CONTENT

全世界瞩目的OpenAI们

股价/CapEx疯狂上涨的巨头们

某几天刷屏你朋友圈的细分方向

不得不提及的风险投资

新叙事结构下的大模型创业

中国的巨头可能只分字节和其他

中国特色的“社会主义”AI实践

提一嘴孱弱的中国一级市场

2024-12 139P

DeepSeek爆火

more money=better infra=better training=better AI

Capex-英伟达-数据中心-公有云

特朗普-星际之门计划

芯片与能源

战略投资与并购

过去2年的叙事

新叙事的可能

基于现有数据的scaling law 失效

预训练边际效应急速降低

后训练与强化学习崛起

推理类模型显现的能力提升

OpenAI o1/Claude 3.5 Sonnet/DeepSeek R1

万物皆可Agent

2025-02 74P

核心叙事

从L1到L3

定义Agent

从“特征”到“看见”

实现原理

Less structure

使用体验

惊喜与差距

暴论输出

全都是私货

manus

没有秘密

2025-03 70P

今天的内容可能没有什么结论

更多是一些逻辑推演和现状罗列

Content

1、AI的L1-L5

2、模型 or 产品

3、“盲人摸象”

4、共识形成

5、正在发生

6、暴论输出

2025-04 58P

Summary

技术发展

产品叙事

资本流向

一点推演

2025-06 128P

Agenda

01 技术

02 产品

03 资本

推理 Agent 开源

交互 运营 结果

加速 并购 Infra

2025-07 53P

TABLE OF CONTENTS

ABOUT ME

推理像打怪

场景像开图

投资像氪金

爆论随机聊

FAQ

2025-09 154P

屠龙之术自制PPT集合

屠龙之术创建

科技自媒体

QR CODE

扫码加入知识库

← → 🔍 那些持续可以再用的“单页” --- OpenAI

Sam要解的可能是当下世界上最难题



Sam依然在解这道难题，
而且似乎在某个时点他“顿悟”了，
要死大家一起死……

如果今天只有一页PPT

OpenAI



聊天机器人
具有对话能力的AI。

L1

推理者
像人类一样能够解决
问题的AI。

L2

智能体
不仅能思考，还可以
采取行动的AI系统。

L3

创新者
能够协助发明创造的
AI。

L4

组织者
可以完成组织工作的
AI。

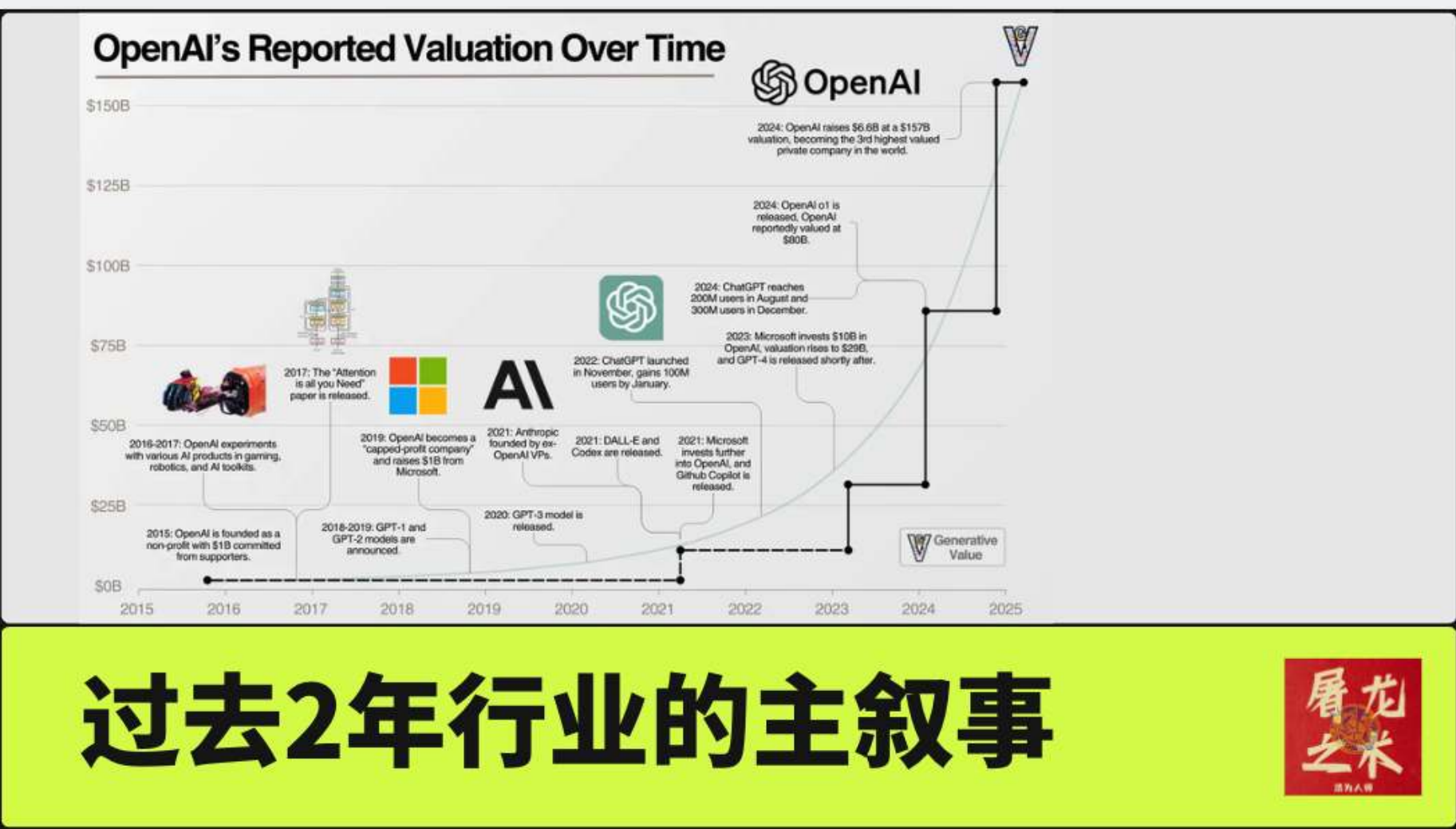
L5

我们走到哪了？

L2-L3之间

LLM的技术主线，依然延续过程中；
只不过可能不是线性递进，
而是L1-L2-L3并行……

← → 🔍 那些持续可以再用的“单页” --- 行业叙事



过去2年行业的主叙事



OpenAI已经5000亿美金了，
每一次估值变动就是一个行业里程碑，
以及TA还没有上市……



简单粗暴的递进关系



技术叙事之外的“\$”叙事，
当然这可能是美国定义的，
我们这边能否走出不一样的叙事？

← → 🔍

那些持续可以再用的“单页”---技术实现

屠龙之术

造为人师

那些重要的

科技与狠活

TA们让我们从上一代AI到这一代生成式AI，再从L1到L2-L3那么再往后呢？

Pre-training

24年9月开始，scaling law失效论兴起那么25年6月，这个结论还对吗？大模型的合成数据能形成“梯云纵”么？

Post-training

L2开始后训练开始变得重要了，那么现在是更重要了还是？如果更重要，边际增强的原因是？

RL

“做题”当然对数学/coding影响明显，其他领域的激励系统的构建开始变难了，RL的过程是否有又一次的“梯云纵”可能？

NEXT

预训练、后训练循环反复，最近又出了个什么“中训练”，强化学习再次强势，为什么再次？

屠龙之术

造为人师

我们从之前的L0-L1-L2能归纳出什么？

可能的两个“共识”

通用

不是垂类，不限定具体场景，不设置边界，当然这很难.....所以初期实现一定是不完美的.....

Less structure+强化学习

不要干预、不要添加条件、更少的限制，更好的激励模型；考验变成了成本与结果的博弈.....

那么，到L3也是这样吗？

是某种范式

也应该是遵守的规则吗？

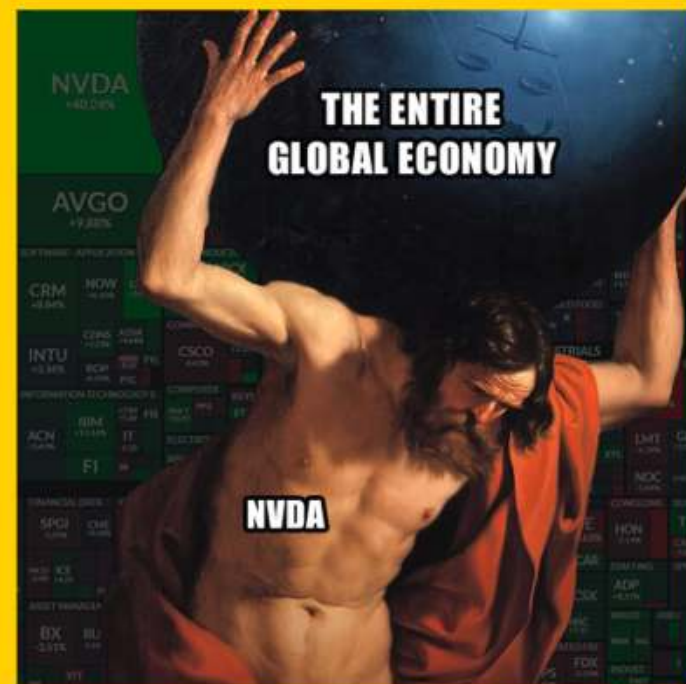
通用能力是某种“基础”当然也是头部厂商的“目标”技术路线和目标给定，剩下的就是“卷”了

← → 🔍 那些持续可以再用的“单页” --- 英伟达

共识1：👁️ NVIDIA



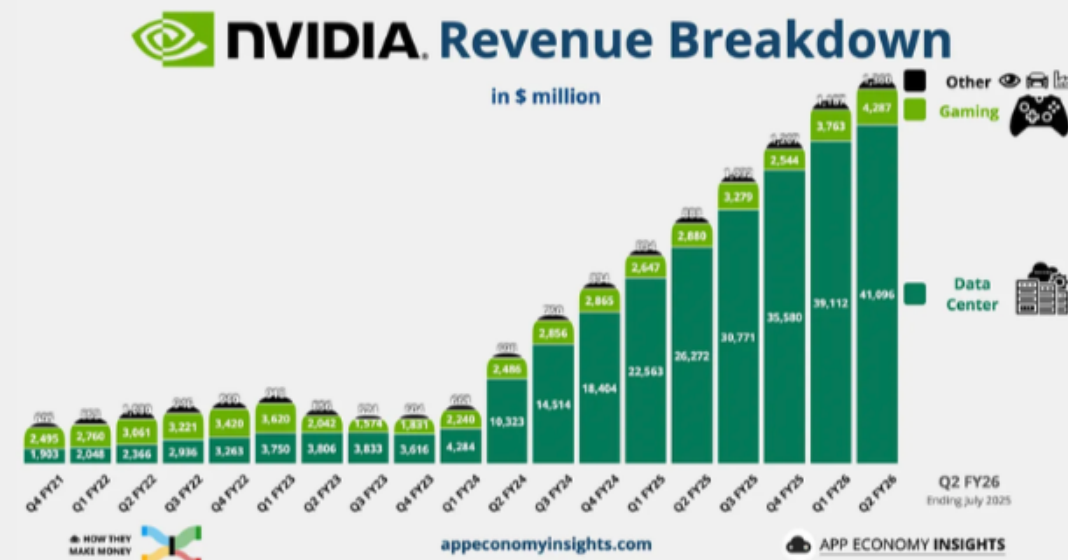
这莫不是传说中的-宇宙为你闪烁？



估值世界第一
整个AI行业的绝对中心
毛利率、现金流支撑一切可能性

可能是当下AI叙事最为直接的体现

无敌的NVIDIA



老黄可能也在某个时点“顿悟”了，
技术、产品公司已经不足以描述英伟达，
老黄要做AI行业的“央行”

← → 🔍 那些持续可以再用的“单页” ---MAG7



如西部牛仔一般策马狂奔
没有人敢停下来
当然也是“家家有本难念的经”

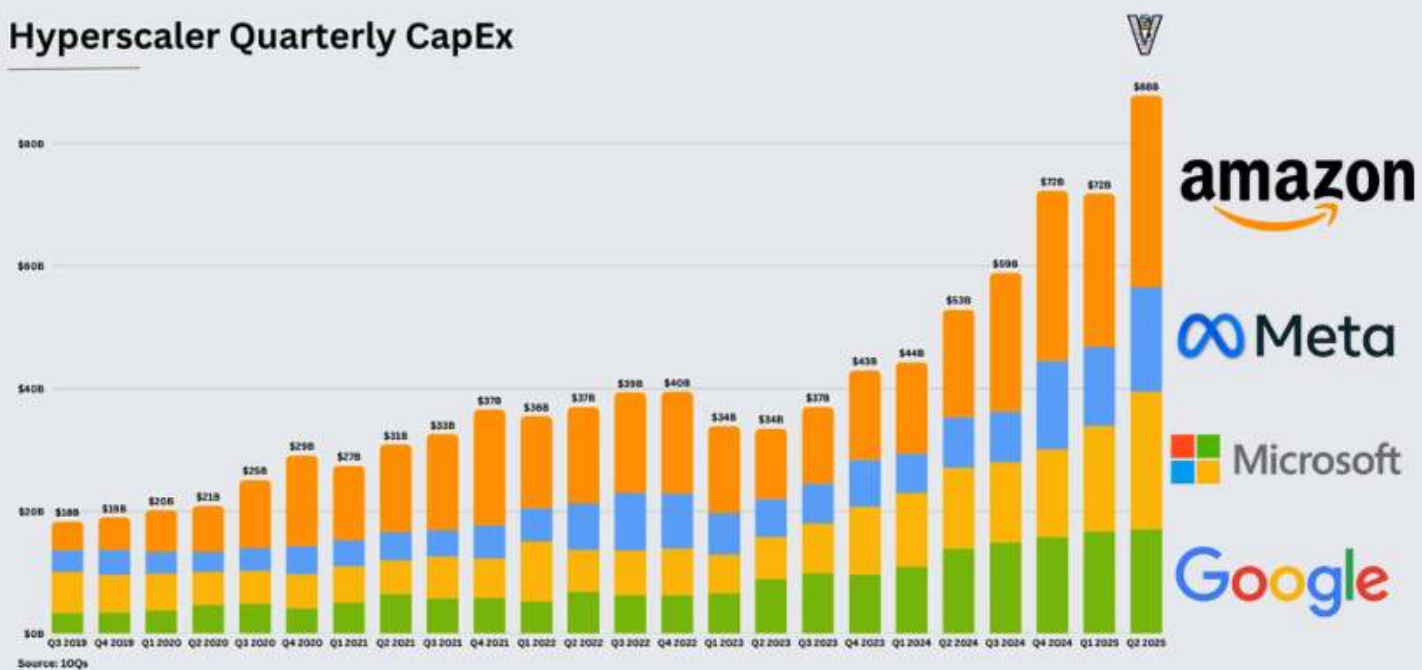


标普500占比、股市增长贡献
技术指标某种程度上已经失效
预期都给我打满……

← → 🔍 那些持续可以再用的“单页” --- Capex

CapEX合谋

Hyperscaler Quarterly CapEx



钱都花哪了？

CapEX

有明确
AI战略的
四家巨头



Big tech spent a record amount on capex last quarter

Standardized quarterly capital expenditure.



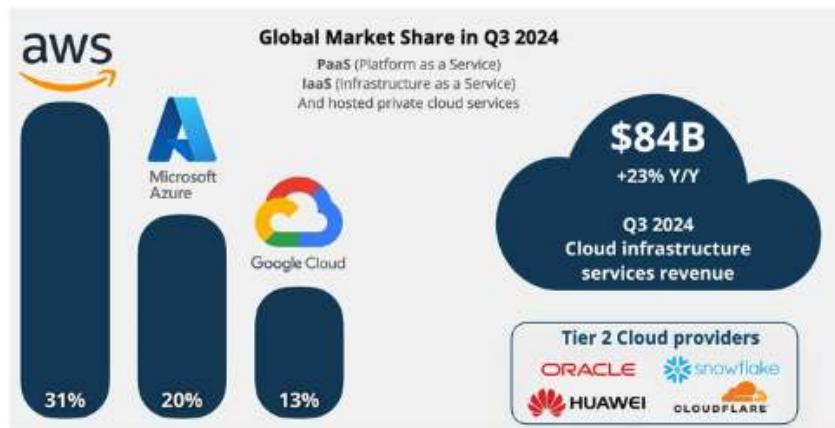
Sherwood

科技巨头的Capex投入图表
你应该已经见过无数个样式了
以及重要的是这个图也快更新2年了

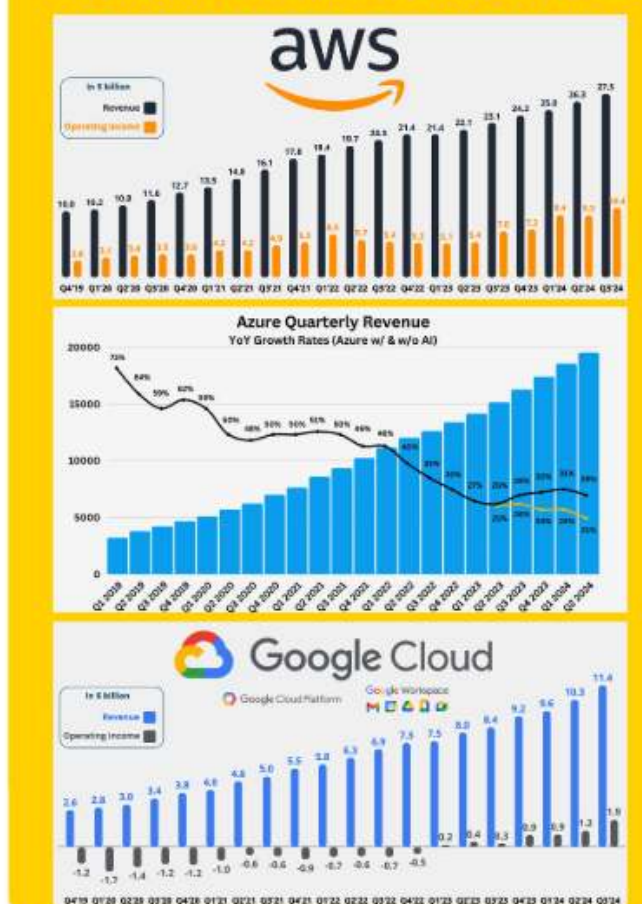
这四家是绝对意义上的“大家伙”
今年的叙事里其实又增加了很多公司
这个数字和公司的自由现金流对比是有意义的

← → 🔍 🔍 那些持续可以再用的“单页” --- 云计算

共识2：云

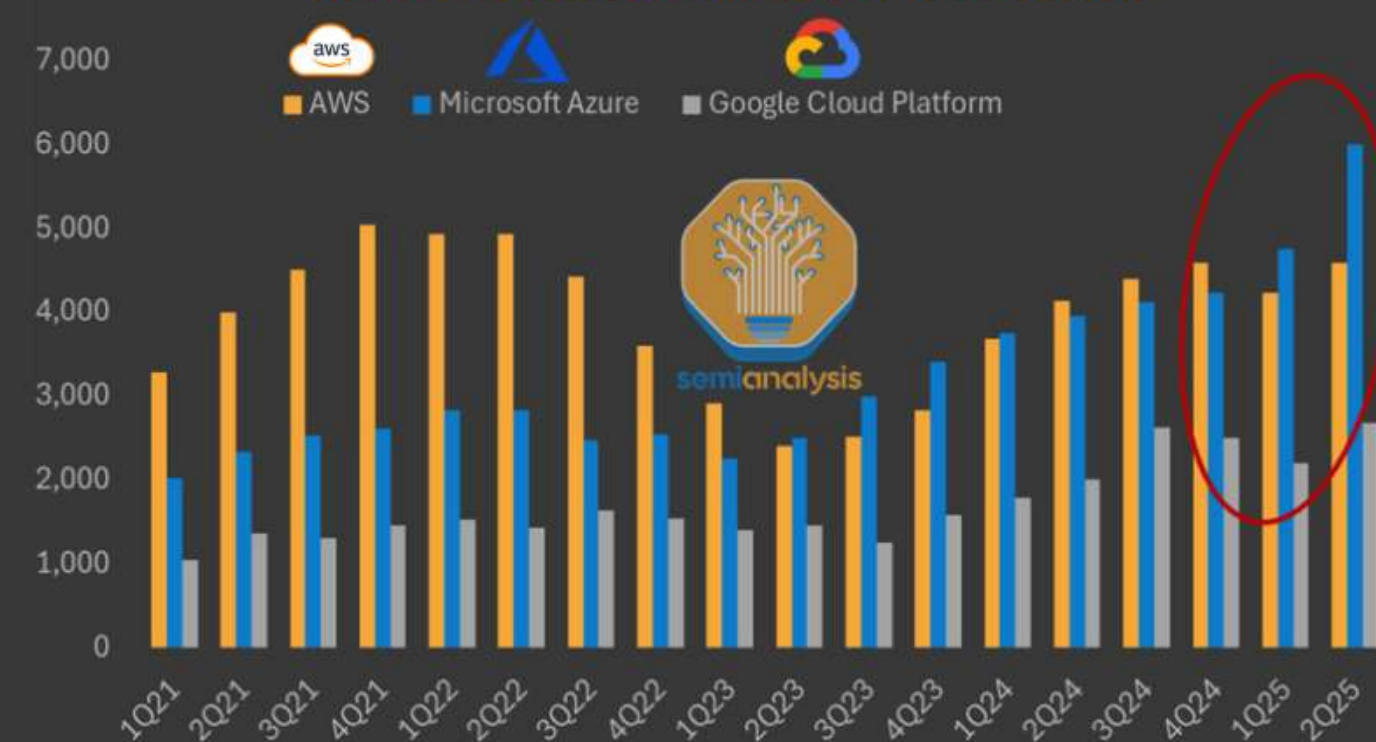


AI的普及某种程度上是靠云收入提升来体现的



“Cloud”合谋

Twelve-months revenue added - USD Million



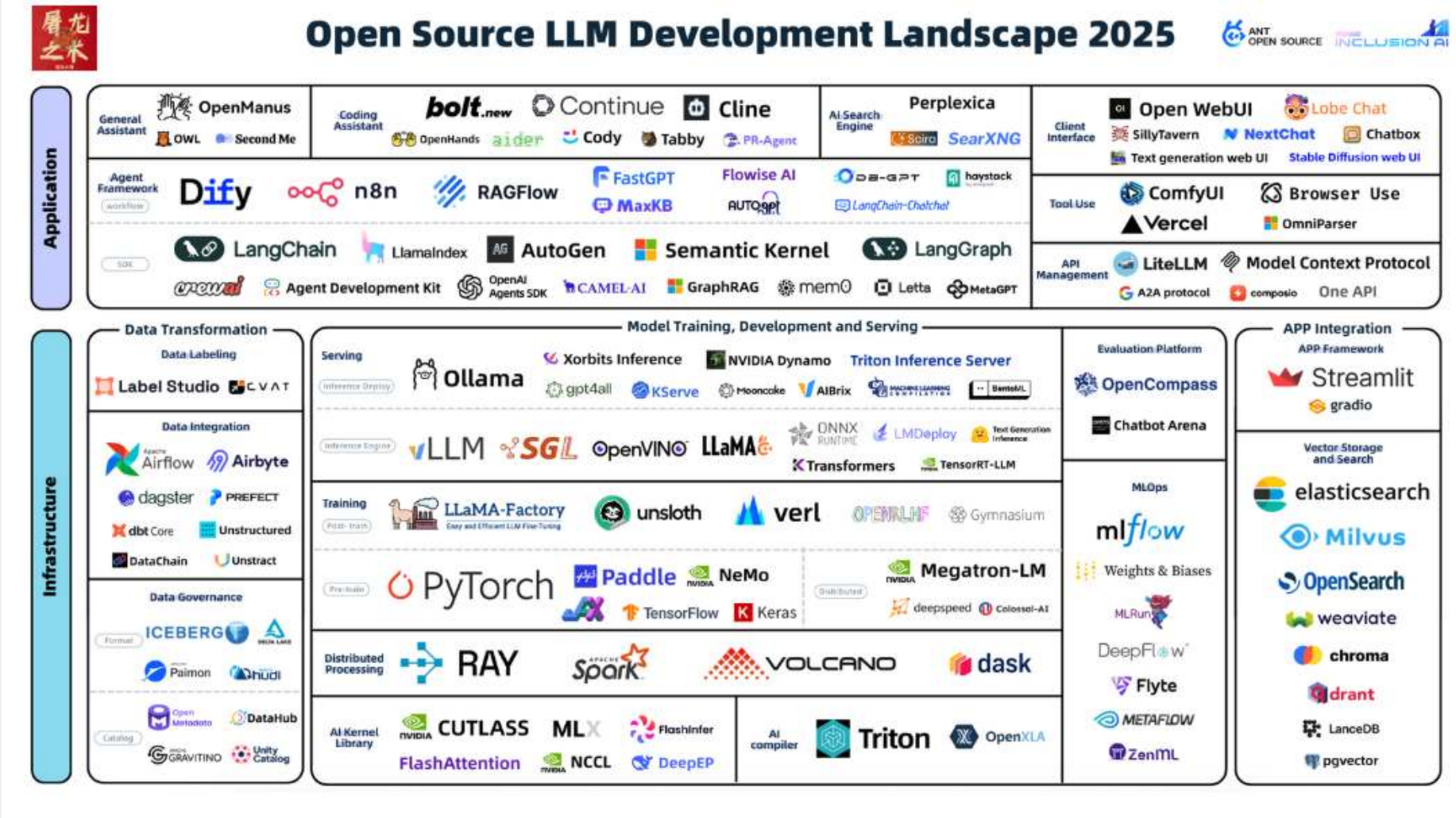
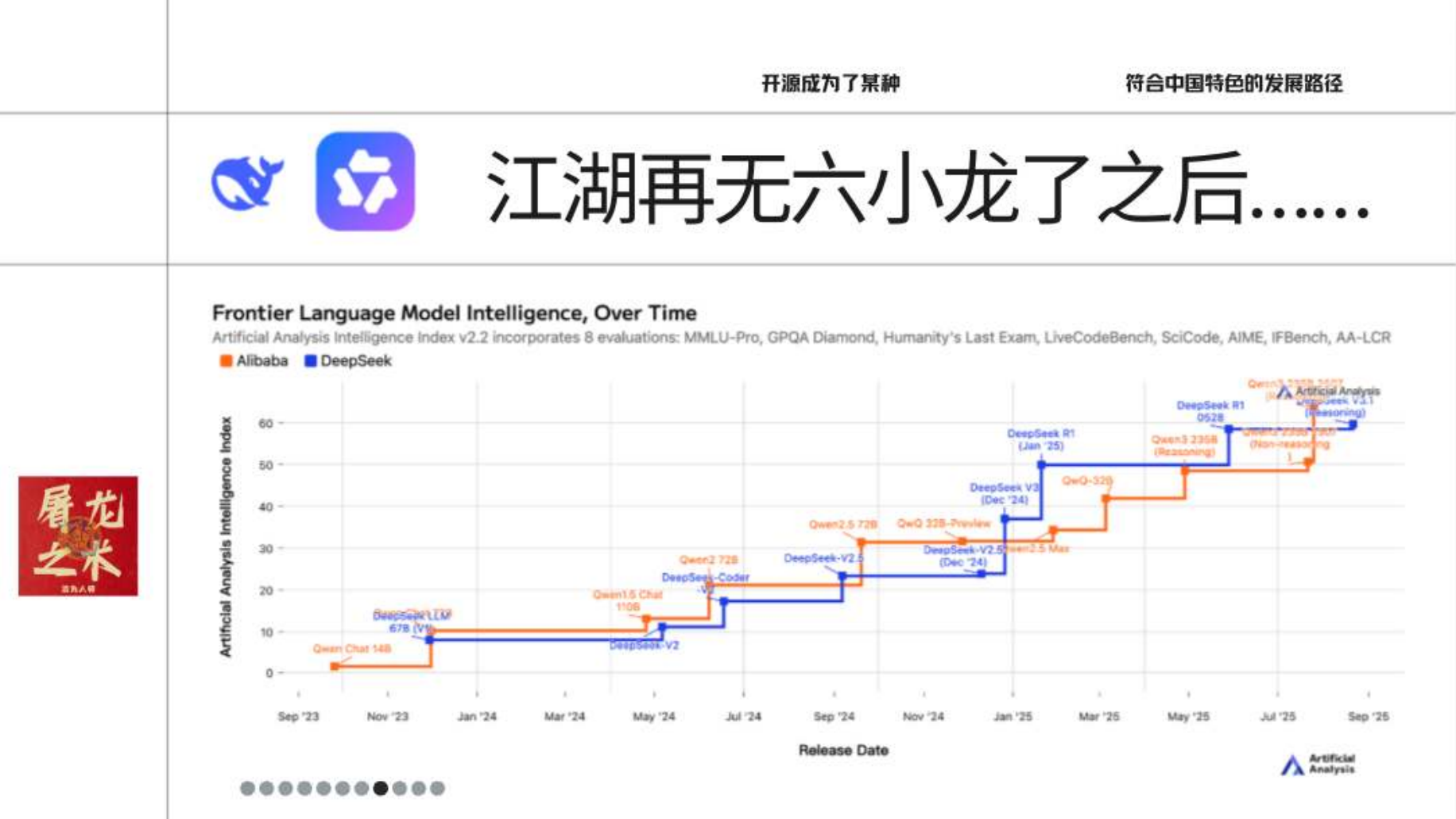
IT和Capex

一体两面

云业务是承接AI落地的第一站
这个共识的达成其实在去年
阿里云讲的也是这个逻辑

AWS依然是领先者
但边际增量上看，微软靠OpenAI强势崛起，
可之后呢？

← → 🔍 🔍 那些持续可以再用的“单页”---开源



开源成为了某种“正确”叙事
且这种叙事可能是由中国厂商引导的
是原因也是结果，也可能是某种“开始”

开源不等于没有商业化
开源生态的繁荣成为了大家追逐的目标
细分板块的发展趋势基本等同行业趋势

← → 🔍 🔍 那些持续可以再用的“单页” ---中美对抗



AI这事儿，大概率就是中美两国的事情了
技术、人才、资本全方位竞争
以及其他国家呢？

屠龙之术

是阶段性的结果？还是某种趋势的开始？

Design Arena	
#1	DeepSeek-R1-0... 1329
#2	Qwen3-Coder-4... 1318
#3	GLM-4.5 1312
#4	DeepSeek-V3-0... 1292
#5	DeepSeek-V3.1-L... 1281
#6	Qwen3-235B-A... 1278
#7	GLM-4.5-Air 1273
#8	DeepSeek-Coder 1268
#9	Qwen3-Coder-3D... 1258
#10	DeepSeek-V3.1 1253
#11	GLM-4-92B 1237
#12	Kimi-K2 1230
#13	Qwen3-235B-A2... 1177
#14	Qwen3-235B-A... 1168
#15	Qwen3-30B-A3B 1154

Business | DeepSeek aftershocks

China is quietly upstaging America with its open models

The Economist

How worried should OpenAI and other labs be?

← r/Futurology · 4 days ago

lughnasadh · transit umbra, lux permanent · Top 1% Poster

a16z

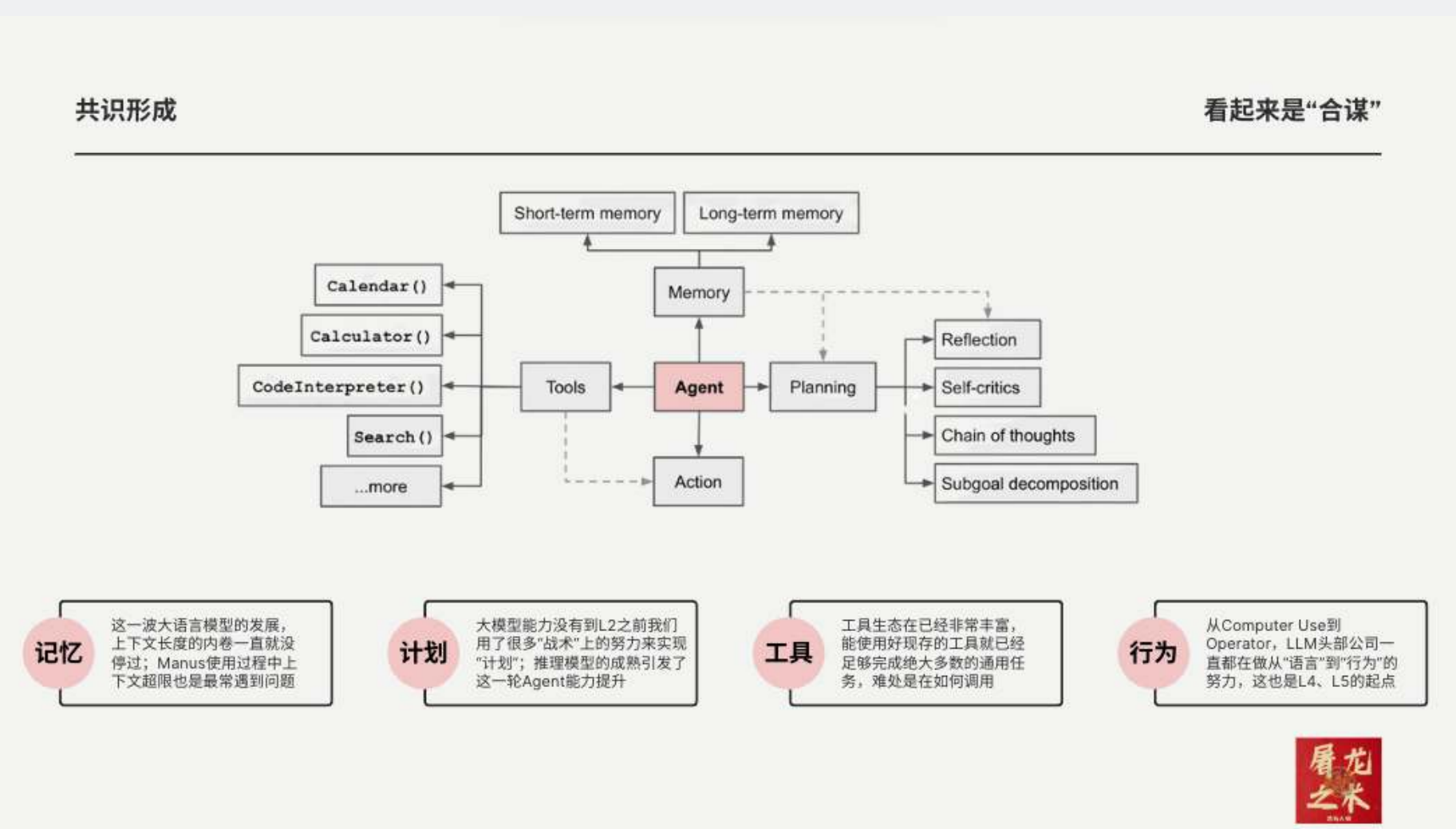
80% of AI start-ups applying for VC funding with Andreessen Horowitz are using Chinese Open-Source AI.

AI

"These days, when entrepreneurs pitch at Andreessen Horowitz (a16z), a major Silicon Valley venture-capital firm, there's a high chance their startups are running on Chinese models. "I'd say there's an 80% chance they're using a Chinese open-source model," notes Martin Casado, a partner at a16z."

拧巴的夹缝中，寻求可能的发展空间
对于创业者、投资人而言是巨大挑战
有些无奈，也有些无解，但活儿还是得干

← → 🔍 🔍 那些持续可以再用的“单页” ---Agent



The AI agent tech stack

Foundation models & infrastructure

Agent frameworks & development platforms

Tool integration

Context

Orchestration

Oversight

仔细想想，要做的工作其实还有很多很多……

The Agent Iceberg

Reporting & Audit

Contact Center Warm Handoff

Intent-based routing

Guardrails and Enforced Determinism

Prompt Injection Protection

Compliance Supervisors

Fine-tuning Pipelines + ML Ops

Multi-channel: Chat, Voice, Email, SMS

Parallelism to Mitigate Latency

Pronunciation of Branded Terms

Background Noise Suppression / VAD

Multi-language Support

Custom Voices

Model Migration and Upgrades

Complex Workflows & Orchestration

Observability and Monitoring

Changelegs

Regression Testing

User Simulation

PII Detection and Encryption

Data Warehouse Export

Role-based access controls

Fuzzy-matching User Authentication

Multi-user QA Workflows

PSTN / SIP Transfers

Stratified Sampling for QA

Reasoning Traces for Audit

Staging and Release Management

Model Redundancy and Failover

屠龙之术

关于Agent本身概念的争论已经少了很多

聚光灯的关注似乎也少了很多

但肯定是今年新项目最多的方向

做Agent和给Agent提供Infra都重要

且有可能在某些领域出现“梯云纵”

当然了，大厂肯定是“全都要”

← → 🔍 🔍 那些持续可以再用的“单页” ---ARR

还不是因为实在找不到什么别的角度了



ARR似乎成为了最重要的**衡量标准**



在连分类都很困难的当下，如何评价Agent？
不看收入还能看什么呢？
看毛利吗？



尤其再考虑到CLAUDE CODE的快速增长

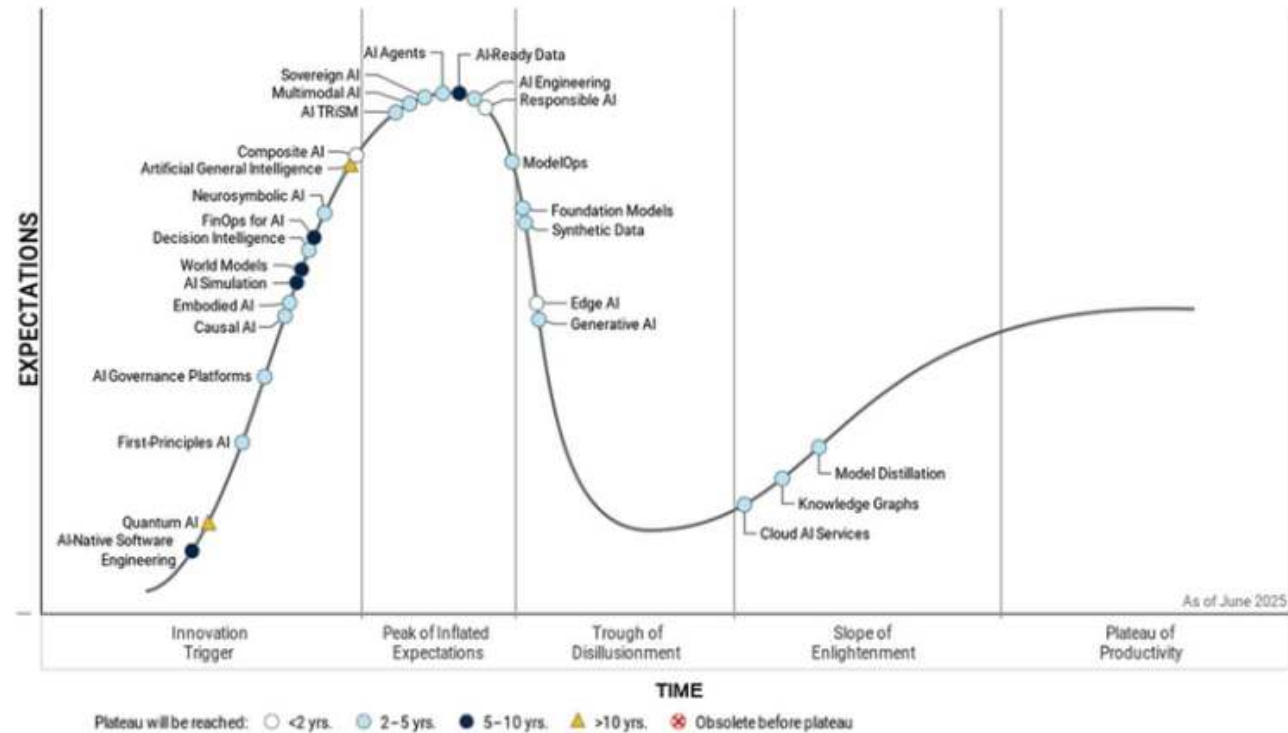
AI是一场巨大的.....**击鼓传花？**



毛利似乎没法看，因为一地鸡毛
单位成本的下降看起来是“幻觉”
在硅谷亏损从来都不是问题……吗？

← → 🔍 🔍 那些持续可以再用的“单页” --- Gartner技术曲线

2025 AI 技术成熟度曲线



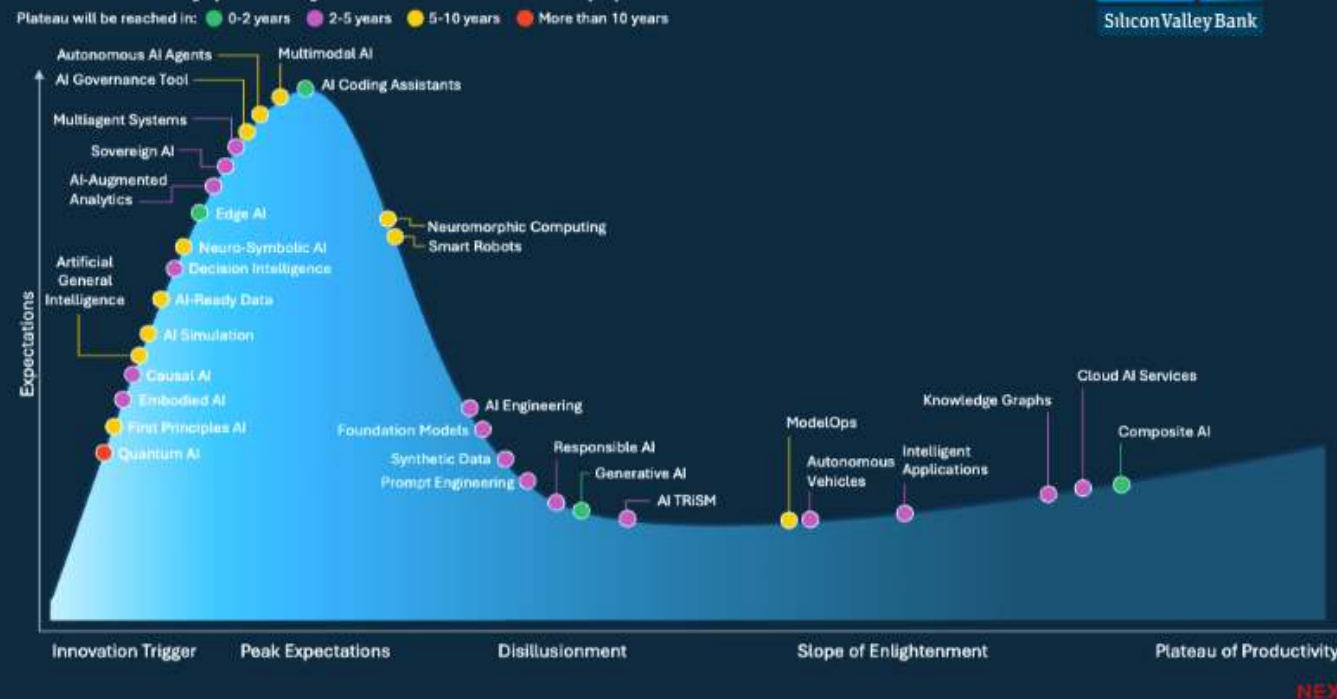
Gartner

讲了许多年了

这个分析框架其实很好用
作为一个心里的标尺
也作为对未来判断的依据

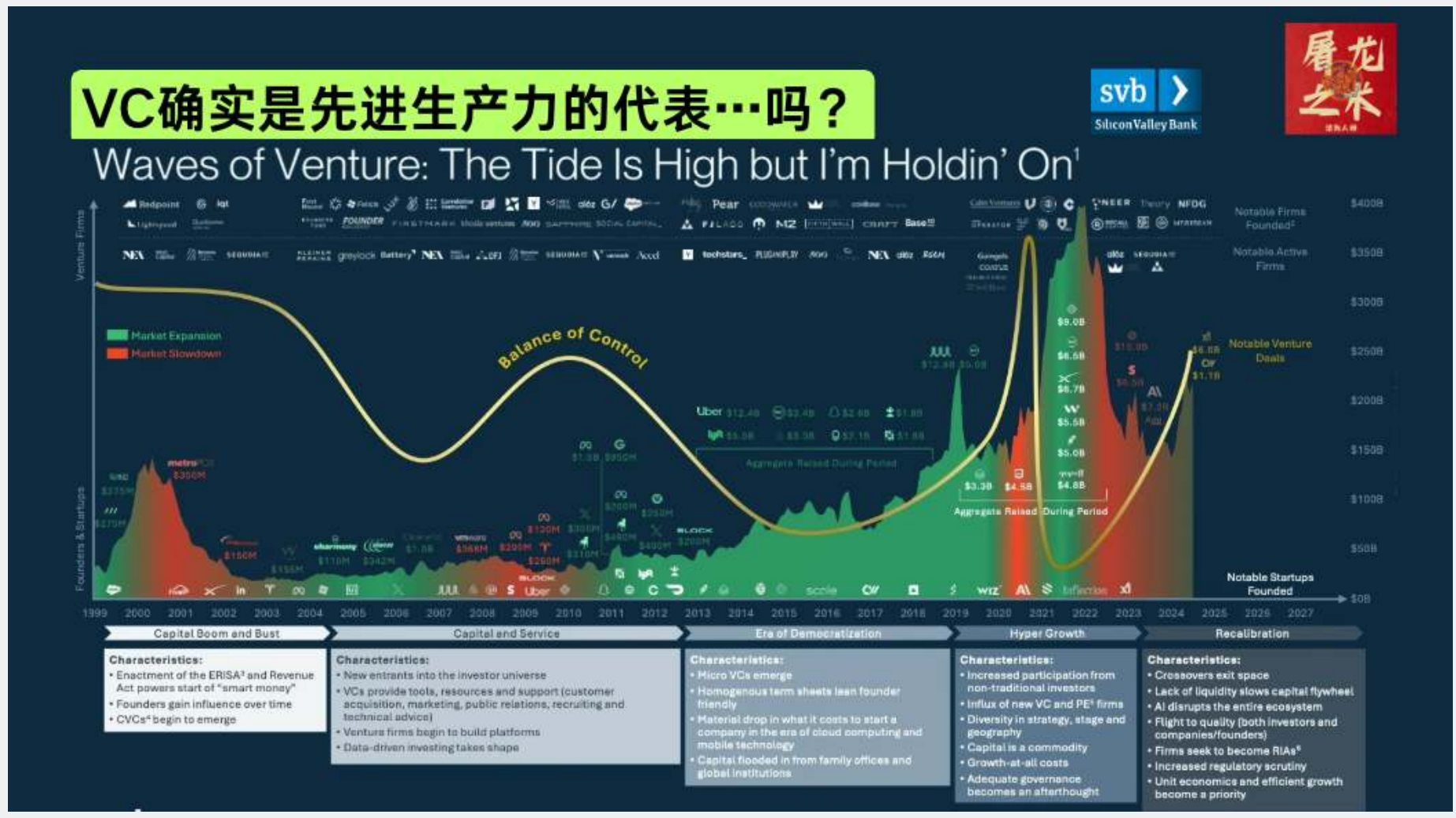
写在最后

The AI Hype Cycle: Future Opportunities

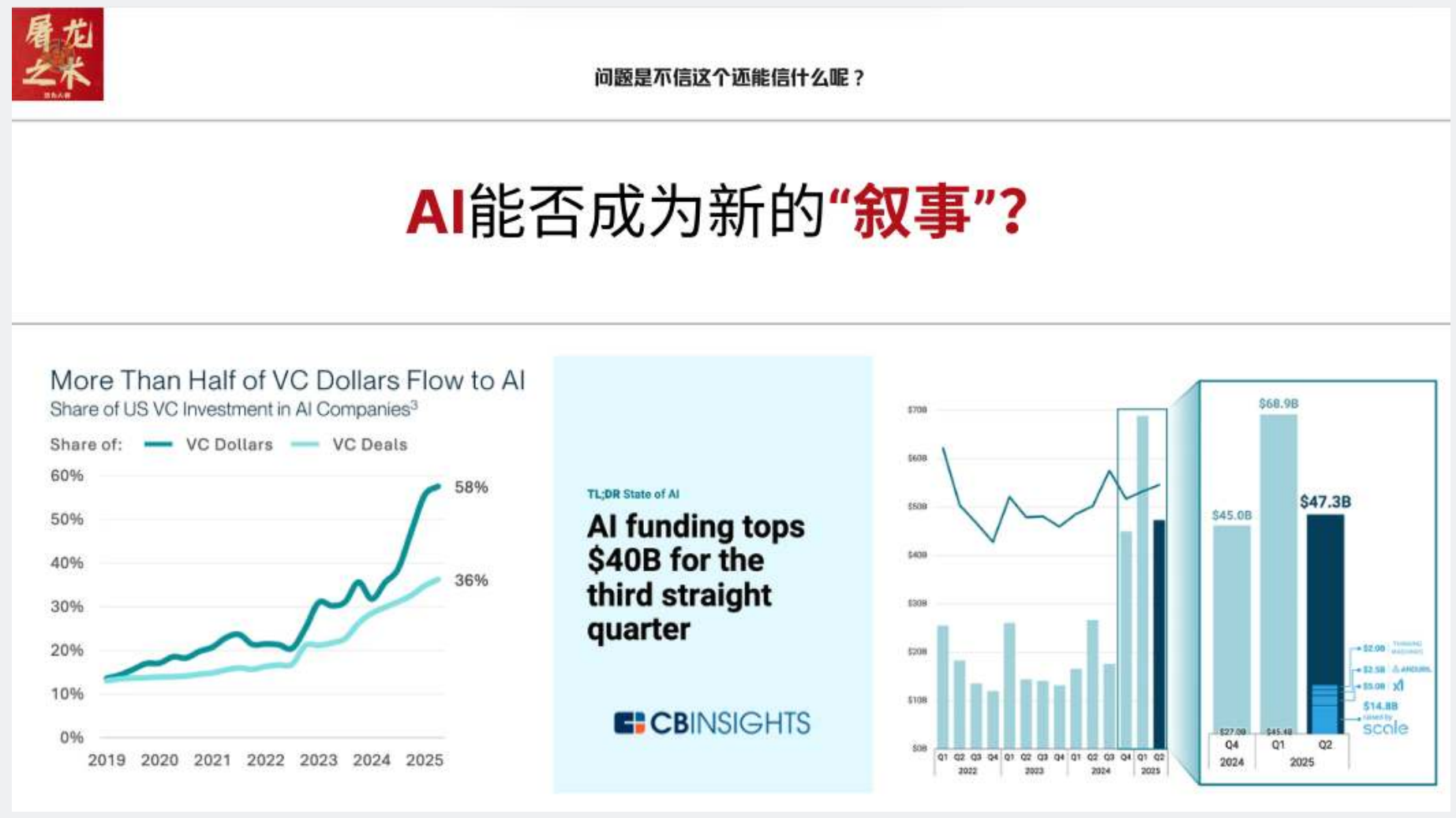


但这个曲线似乎也意味着泡沫是必然的过程
这个曲线一旦和股票曲线联系在一起
很多事情看起来就“合理”了

← → 🔍 🔍 那些持续可以再用的“单页” ---一级市场



你是否还相信技术创新带来经济增长？
你是否还相信VC对于创新的促进？
认清现实是否能依然保持热爱？



AI当然是最重要的叙事，
作为当年从沙子里做出芯片的代表，
面对海量的流动性，不投这个投什么？🐶

← → 🔍 似乎整个故事已经超级完整了……

每次做PPT最痛苦的事情是什么？

行业的发展比你做PPT的速度还快...

← → 🔍 要不怎么说AI-司马迁难做呢，这些事情全部发生在本周

OpenAI | Microsoft



JENSEN HUANG
KEYNOTE

October 2025

Live stream offline



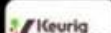
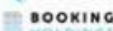
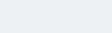
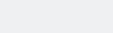
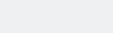
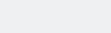
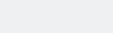
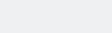
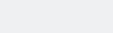
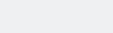
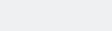
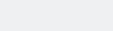
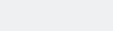
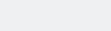
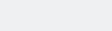
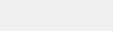
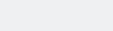
LIVE from Washington, D.C.: NVIDIA CEO Jensen Huang at GTC 2025



Brew Markets

This week is a big one for earnings:

- Apple
- Microsoft
- Alphabet
- Amazon
- Meta

The Transcript		Earnings Calendar				Week 44 of 2025		
Monday 10/27		Tuesday 10/28		Wednesday 10/29		Thursday 10/30		Friday 11/01
Before Open	After Close	Before Open	After Close	Before Open	After Close	Before Open	After Close	Before Open
 Keurig Dr Pepper	 Whirlpool	 United Healthcare	 ENPHASE	 verizon	 Meta	 Lilly	 amazon	 ExxonMobil
 carter's	 UHS	 PayPal	 ONEOK	 BOEING	 Microsoft	 MERCK	 Apple	 Charter COMMUNICATIONS
 Bank of America	 NXP	 jetBlue	 BOOKING HOLDINGS	 CVS Healths	 Alphabet	 insmed	 First Solar	 abbvie
 revvity	 UPS	 SoFi	 SEAGATE	 CENTENE	 STARBUCKS	 COMCAST	 coinbase	 Chevron
 Assurant	 TWO HARBORS	 NUCOR	 Meritage Homes	 CATERPILLAR	 mercado libre	 ROBLOX	 Western Digital	 Dominion Energy
 Bank of Hawaii	 Olin	 MSCI	 VISA	 Kraft Heinz	 AGNICO EAGLE	 reddit	 MicroStrategy	 CN AON
 City Bank	 Rambus	 APPLIED	 Bloomenergy	 66	 CARVANA	 Baxter	 Roku	 Cboe
 DAQO	 NET INCORP	 CElestica	 GREENBRIER	 GE HealthCare	 SPROUTS	 Advance Auto Parts	 SHAKE SHACK	 BLUE OWL
	 welltower	 IP HORTON	 Mondelez	 Iserv	 ebay align	 servicenow	 CLOUDFLARE	 appfolio
	 cadence	 VF CORPORATION	 TERADYNE	 Brinker	 NiSource	 eBay	 SHAKE SHACK	 BLUE OWL
	 CONFLUENT	 ROYAL CARIBBEAN	 EDISON	 GARMIN	 equis	 TransMedics	 twilio	 MPS
	 NOY	 NEXTERA ENERGY	 equinor	 KLA		 Cardinal Health	 elderado gold	 ROCKET Companies, Inc.

42

305

1.7K

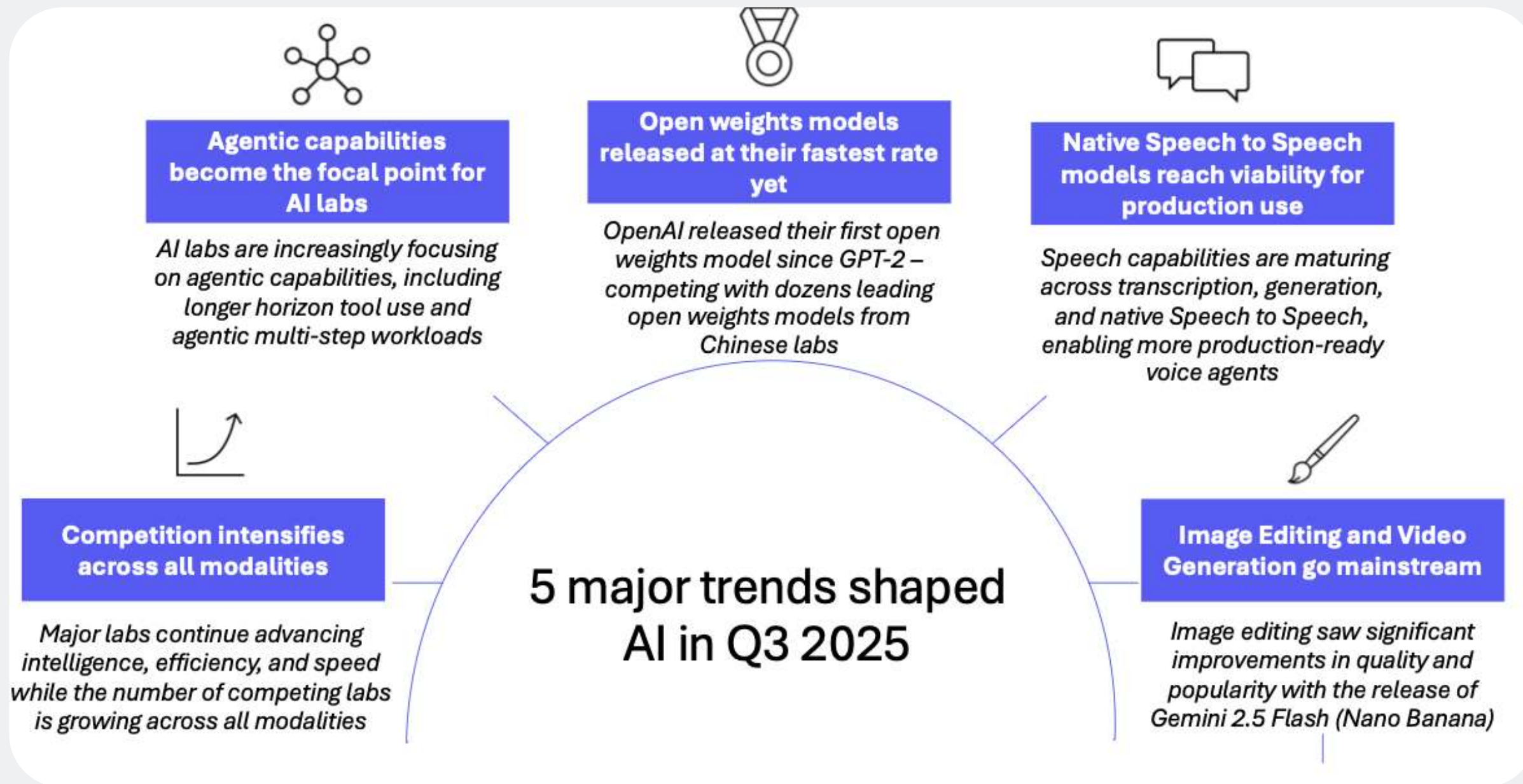
160K



← → 🔍 虚无的AGI似乎已经无人提及，但渐进式的技术演进依然持续

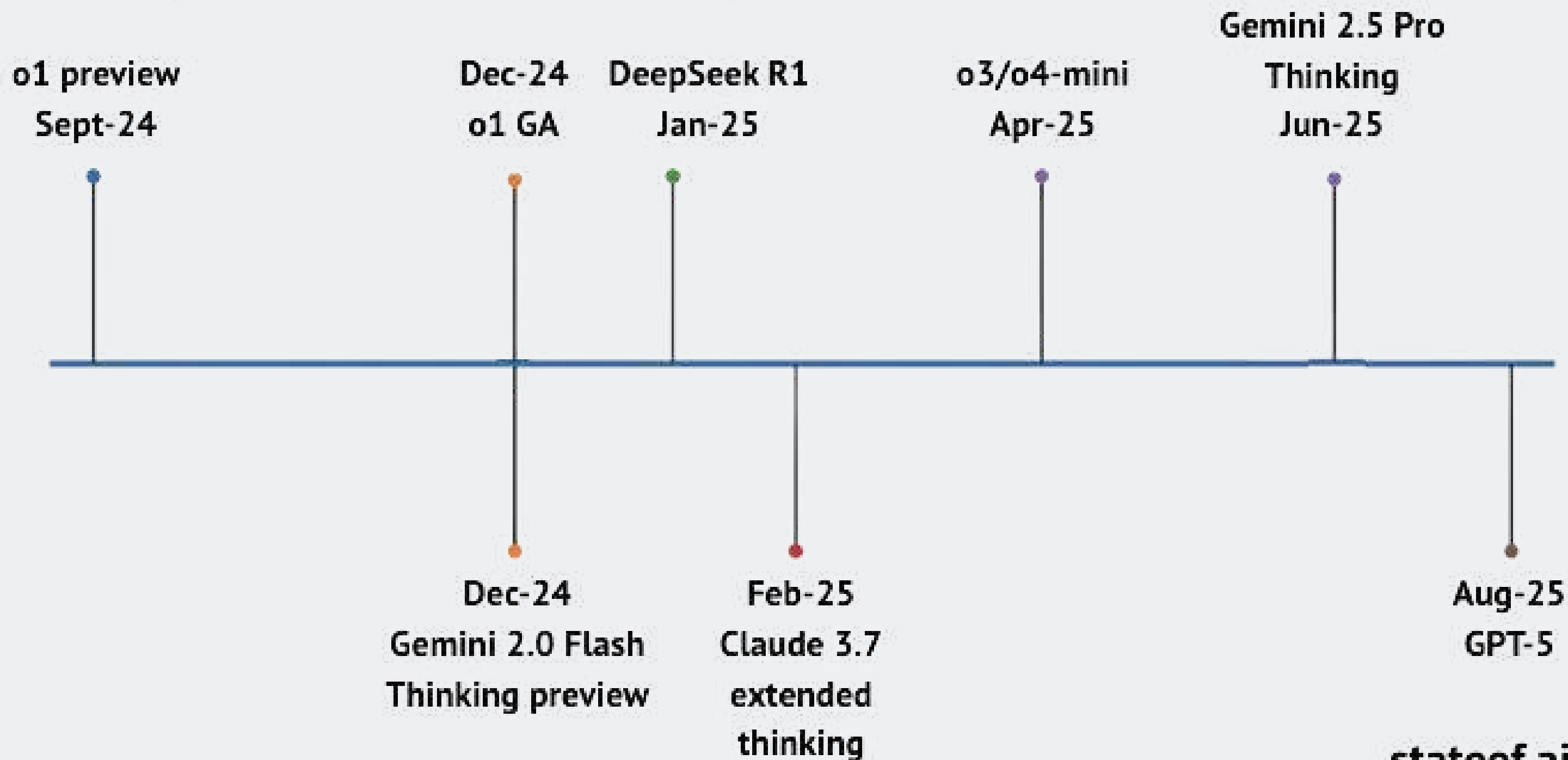
技术演进

- LLM
- 开源
- 多模态
- Agent



← → 🔍 2025年LLM主航道---推理模型

The reasoning timeline: from o1 “thinking” to R1, GPT-5 and parallel compute routing



stateof.ai 2025

← → 🔍 在纯LLM的研发上，我们遇到了什么问题？

激励模型

- 可验证结果---模糊/主观目标
- 二元结果---横向人类偏好
- 不可验证的创造性任务
- 过程激励 or 结果激励
- Reward Hacking
- 更广泛意义上的数据获取
-

记忆

- 上下文长度可能仅仅是开始
- 长时与短时记忆的取舍
- 可控性差
- 语义漂移与记忆冲突
- 检索/访问效率低
- 纯文本之外的可能性
-

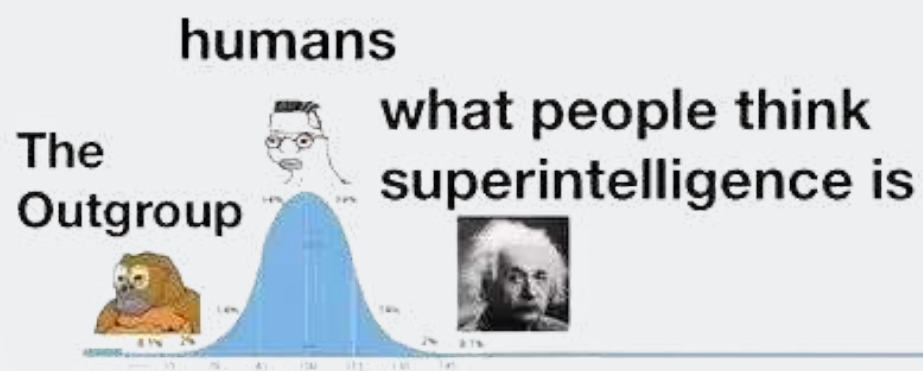
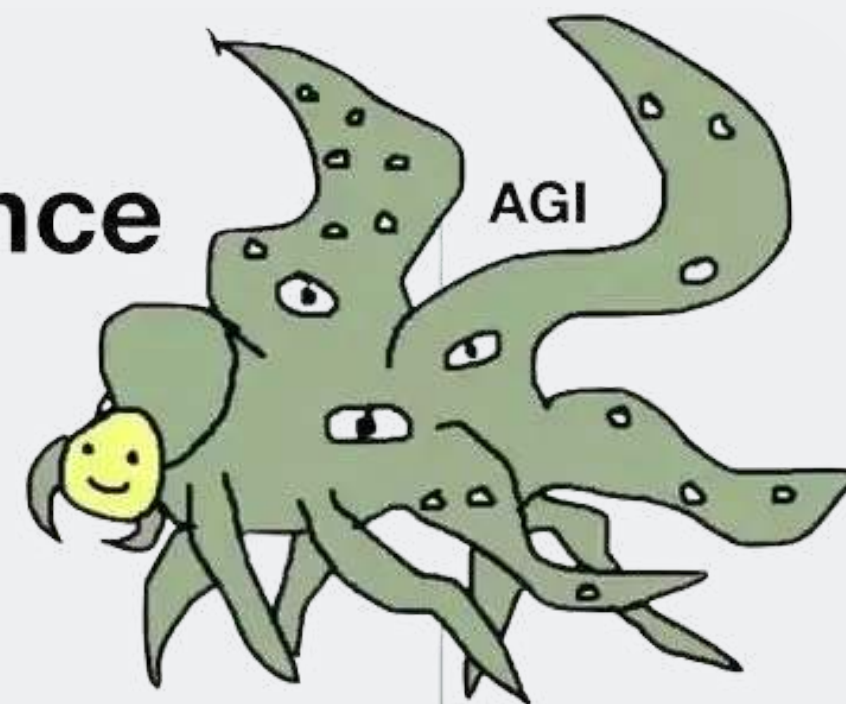
Benchmark

- 基准指标快速失效
- 训练/评估范式与现实脱轨
- 应用导向 VS 基准导向
- 数据泄漏/模型过拟合严重
- 动态变化技术能力 VS 静态评估
- “长期有效”的 benchmark 很难
-

← → 🔍 数据和benchmark层面的“难”

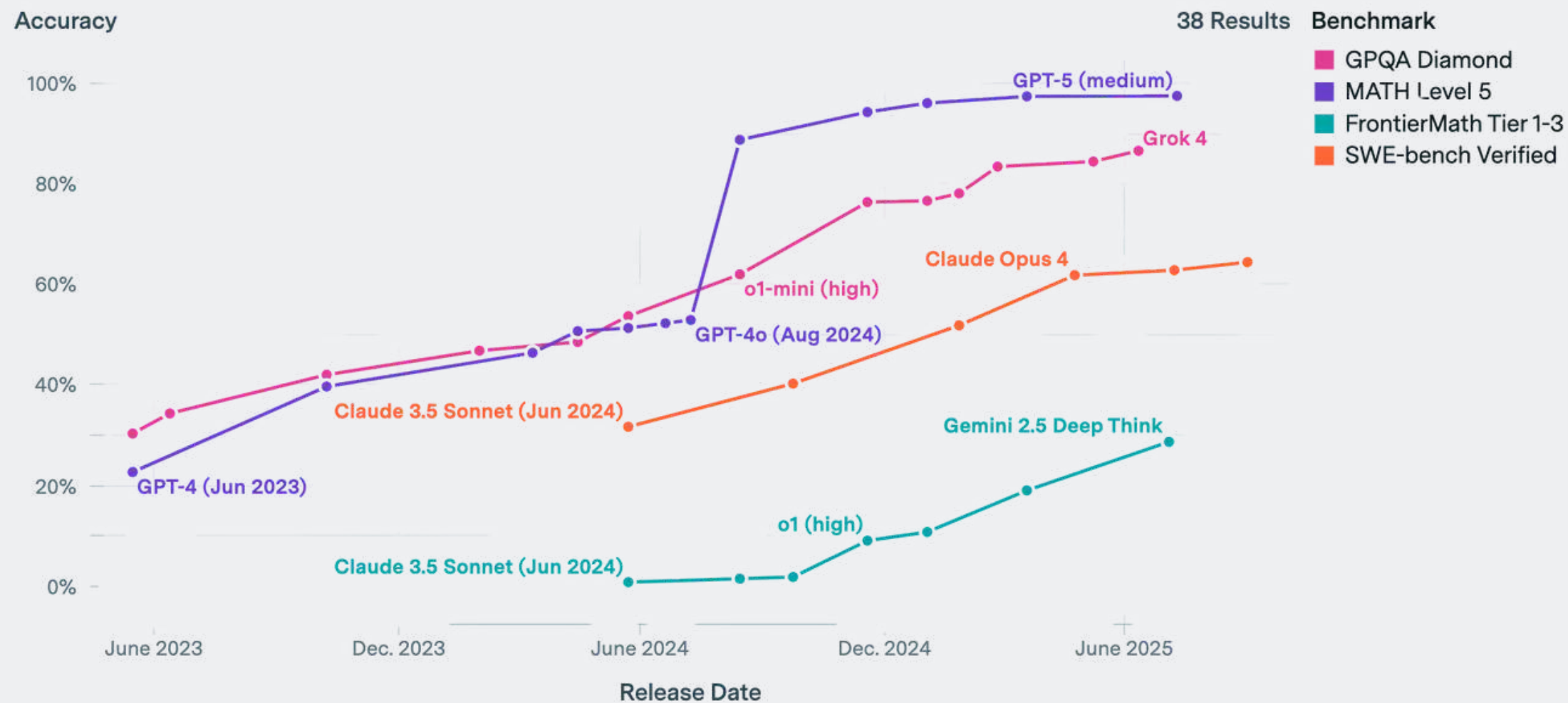


actual fucking
superintelligence



← → 🔍 在美国，纯基础模型的战役参与方就是这老几位了……

Frontier performance across benchmarks



OpenAI



Claude



Gemini



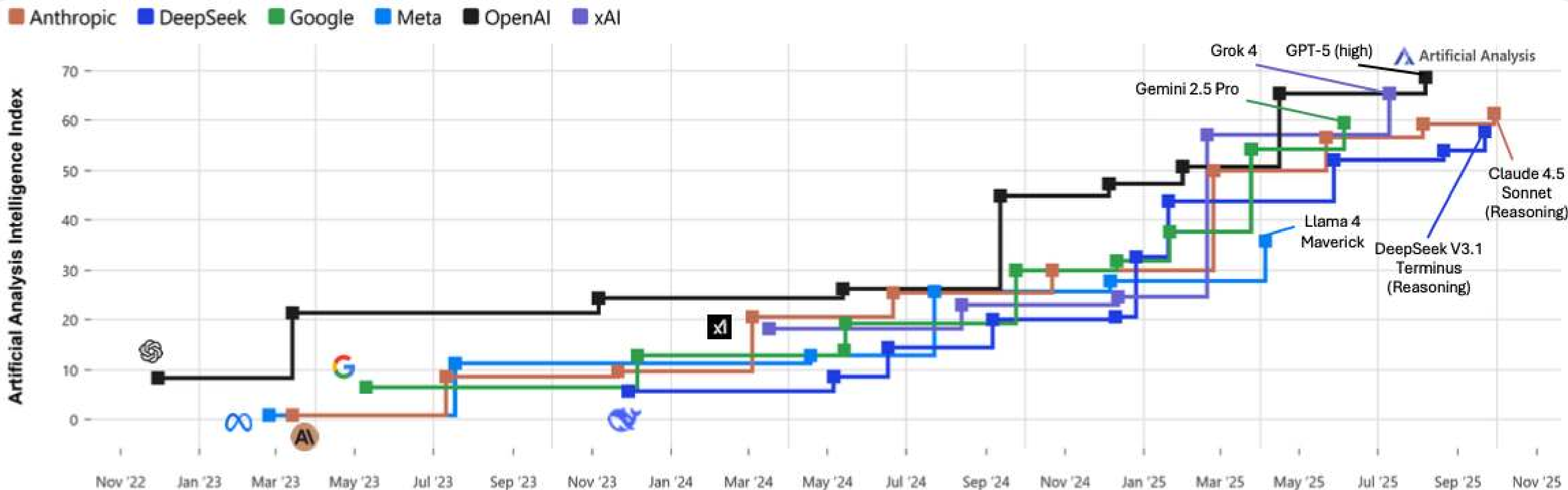
Grok

LLAMA 4



哥，我呢？

← → 🔍 🔍 方向和方法都给定，剩下的就是卷了……

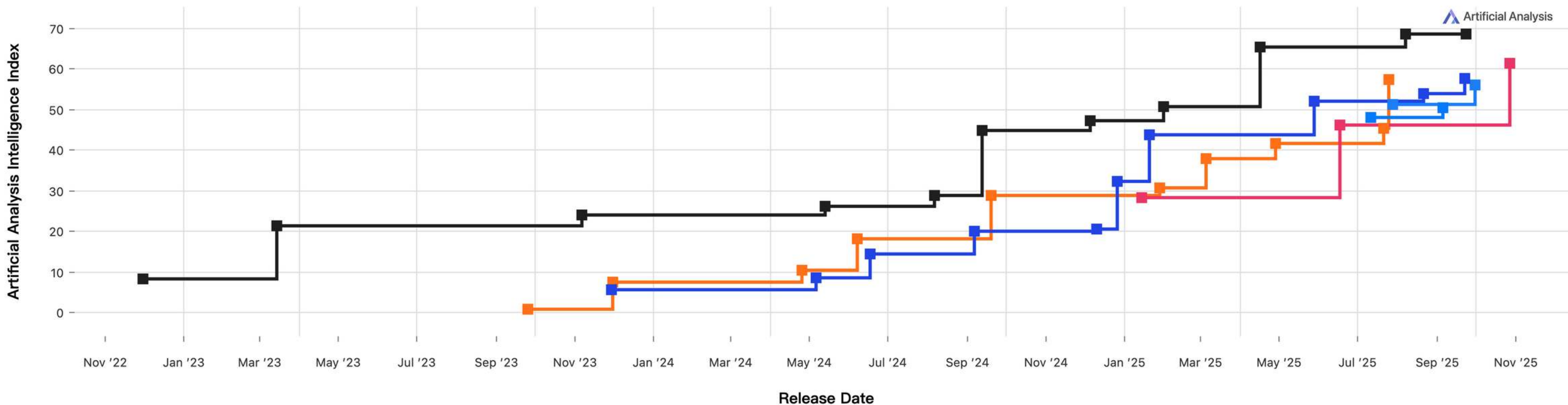


← → 🔍 六小龙叙事之后，**开源**似乎成为了我们这边的“主叙事”

Frontier Language Model Intelligence, Over Time

Artificial Analysis Intelligence Index v3.0 incorporates 10 evaluations: MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIME 2025, IFBench, AA-LCR, Terminal-Bench Hard, τ^2 -Bench Telecom

Alibaba DeepSeek MiniMax Moonshot AI OpenAI Z AI



一个小问题：OpenAI-5000亿美金，中国头部模型厂商的合理估值应该是多少？

← → 🔍 🔍 开源甚至被写进了国务院-“人工智能+”行动意见中

制，推进跨行业、跨领域、国际化标准联动。

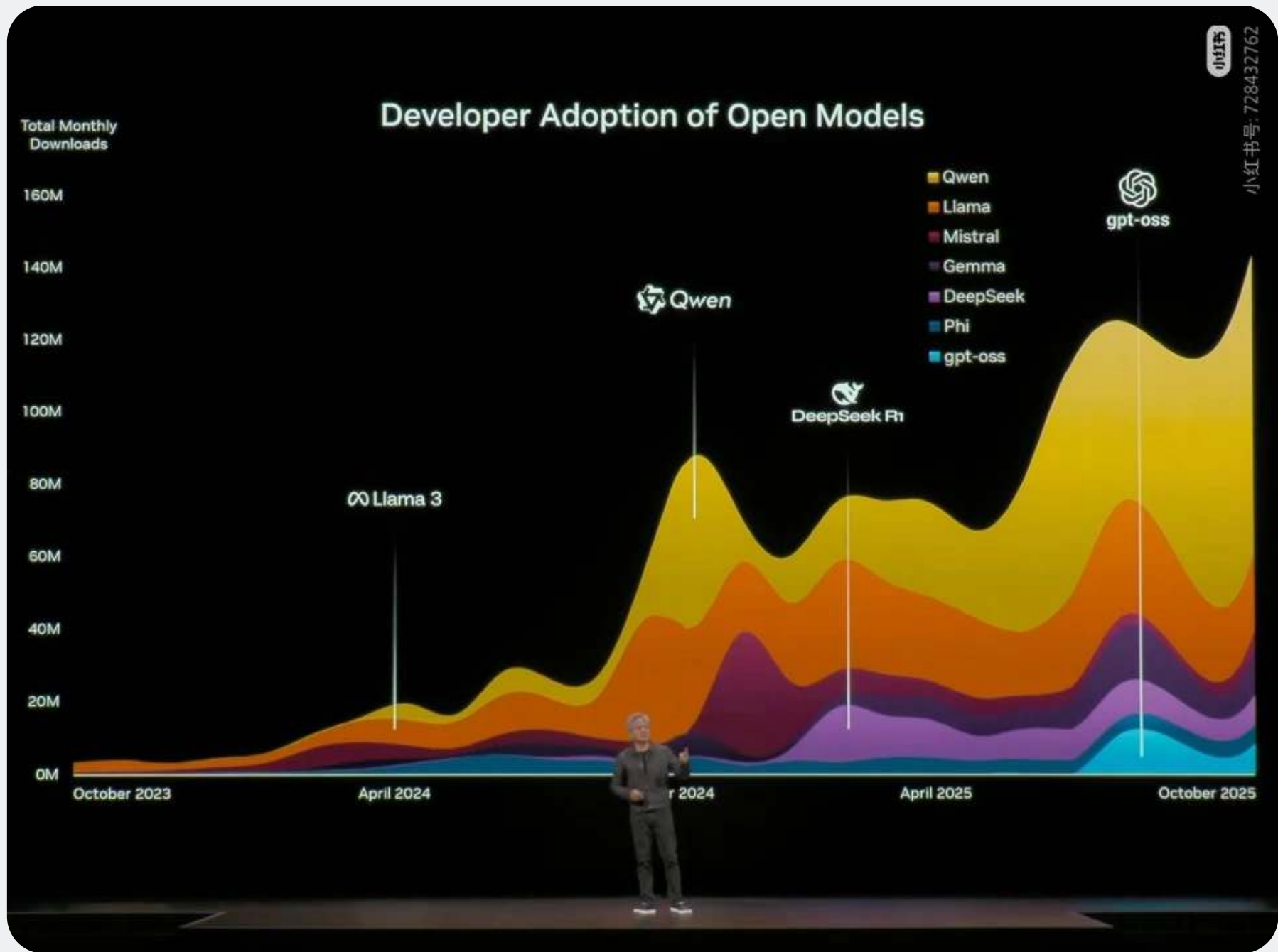
（十一）促进开源生态繁荣。支持人工智能开源社区建设，促进模型、工具、数据集等汇聚开放，培育优质开源项目。建立健全人工智能开源贡献评价和激励机制，鼓励高校将开源贡献纳入学生学分认证和教师成果认定。支持企业、高校、科研机构等探索普惠高效的开源应用新模式。加快构建面向全球开放的开源技术体系和社区生态，发展具有国际影响力的开源项目和开发工具等。

（十二）加强人才队伍建设。推进人工智能全学段教育和全社会通识教



信息标准化咨询服务

国务院关于深入实施“人工智能+”行动的意见



Nathan Lambert
@natolambert

Subscribe



Airbnb CEO Brian Chesky: “We’re relying a lot on Alibaba’s Qwen model. It’s very good. It’s also fast and cheap... We use OpenAI’s latest models, but we typically don’t use them that much in production because there are faster and cheaper models.”

The valley is built on Qwen?

11:28 PM · Oct 21, 2025 · 883.9K Views



111



426



4.1K



1.3K



Nathan Lambert
@natolambert · Oct 21



From latimes.com



1



5



95

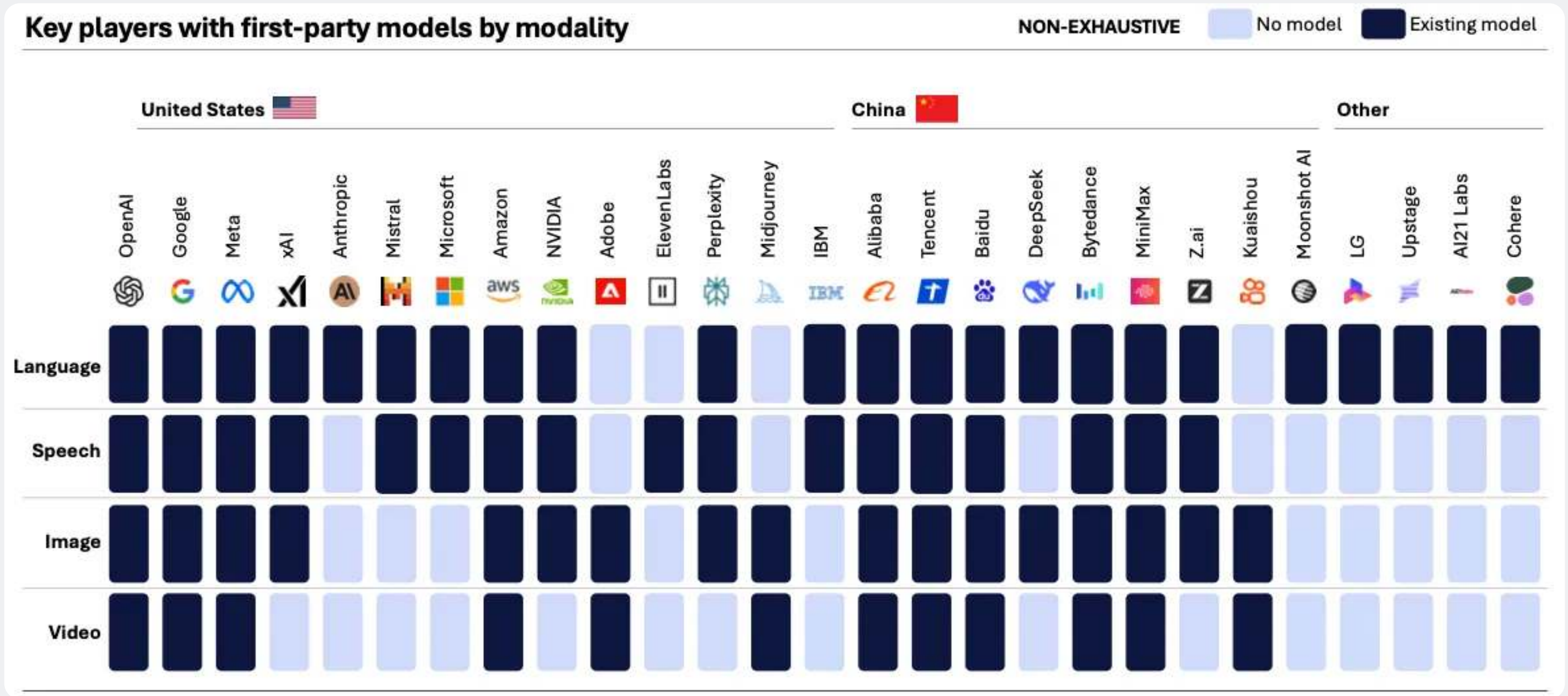


49K



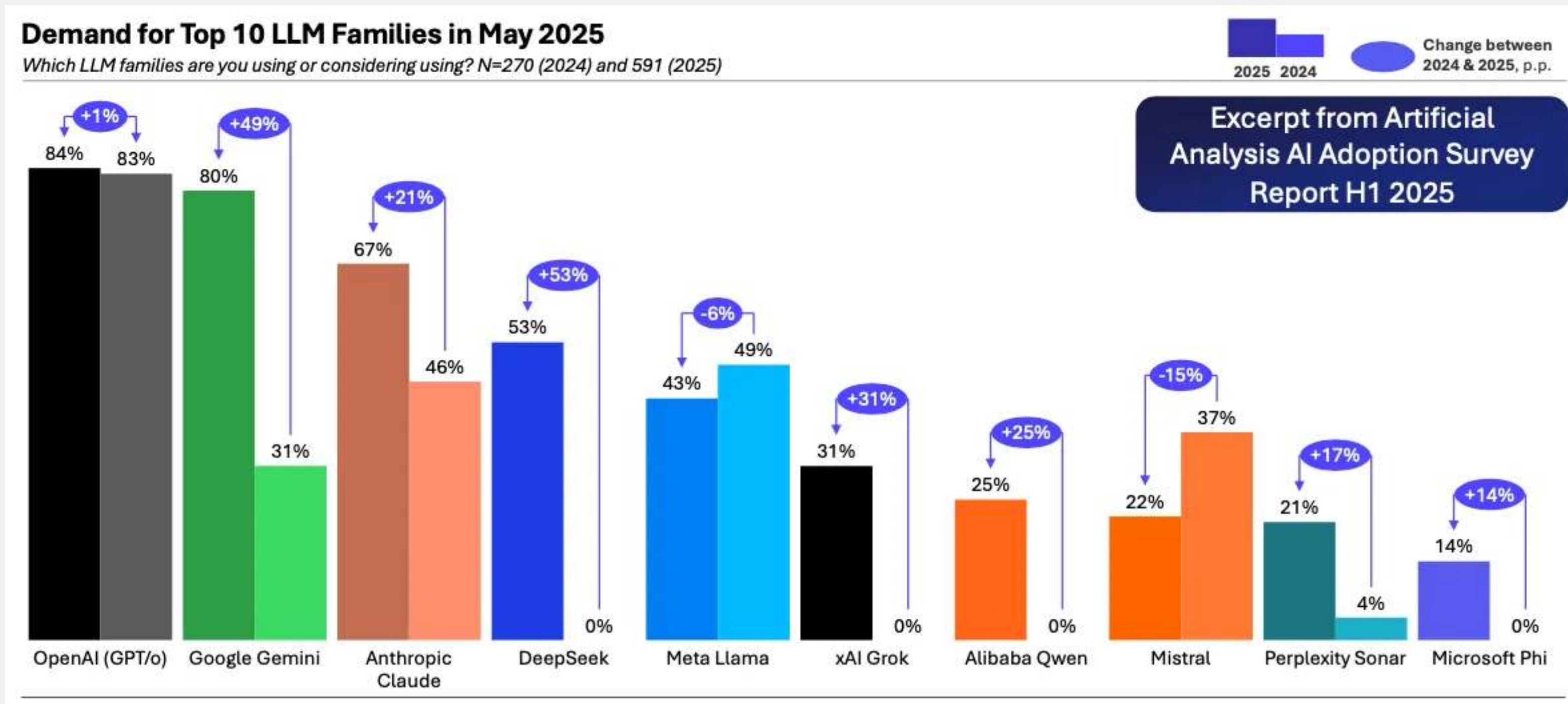
1

← → 🔍 LLM主桌上的也就是这几位了，且只会越来越少



小问题+1： Mistral都140亿美金了， 中国头部模型厂商的合理估值应该是多少？

← → 🔍 竞争残酷，位次/排名都不稳定

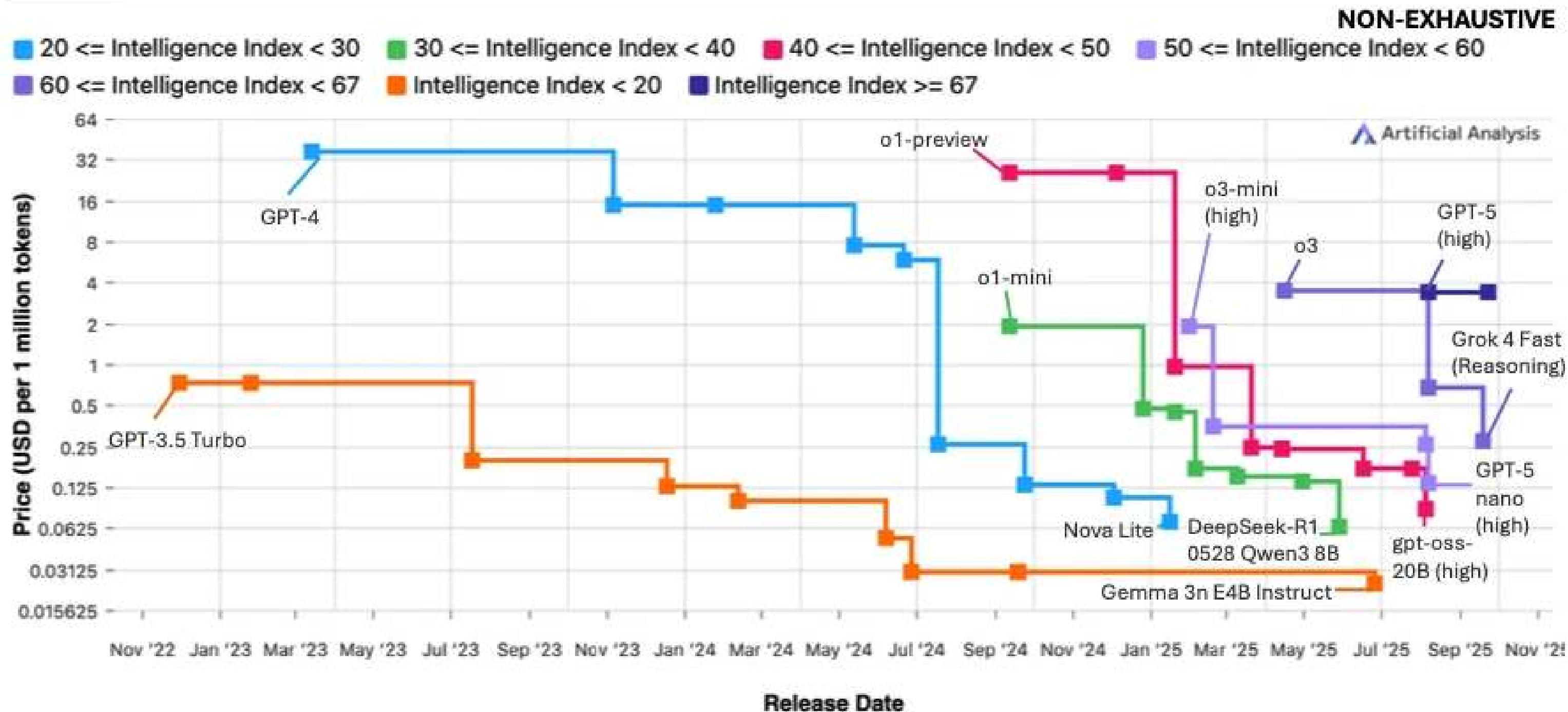


小问题+1： Mistral都140亿美金了，中国头部模型厂商的合理估值应该是多少？

← → 🔍 单token**成本下降**，除了最好的模型……

Language Model Inference Pricing by Intelligence Class, Over Time

Price in USD per 1 million tokens (blended input to output token price 3:1) ; Artificial Analysis Intelligence Index v3 (includes 10 evaluations)



← → 🔍 🔍 回头看25年，印象最深的一次模型发布可能不是GPT5



deepseek

9月29日

DeepSeek-V3.2-Exp 发布，训练推理提效，API 同步降价

阅读 10万+ 赞 5677 5个朋友读过

V3.2 Exp

9月22日

DeepSeek-V3.1 版本更新

阅读 10万+ 赞 5602 2个朋友读过

V3.1 更新

	Benchmark	DeepSeek-V3.1-Terminus	DeepSeek-V3.2-Exp
General	MMLU-Pro	85.0	85.0
	GPQA-Diamond	80.7	79.9
	Humanity's Last Exam	21.7	19.8
Search Agent	BrowseComp	38.5	40.1
	BrowseComp-zh	45.0	47.9
	SimpleQA	96.8	97.1
Code	LiveCodeBench	74.9	74.1
	Codeforces-Div1	2046	2121
	Aider-Polyglot	76.1	74.5
Code Agent	SWE Verified	68.4	67.8
	SWE-bench Multilingual	57.8	57.9
	Terminal-bench	36.7	37.7
Math	AIME 2025	88.4	89.3
	HMMT 2025	86.1	83.6



rosicky311 (明浩) 上海

14 1

原来3.1终极版之后是3.2 还以为是4 终究是我自作多情了



2条回复



污喵王 四川

3 1

没毛病啊，3.1终极版，现在不换成3.2的嘛



赛博禅心

未来两周：DS-v3.2/GLM-4.6/OpenAI... 集中发布

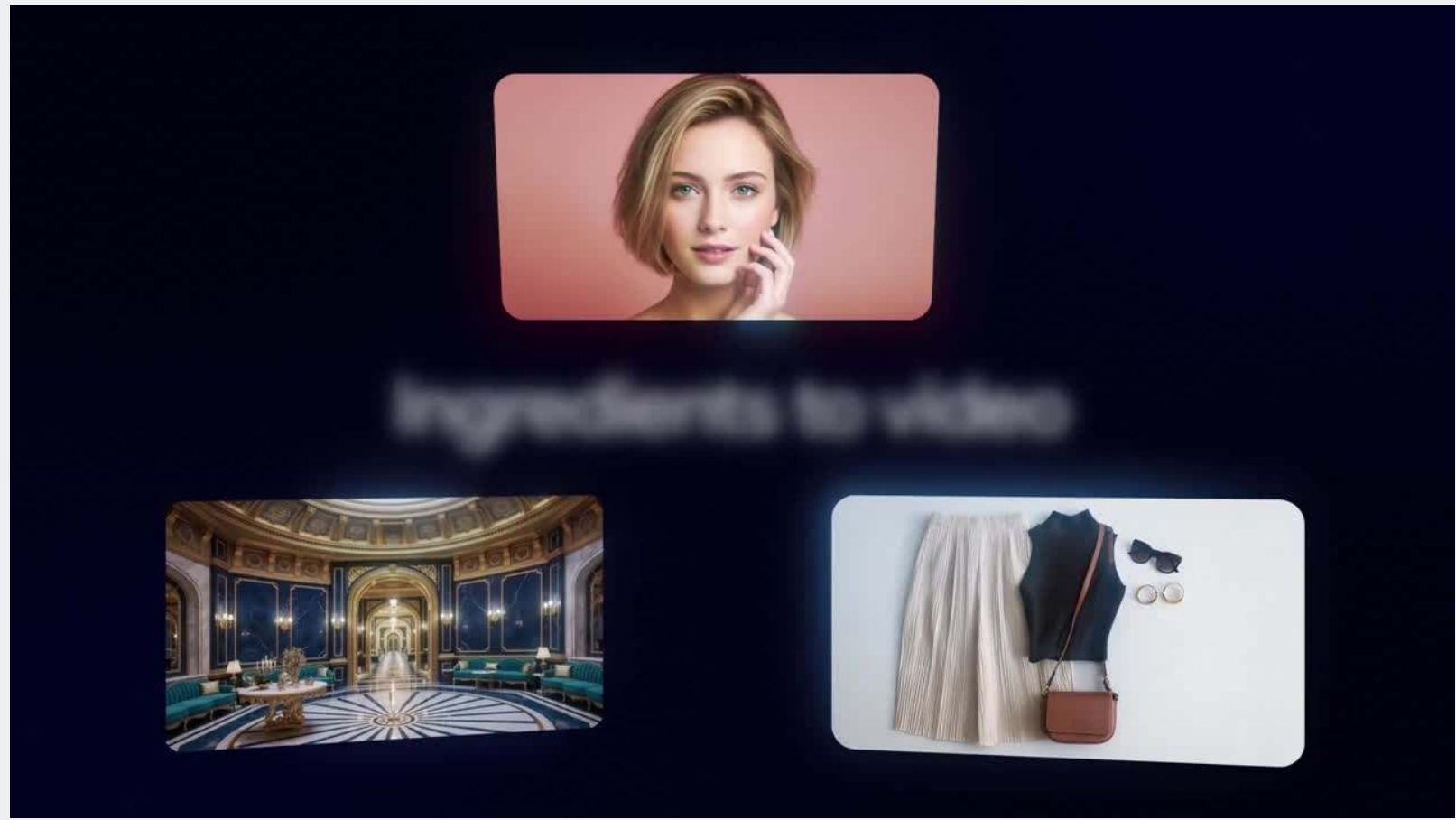
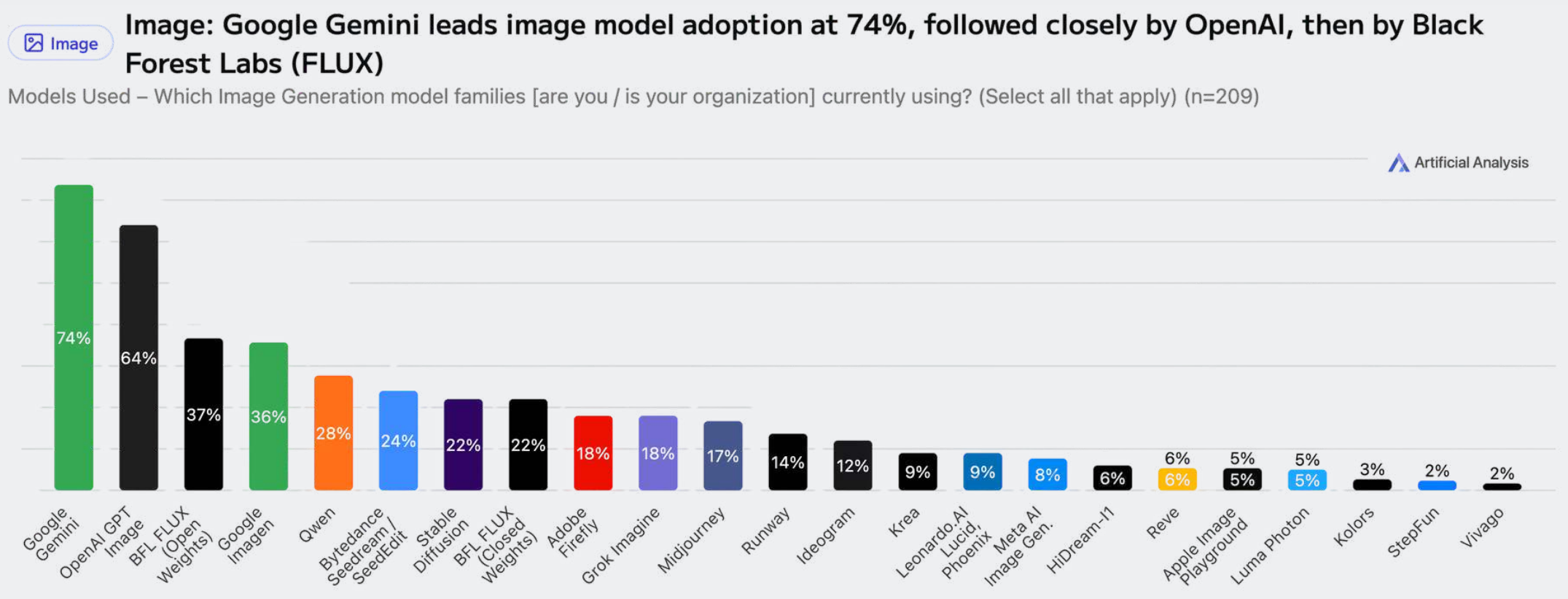
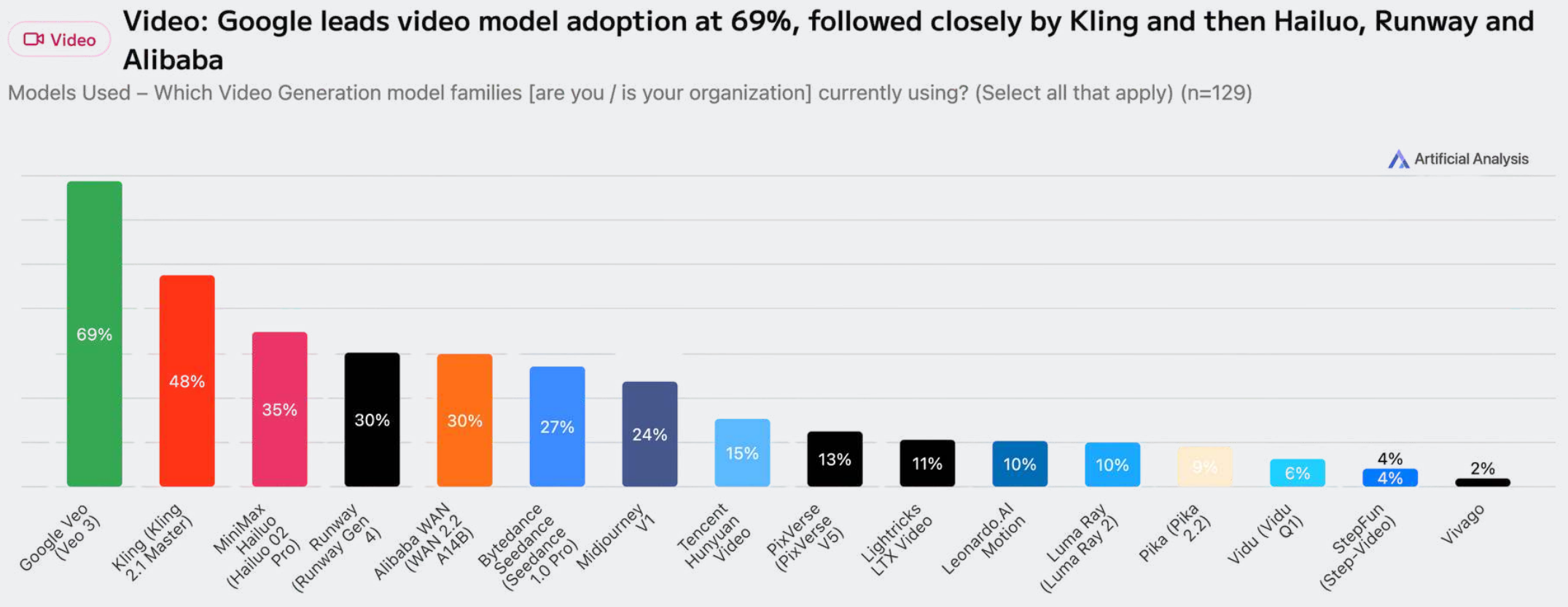


小问题+1：DeepSeek R2会是基于V3.2而来么？

← → ↺ 🔍 多模态，借用一张之前的PPT

[illegible]

图与视频已经进化到 “Media”



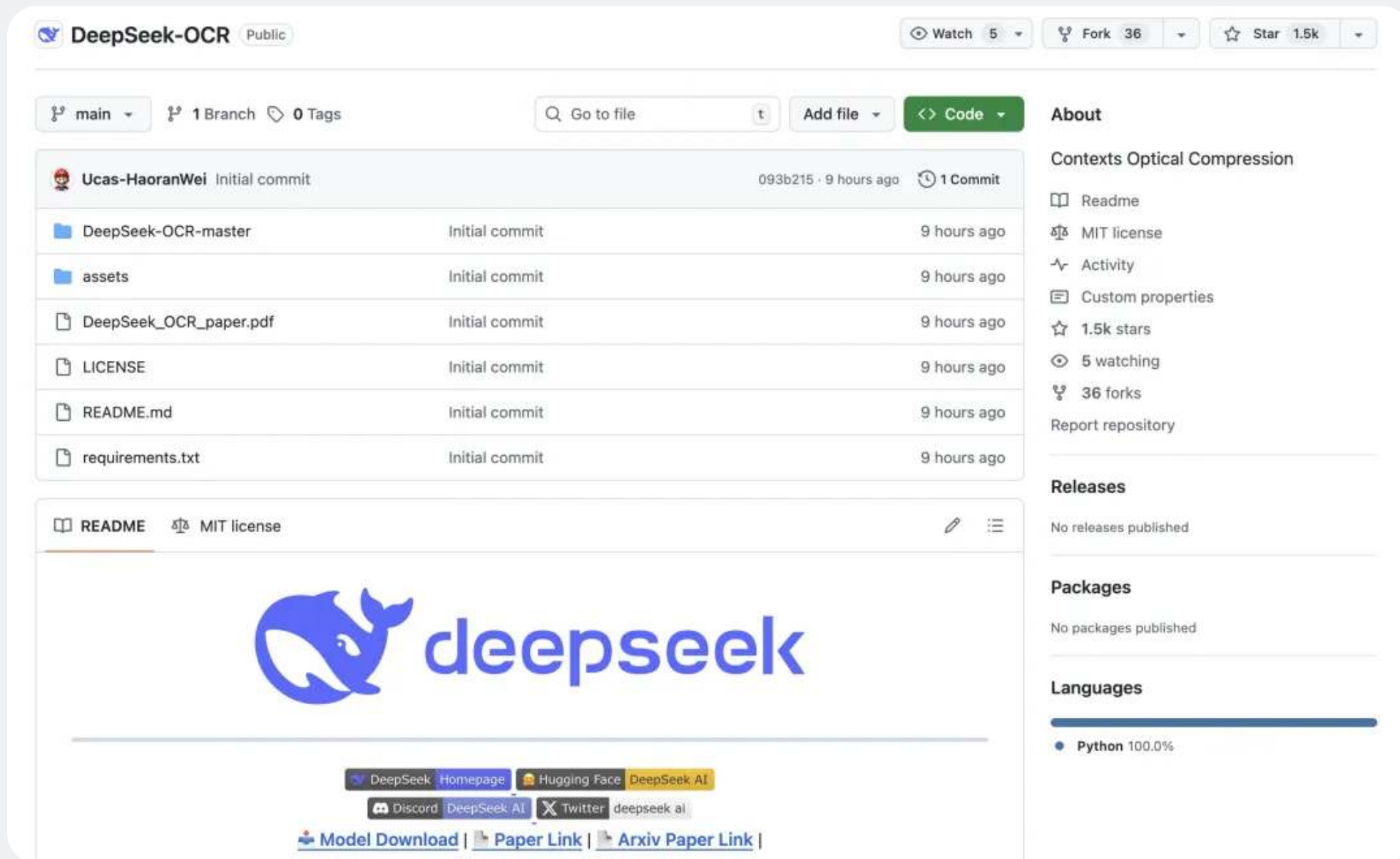
Vevo 3.1

← → 🔍 “孕育AGI的子宫” ---世界模型



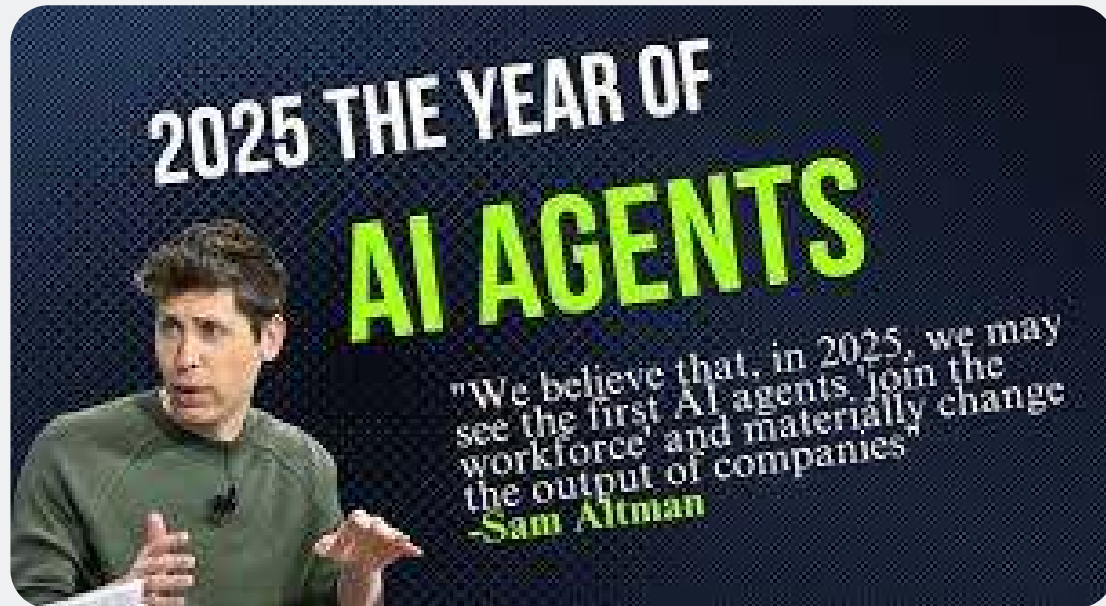
小问题+1：
这玩意真的做成“极致”状态，
元宇宙的叙事是不是就成了？

← → 🔍 🔍 视觉和语言的关系到底是什么？



小问题+1：
DeepSeek为什么出了一个OCR模型？
差不多时间智谱也出了，他们在解决的是什么问题？

← → 🔍 2025年是Agent元年……吗？ 以及会有几个元年？



要想实现AGI，达到能替代80%的人类白领工作的Agent，是需要非常漫长的时间，可能要5-10年甚至更久，除非有全新的颠覆性技术出现。



Flood Sung

知乎

论Agent的“持久战”

扫码 打开知乎 · 与我共读



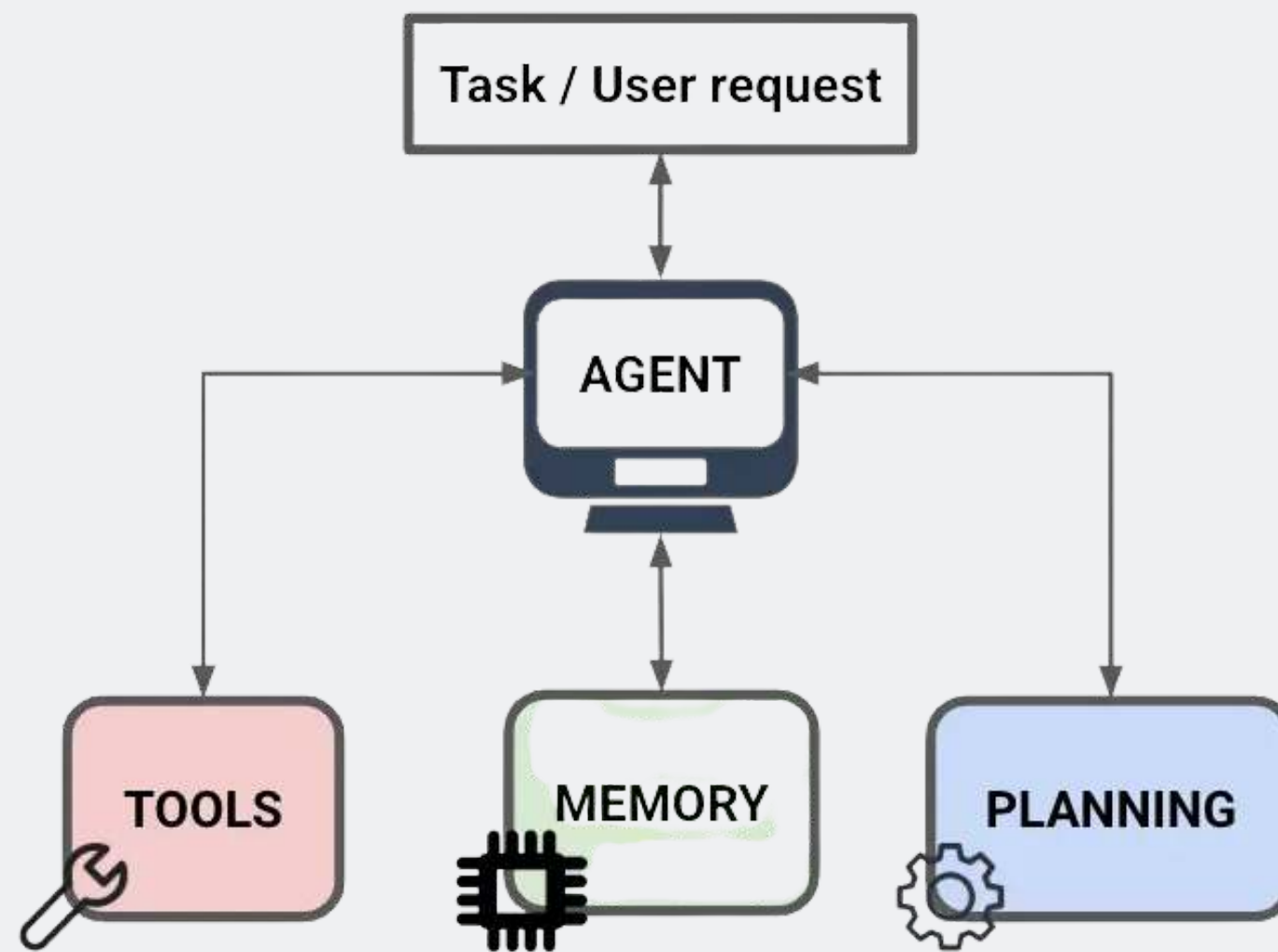
AI won't work like us -
for at least 10 years



Andrej Karpathy

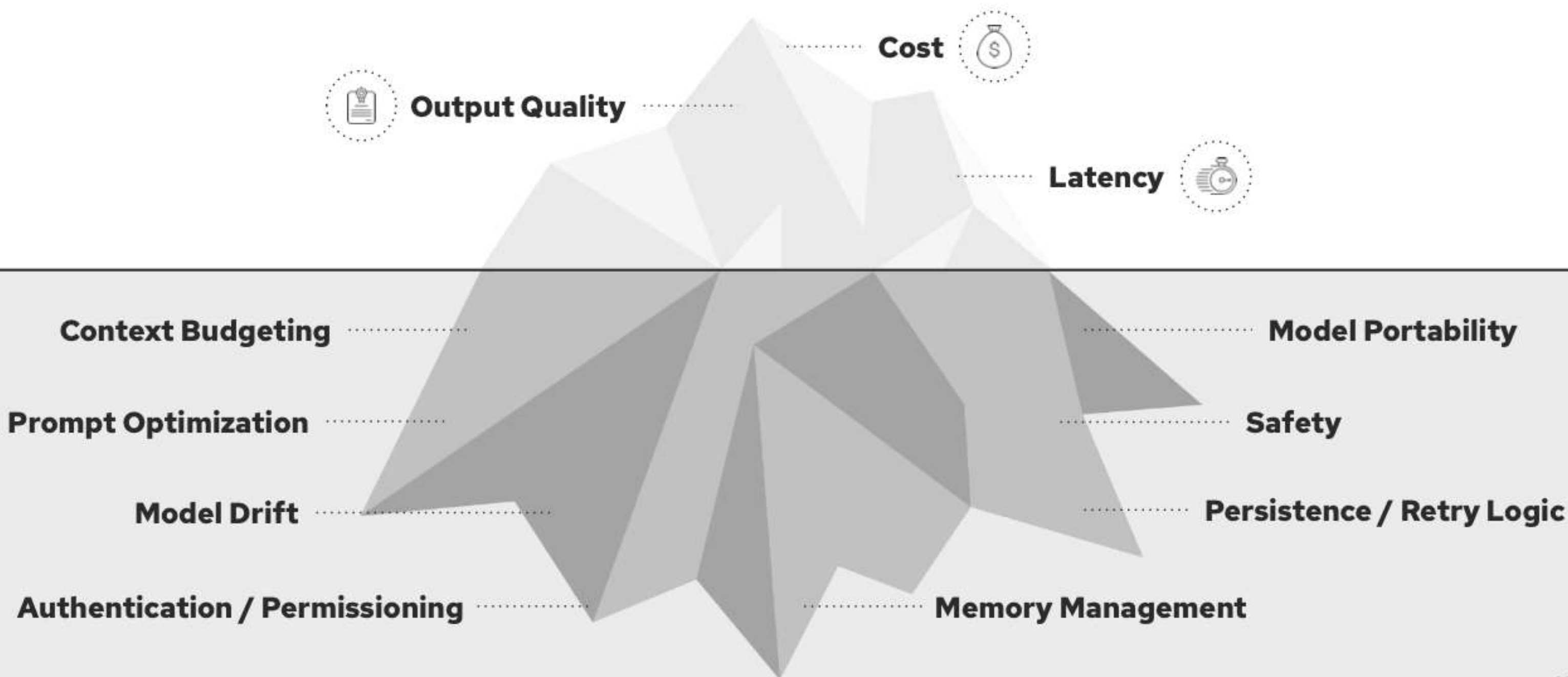
← → 🔍 🔍 为什么今天的Agent还和脑血栓患者一样?

AI Agent 现状



← → 🔍 🔍 因为Agent要做的事情就不只是Agent自己了

Tons of unsolved problems.



← → 🔍 这些“模式”，到底是技术范式还是产品范式？

Domain	Illustrative products & players		NON-EXHAUSTIVE Frontier lab product	
A Coding	GitHub Copilot, CURSOR, OpenAI Codex, replit, GEMINI, jules, CLAUDE CODE, Amp, KIRO, V2			
B Deep research	Gemini, GPT-4o, perplexity, Claude, Grok, Qwen3			
C Computer use	Gemini, GPT-4o, comet, SURFER, Browser Use, AI, Opera Neon			



← → 🔍 我们到底该如何看待这样一家其实已经创造了无数历史的公司

产品叙事

OpenAI是家什么公司?

- 一家5000亿美金的初创公司
- 能用1-2个描述就概括了么?
- 技术 or 产品?
- 还是全都要?
-

热门产品方向

- Chatbot
- Coding
- Agent
- Social
-

ARR的叙事

- ARR依然是有效手段吗?
- ARR之外还有什么指标?
- 毛利? 净利? RPO?
- 如果不增长怎么办?
-

← → 🔍 作为绝对绝对中心的**OpenAI**，这一年都发布了啥？

o3/o4系列

GPT 4.1

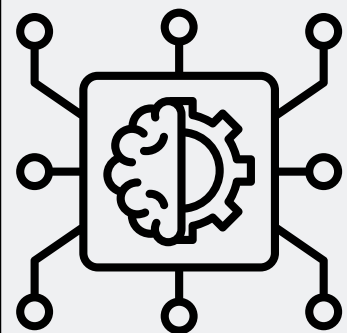
DALL.E 3

Realtime

GPT 5

Sora 2

.....



CodeX

Pulse

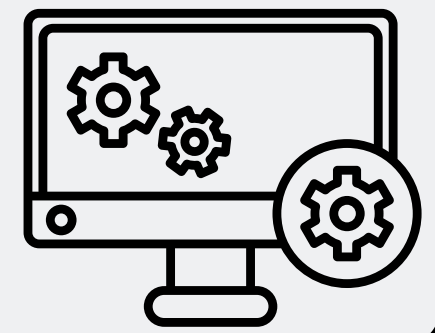
Agent kit

App inside

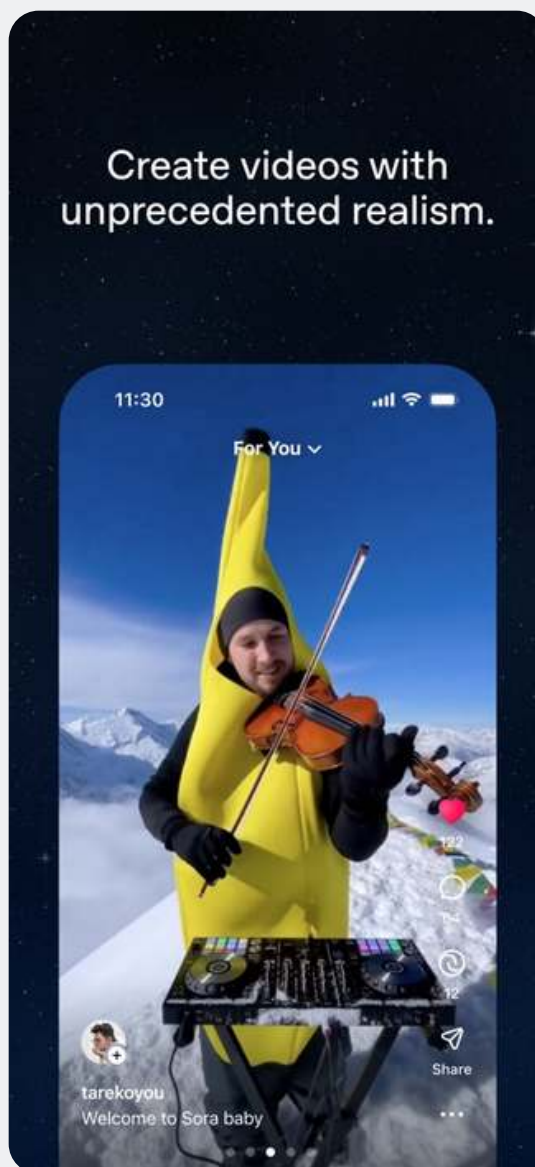
Sora APP

Atlas

.....

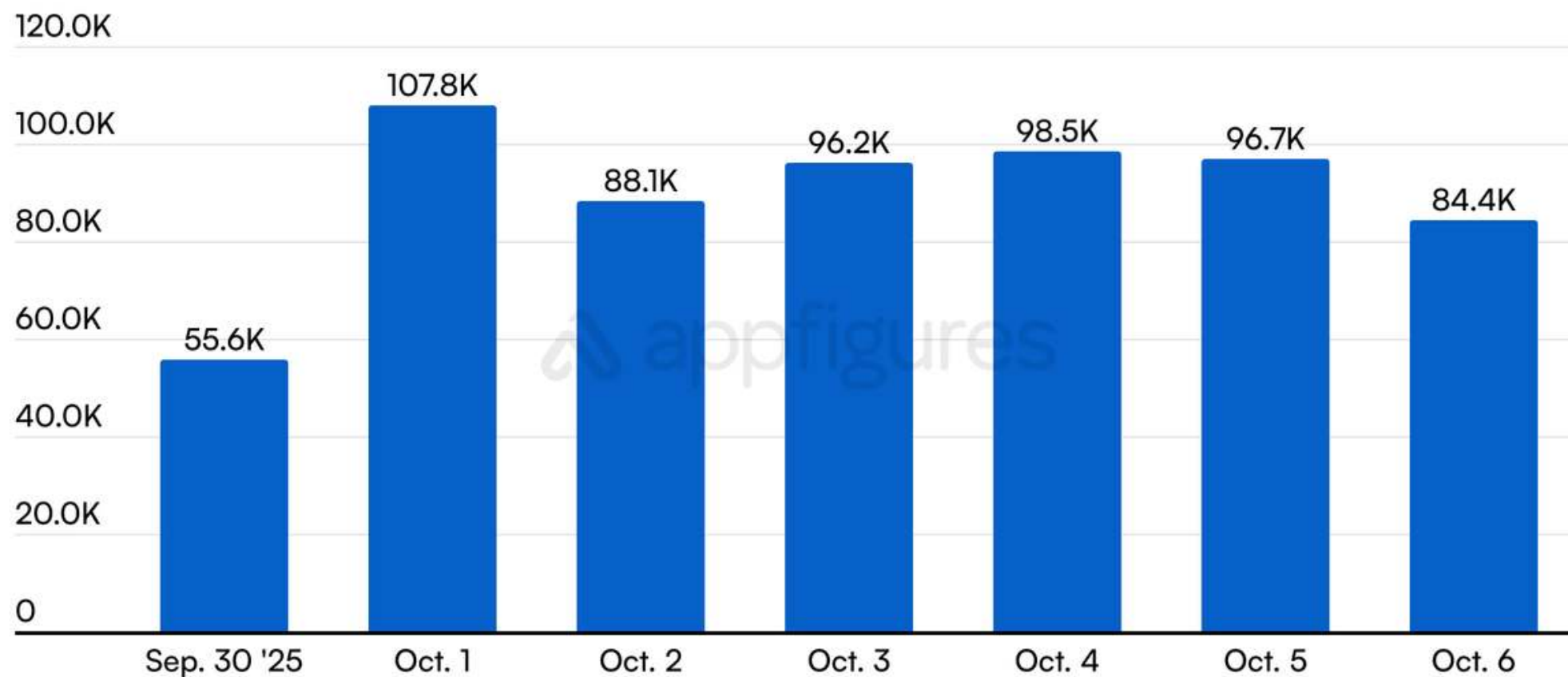


← → 🔍 世界当然不需要一个AI版的抖音，但OpenAI需要



Sora Just Beat ChatGPT's First Week Downloads

Sora Estimated iOS App Store Downloads

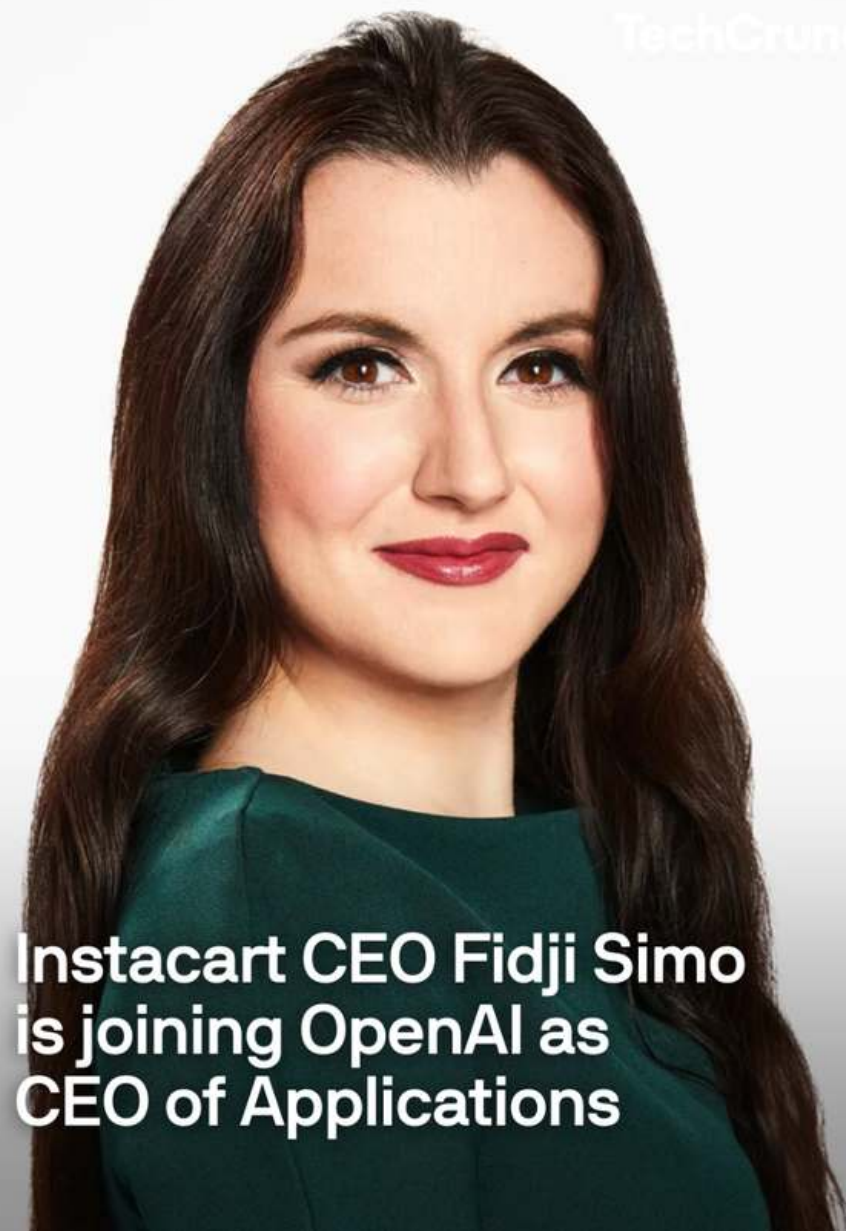


Source: Appfigures Intelligence



← → 🔍 当然这肯定有组织架构的调整，以及引发新的问题

TechCrunch



Instacart CEO Fidji Simo
is joining OpenAI as
CEO of Applications

“one in five OpenAI employees
are former Meta hires,
bringing with them Silicon
Valley’s most aggressive
scaling playbooks — and
raising questions internally
about culture, strategy, and
control.”

AJ Green
When Meta DNA Meets OpenAI’s Mission — Who
Wins?

AJSAI.SUBSTACK.COM

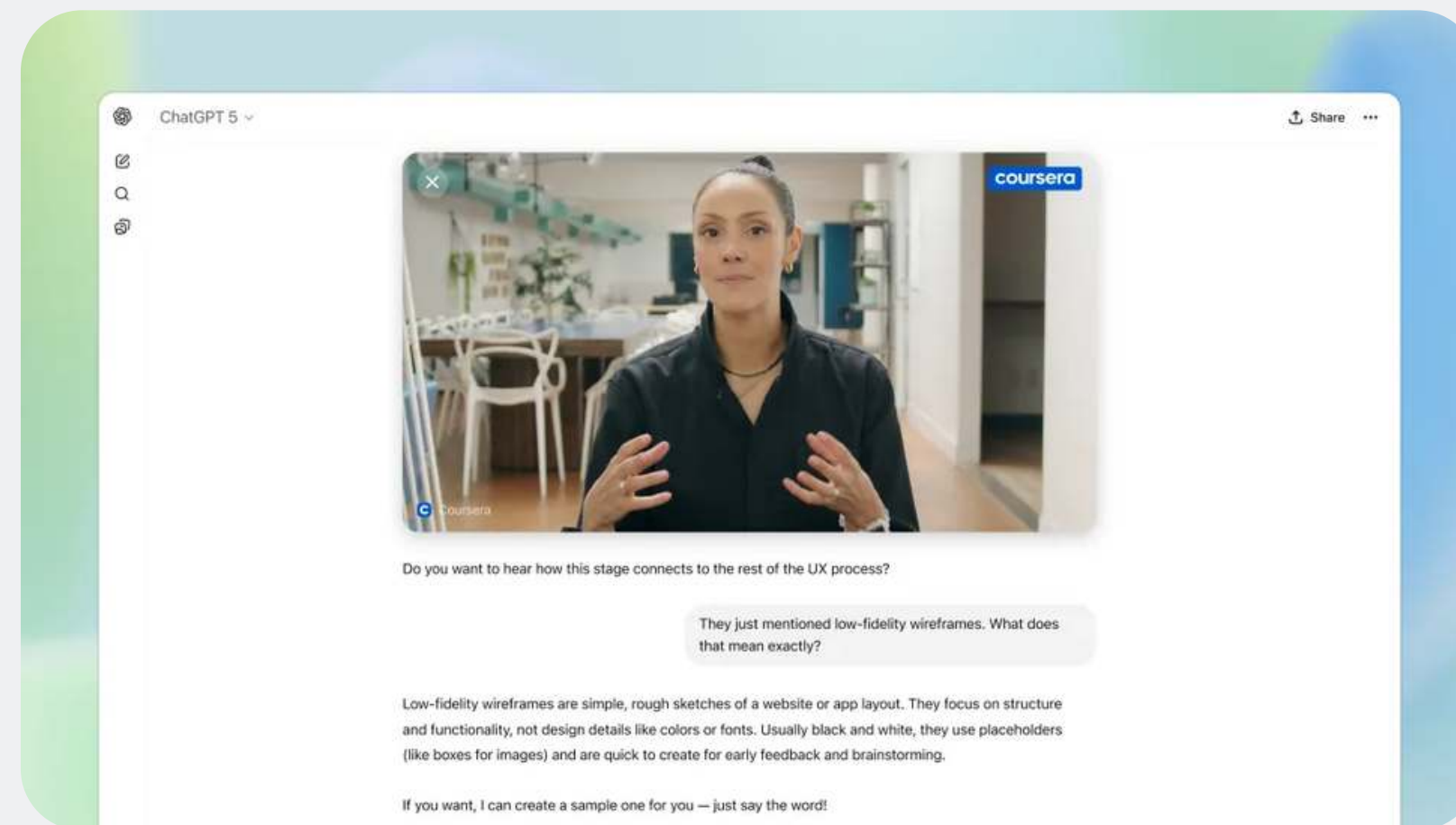
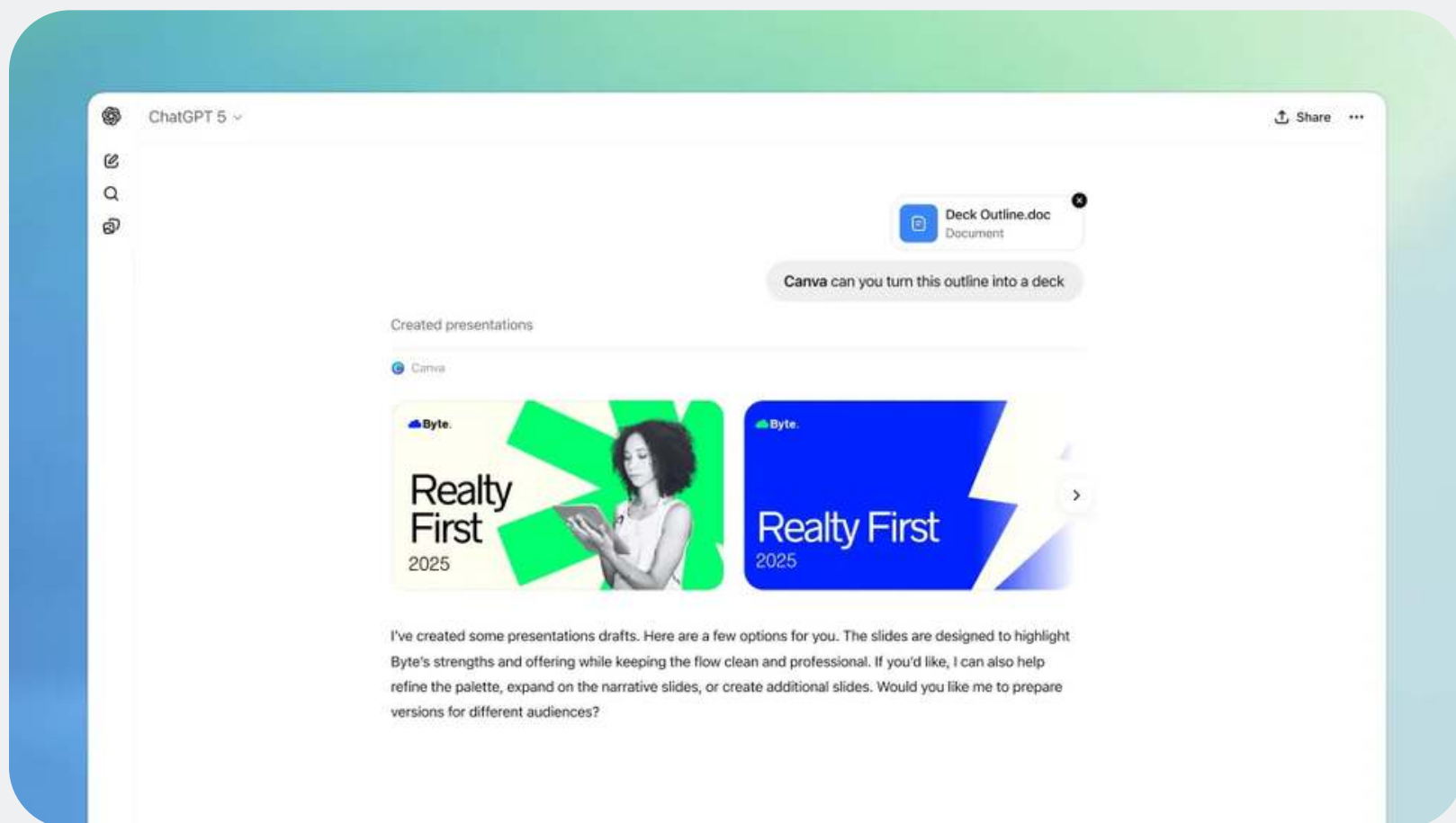


OpenAI Readies Itself for Its Facebook Era

By Kalley Huang, Erin Woo and Stephanie Palazzolo



← → 🔍 App inside ChatGPT = **All in one**



← → 🔍 ChatGPT凭什么呢?

ChatGPT has reached 800M weekly active users, Sam Altman also announced

ChatGPT's WAUs +14% to 800M in 2 months

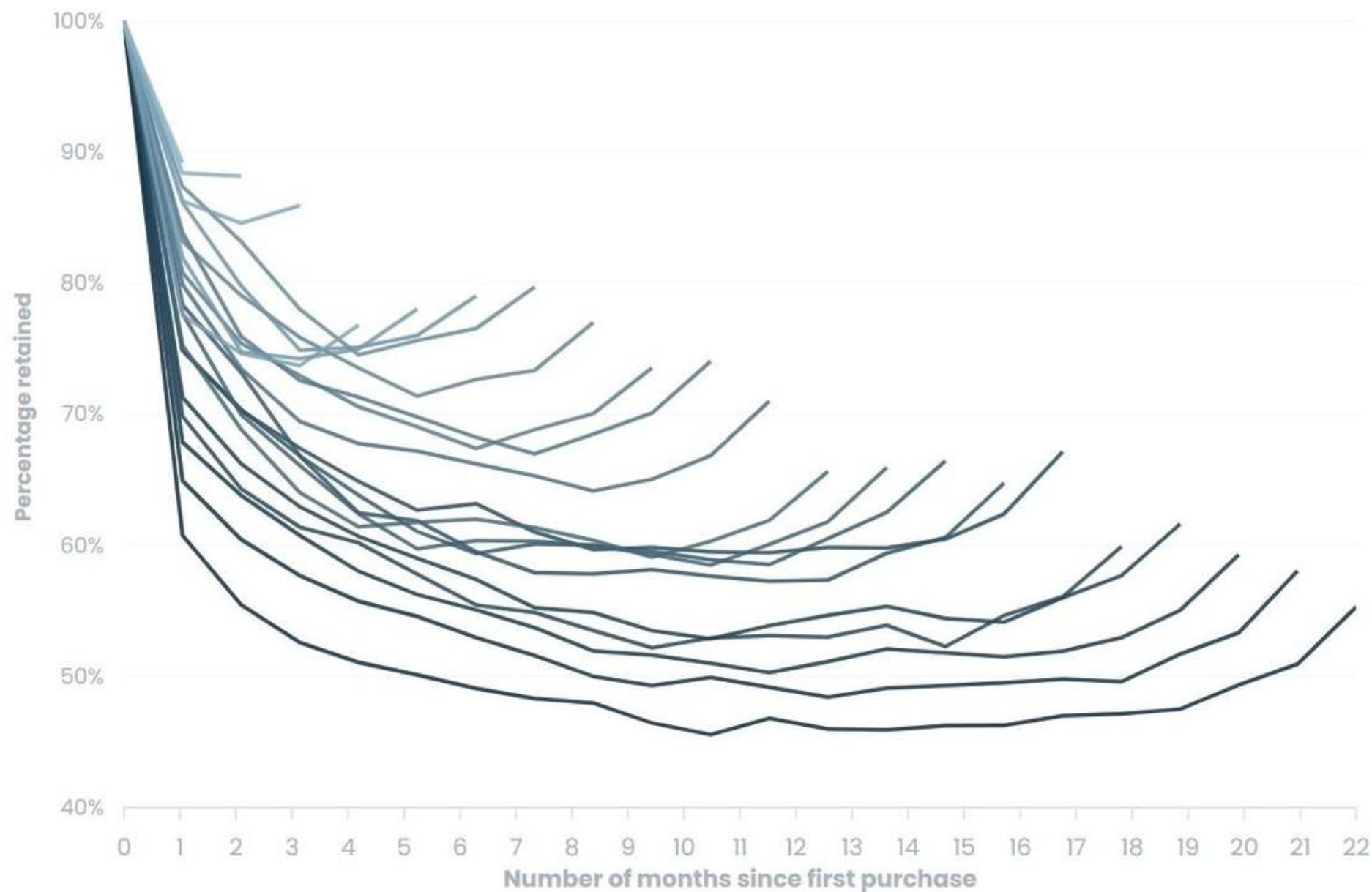
Weekly active users, in millions.



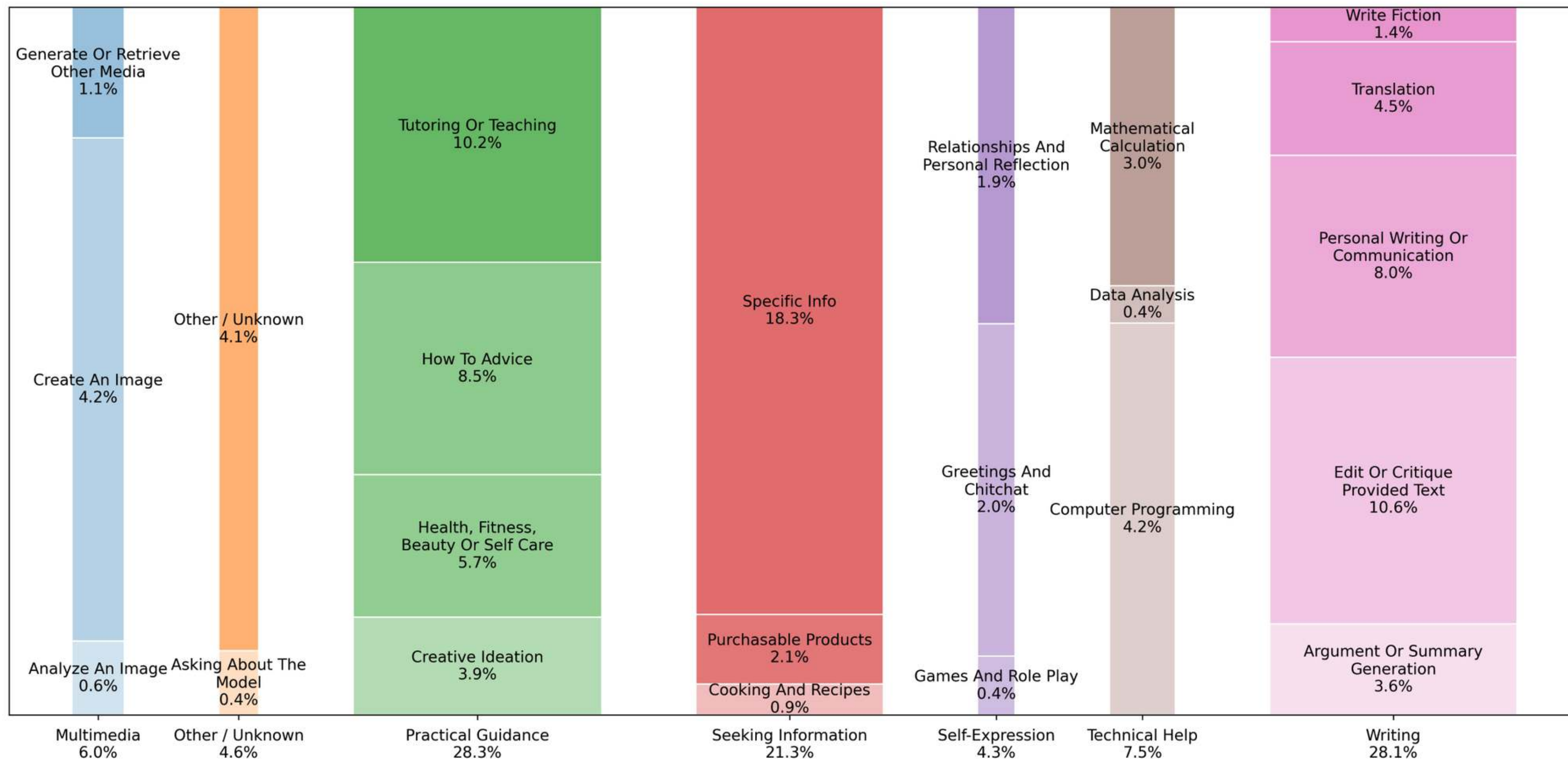
Source: OpenAI, Bay Area Times analysis.

Bay Area Times

Customer Retention (ChatGPT)

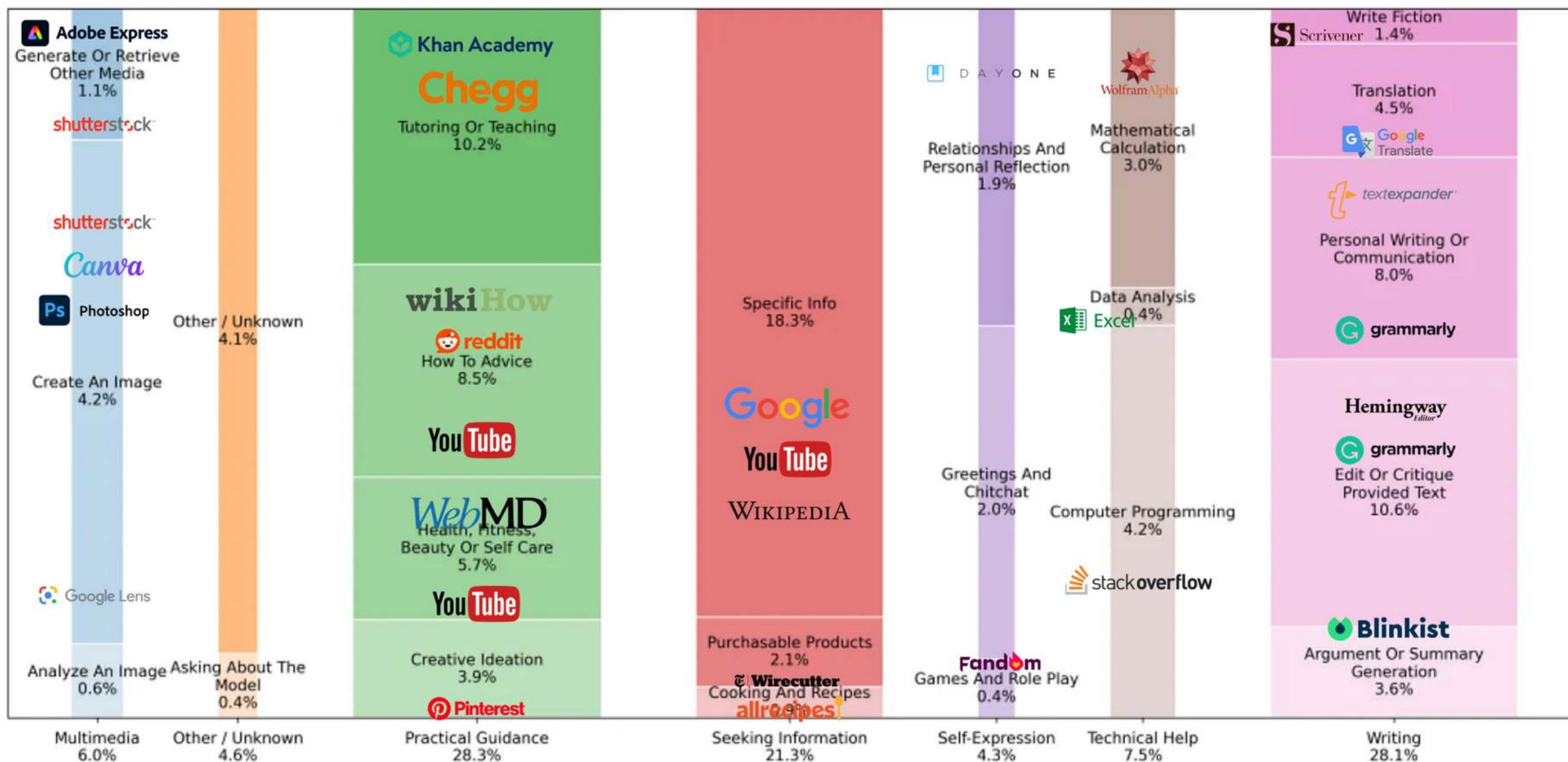


← → 🔍 每周8亿人到底用ChatGPT来干嘛？



← → 🔍 这个图如果重新画一下……

ChatGPT and the great bundling



@tanayj

← → 🔍 🔍 所以在今天，OpenAI到底是一家什么样的公司？

Product Parade

OpenAI's product ambitions have quickly grown, as the company looks to develop everything from humanoid robots to AI coding assistants

- **Collaboration tools** so multiple ChatGPT users can work together on projects and chat about them.

- **New AI that combines** traditional large language models with reasoning AI

- **ChatGPT "agents"** that let customers create and edit spreadsheets and presentations, and generate sophisticated reports

- **A web browser** combined with ChatGPT

- **A-SWE, an AI coding assistant** that replicates a senior software engineer and handles tasks that takes a human programmer hours or days

- **Software and hardware for robots**—possibly humanoid ones

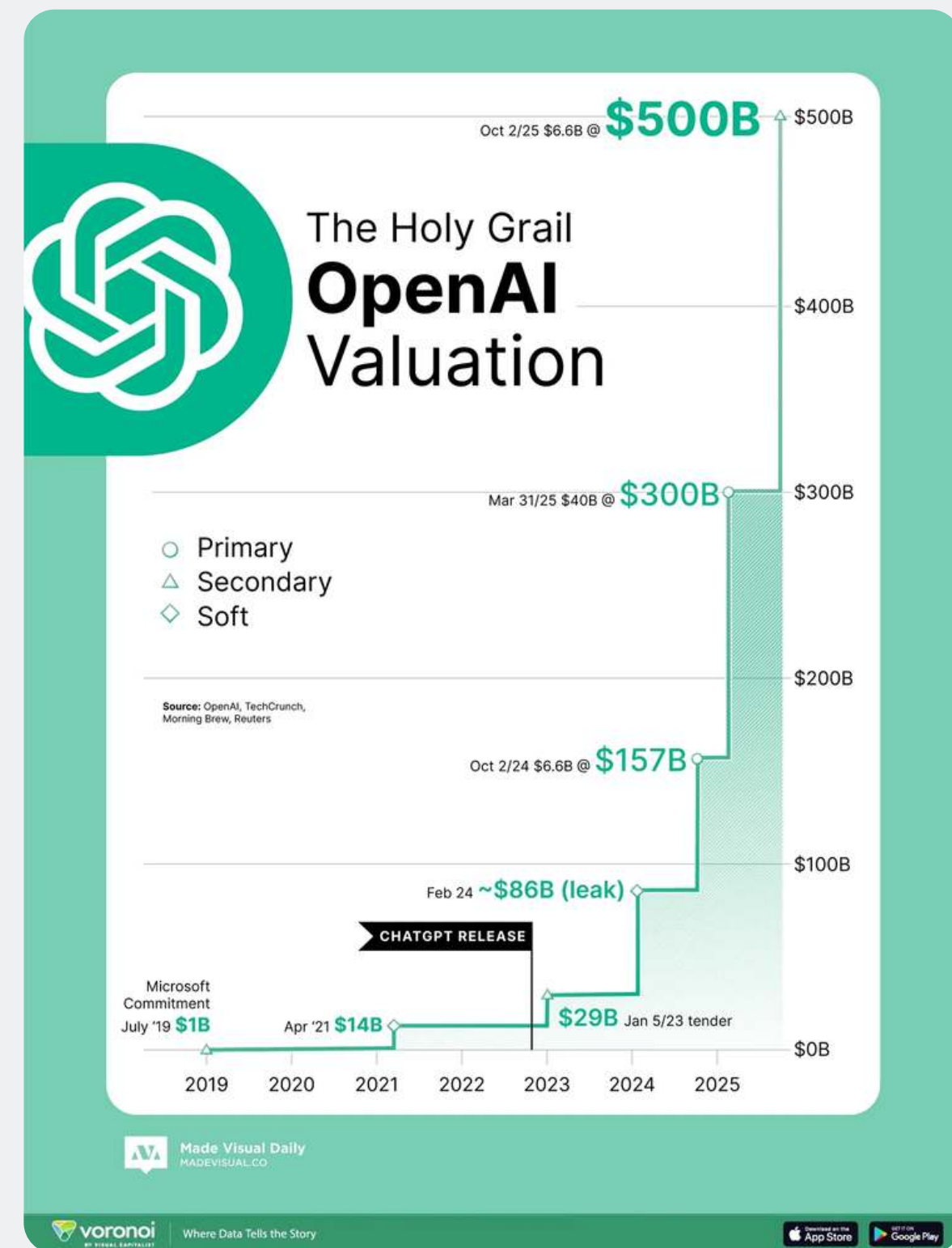
- **An AI-powered personal device** via an acquisition of Jony Ive and Sam Altman's device startup

- **Social media feed in ChatGPT** where users can share how they're using the chatbot to solve problems or create images

- **Shopping recommendation features in ChatGPT**, which provide personalized product suggestions and let users buy them through ChatGPT

- **Custom models**, which incorporate a customer's unique data and business context to power bespoke internal AI tools

- **Music-generating AI**, which could help users generate music from scratch



← → 🔍 虽然产品的发布很多，但在成本占比上……

EPOCH AI

R&D compute
\$5 billion

\$4.5 billion

Other R&D compute

Includes basic research, experimental/derisking runs for final training runs, and unreleased models.

Inference compute
\$2 billion

Excludes Microsoft inference to run OpenAI models for Microsoft products.



小问题+1：

所以OpenAI就是
模型研究实验室+产品公司？
还会有其他的么？

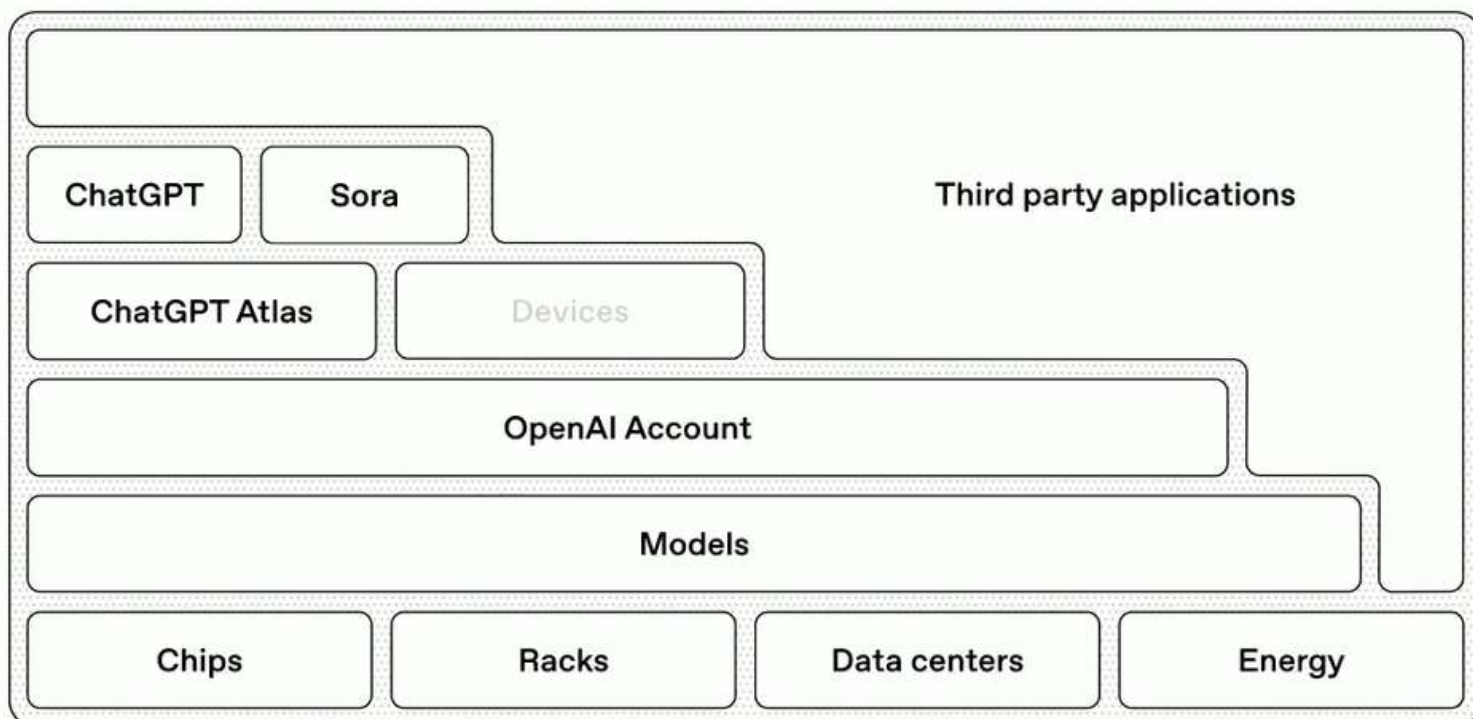
\$400 million

GPT-4.5 final training run

\$80 million

Other models

← → 🔍 这页的制作时间是10.29，截图来自和微软新合作官宣后的Sam直播



一个能预见人们借助AI构建卓越服务的新纪元
getting into a world where we can see that people
are going to be able to build incredible services



景，那就是：我们要创造工具，然后希望人们使用这些工具去开创未来。

我们为OpenAI的研究、产品与基础设施设想了三大核心支柱：我们必须成功完成构建AGI所需的研究；必须打造一个既便捷又强大的使用平台；同时还要建设足够的基础设施，让人们能够以低廉成本使用所有他们喜爱的卓越AI技术。

我们对整体架构的构想：最底层是芯片组、机架及相关系统，它们所在的数据中心以及能源支撑。在此基础上，我们训练模型，然后构建开放平台。

我们现在推出了名为Atlas的浏览器，未来几年还将推出便携设备，让



腾讯科技

1.4万亿投资、GPT-6、IPO进程，奥特曼回应“新OpenAI”的一切 | 1小...



← → 🔍 分类依然很难，想到哪算哪吧.....

产品叙事

OpenAI是家什么公司?

- 一家5000亿美金的初创公司
- 能用1-2个描述就概括了么?
- 技术 or 产品?
- 还是全都要?
-

热门产品方向

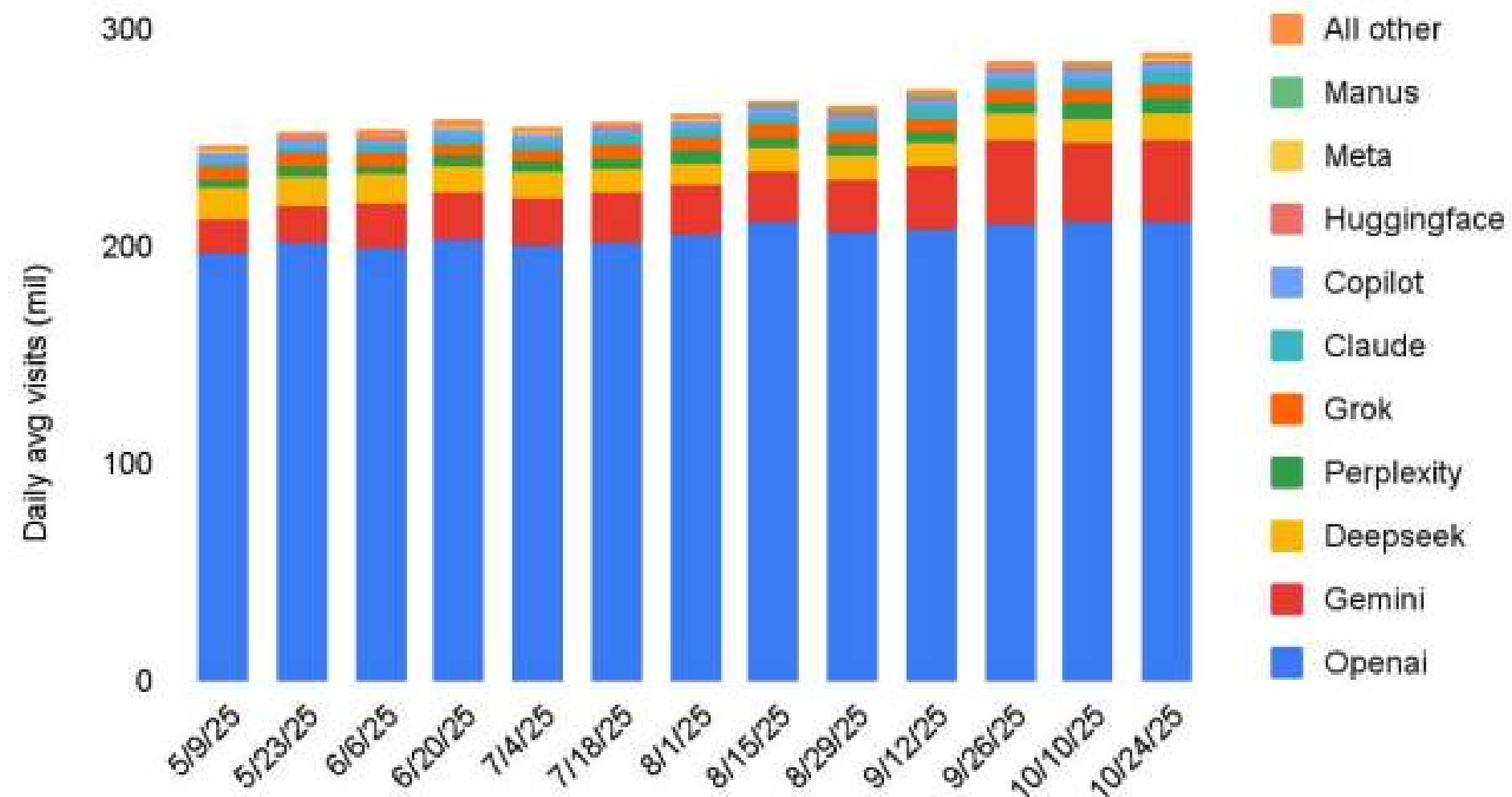
- Chatbot
- Coding
- Agent
- 那些数据
-

ARR的叙事

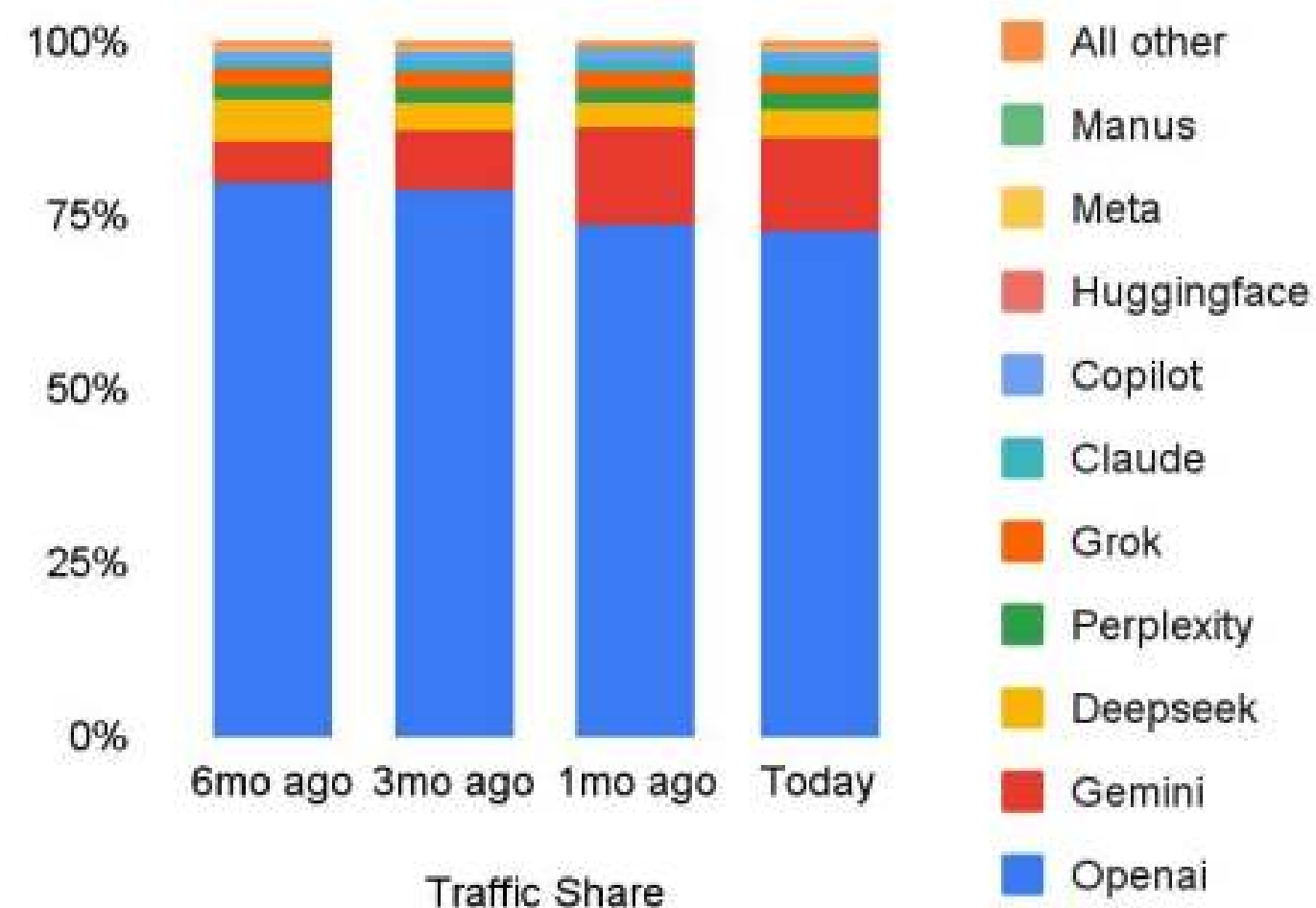
- ARR依然是有效手段吗?
- ARR之外还有什么指标?
- 毛利? 净利? RPO?
- 如果不增长怎么办?
-

← → 🔍 Chatbot的战役已经打完了…吗?

General AI Tools Traffic



General AI Traffic Share

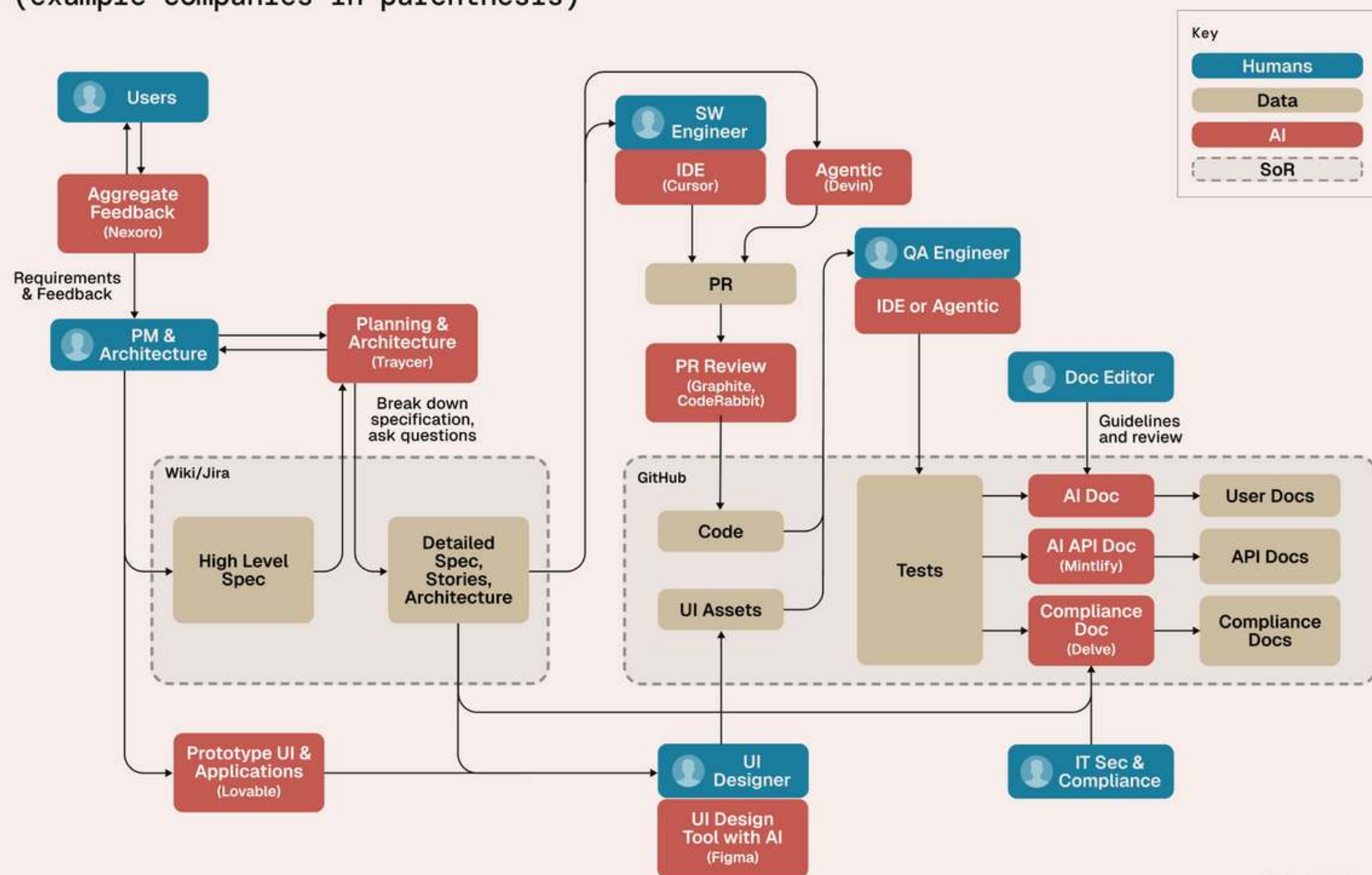
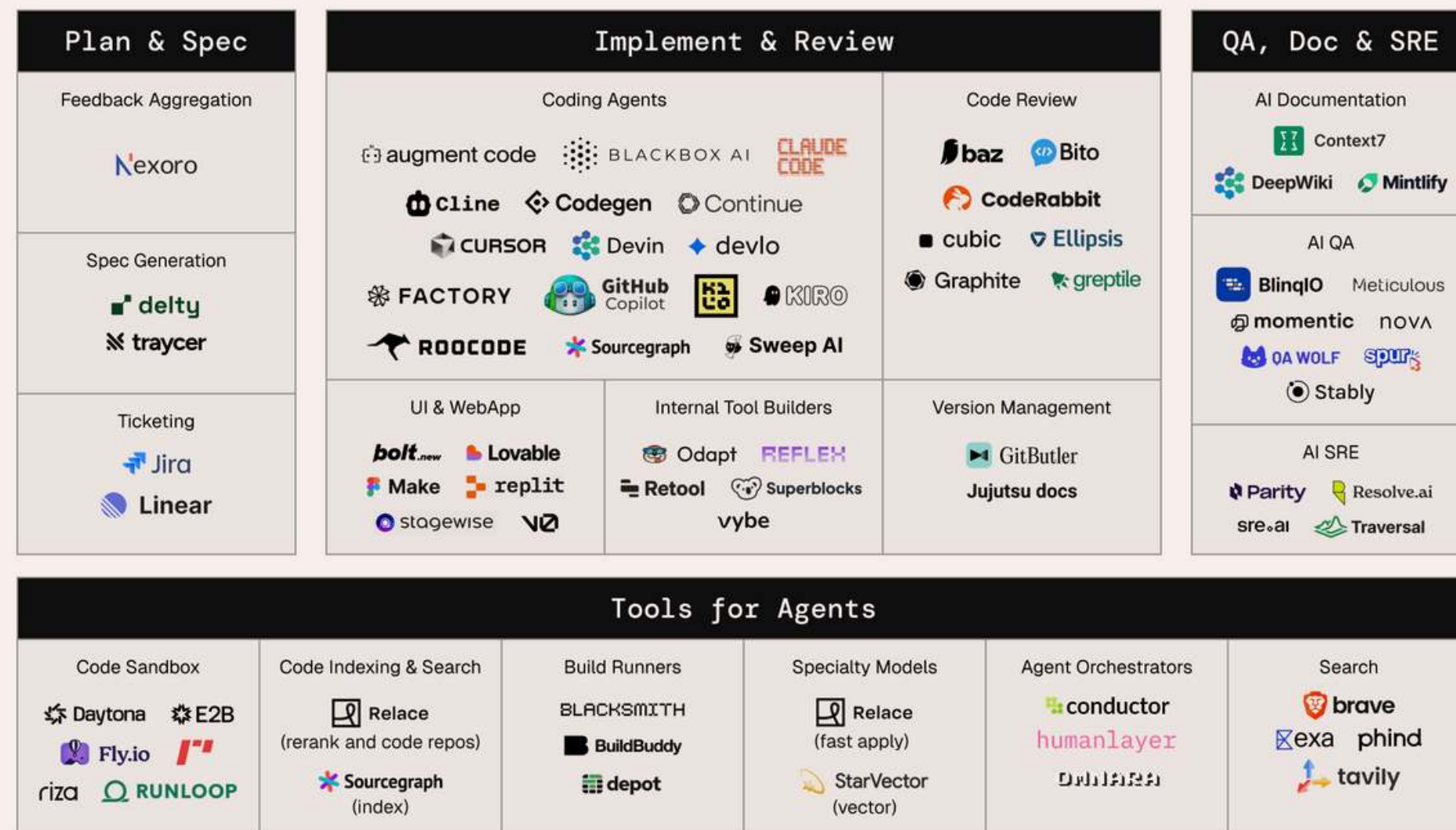


← → 🔍 Chatbot的战役已经**打完了…吗?**

12wk Change	5/23	6/6	6/20	7/4	7/18	8/1	8/15	8/29	9/12	9/26	10/10	10/24
Openai	26%	23%	23%	6%	3%	4%	5%	4%	3%	5%	5%	3%
Gemini	58%	95%	88%	65%	63%	51%	39%	20%	32%	78%	64%	69%
Deepseek	-26%	-20%	-20%	-26%	-27%	-27%	-28%	-22%	-17%	-8%	-8%	12%
Perplexity	0%	1%	12%	13%	12%	26%	14%	17%	20%	37%	66%	39%
Grok	64%	-2%	-25%	-27%	3%	9%	7%	14%	16%	6%	-13%	5%
Claude	14%	12%	14%	19%	26%	33%	46%	42%	34%	35%	47%	56%
Copilot	30%	18%	16%	-8%	3%	6%	0%	11%	12%	19%	15%	13%
Huggingface	-24%	-12%	-9%	-5%	-2%	3%	12%	3%	-6%	-4%	-14%	-12%
Meta	62%	68%	79%	38%	-17%	-22%	-21%	-21%	-21%	-15%	102%	82%
Manus	na	-46%	-20%	-16%	-1%	25%	2%	-11%	-19%	-14%	-23%	-17%

← → 🔍 🔍 **AI-Coding**, 突飞猛进的一年……

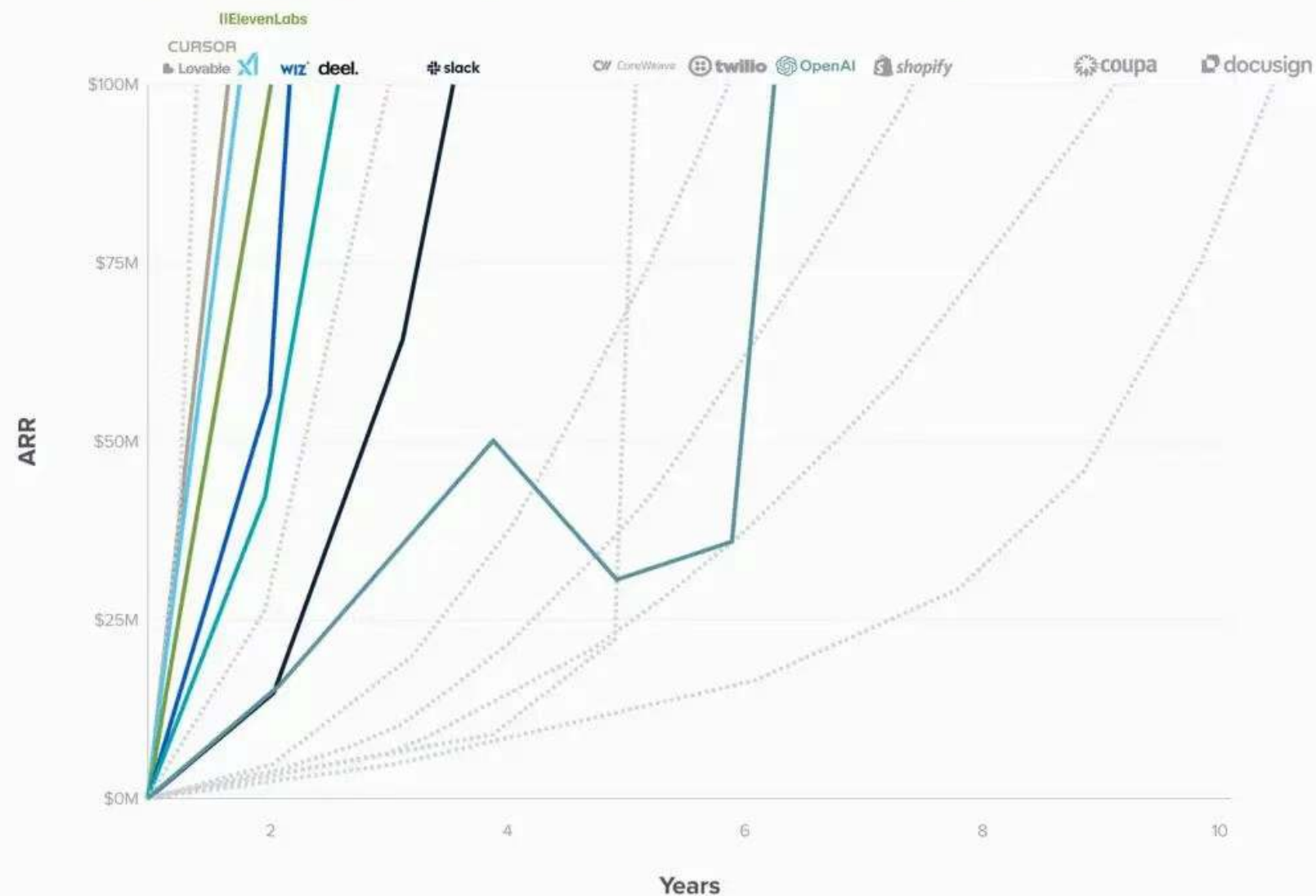
AI Software Development (example companies in parenthesis)

A16Z
INFRAA16Z
INFRA

← → 🔍 🔍 我们一次次的在这里见证**增长神话**

Top AI startups rapidly achieving \$100M ARR milestone

a16z Growth

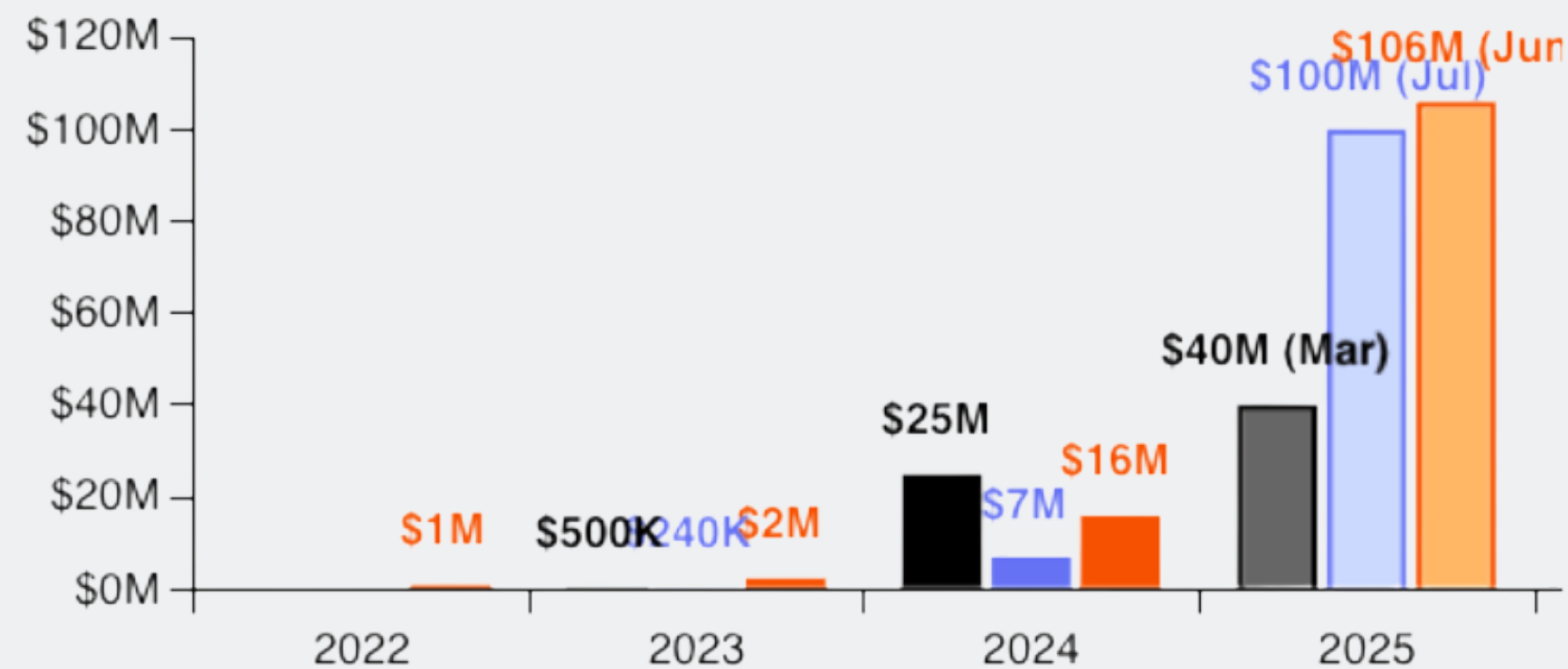


Source: a16z Growth and public data.

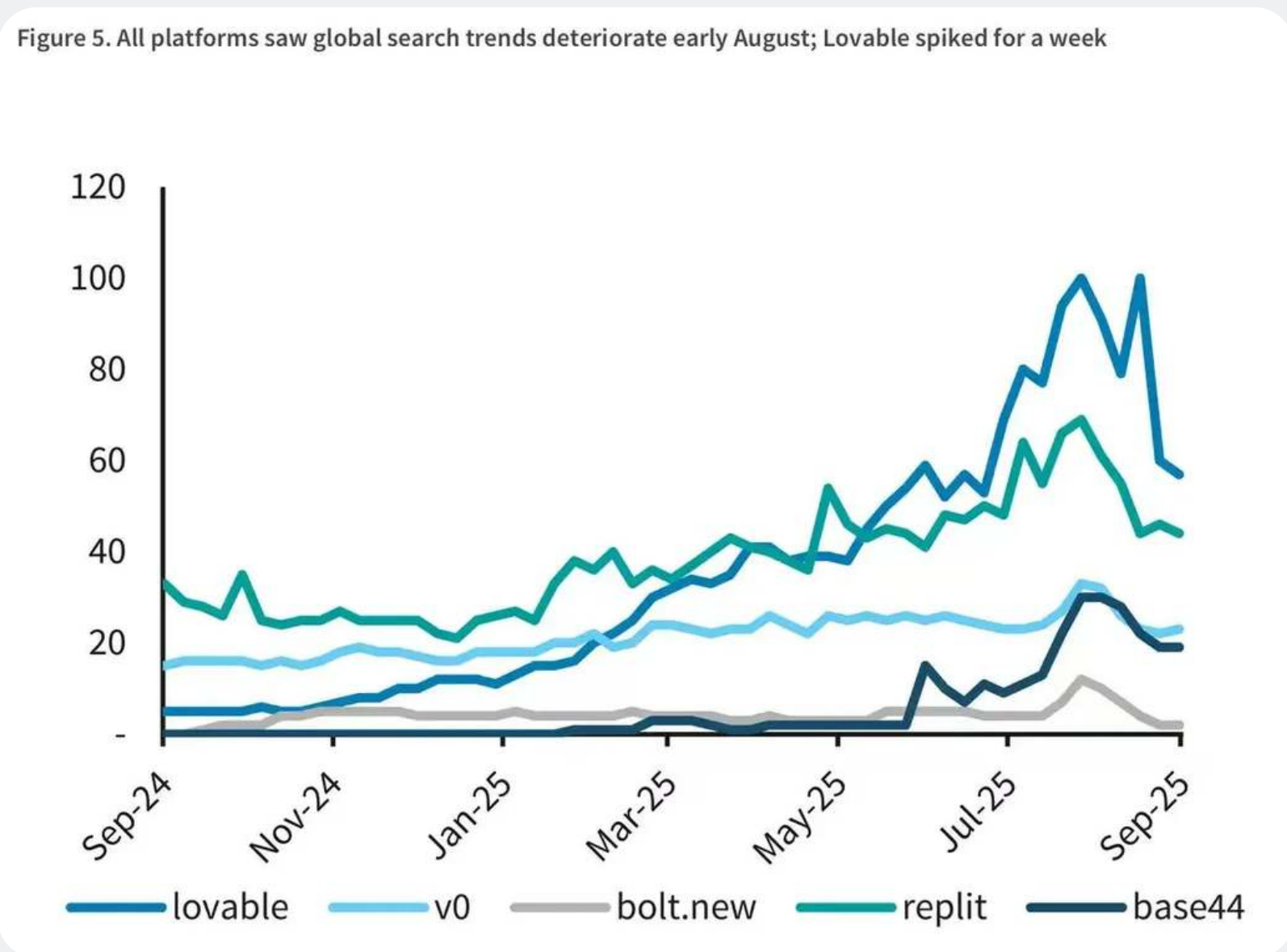


bolt vs. **Lovable** vs. **replit** Annual recurring revenue

■ Bolt.new: ARR (\$M) ■ Lovable: Revenue (\$M)
■ Replit: ARR (\$M)



← → 🔍 🔍 但最近，事情似乎出现了变化……

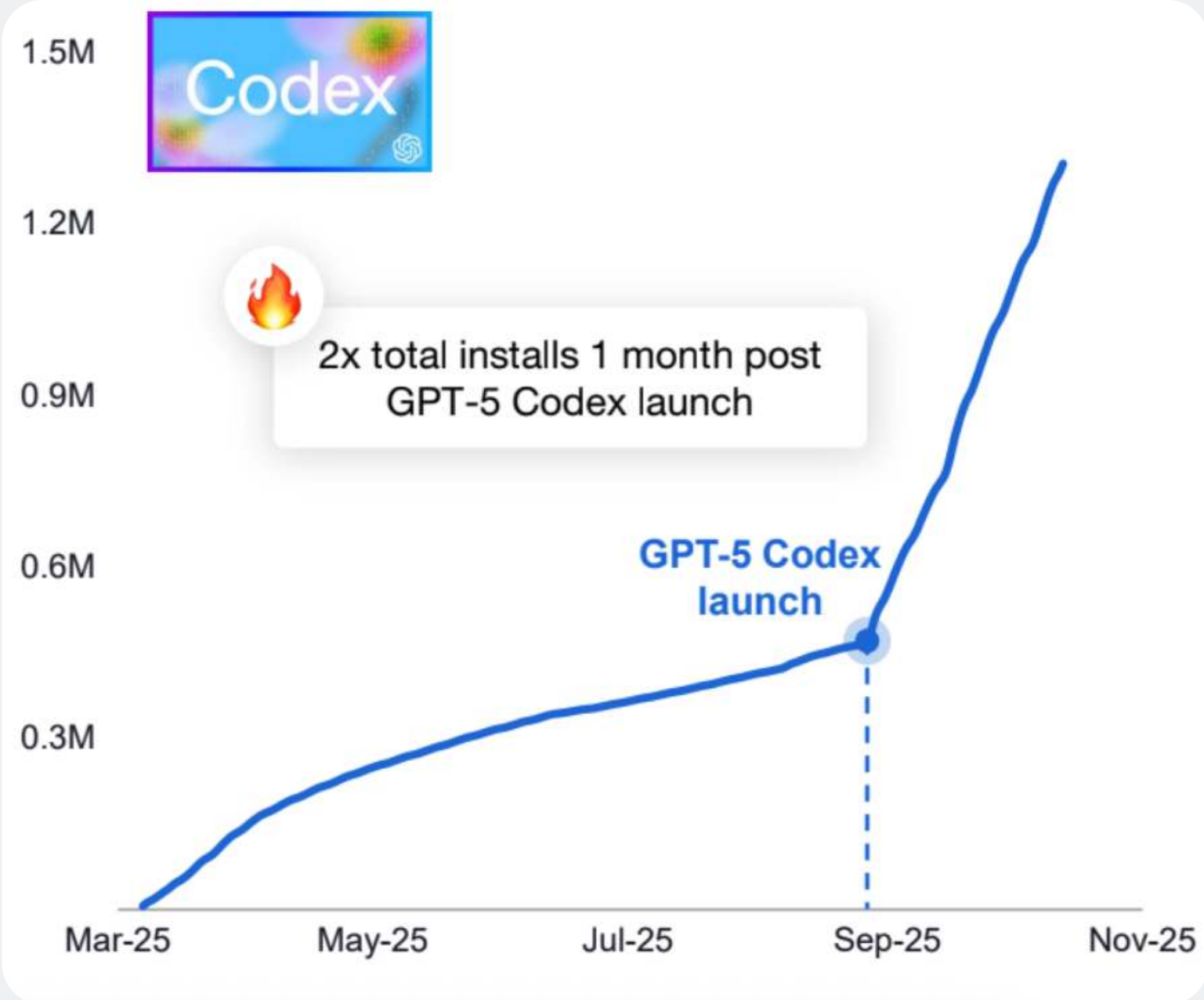
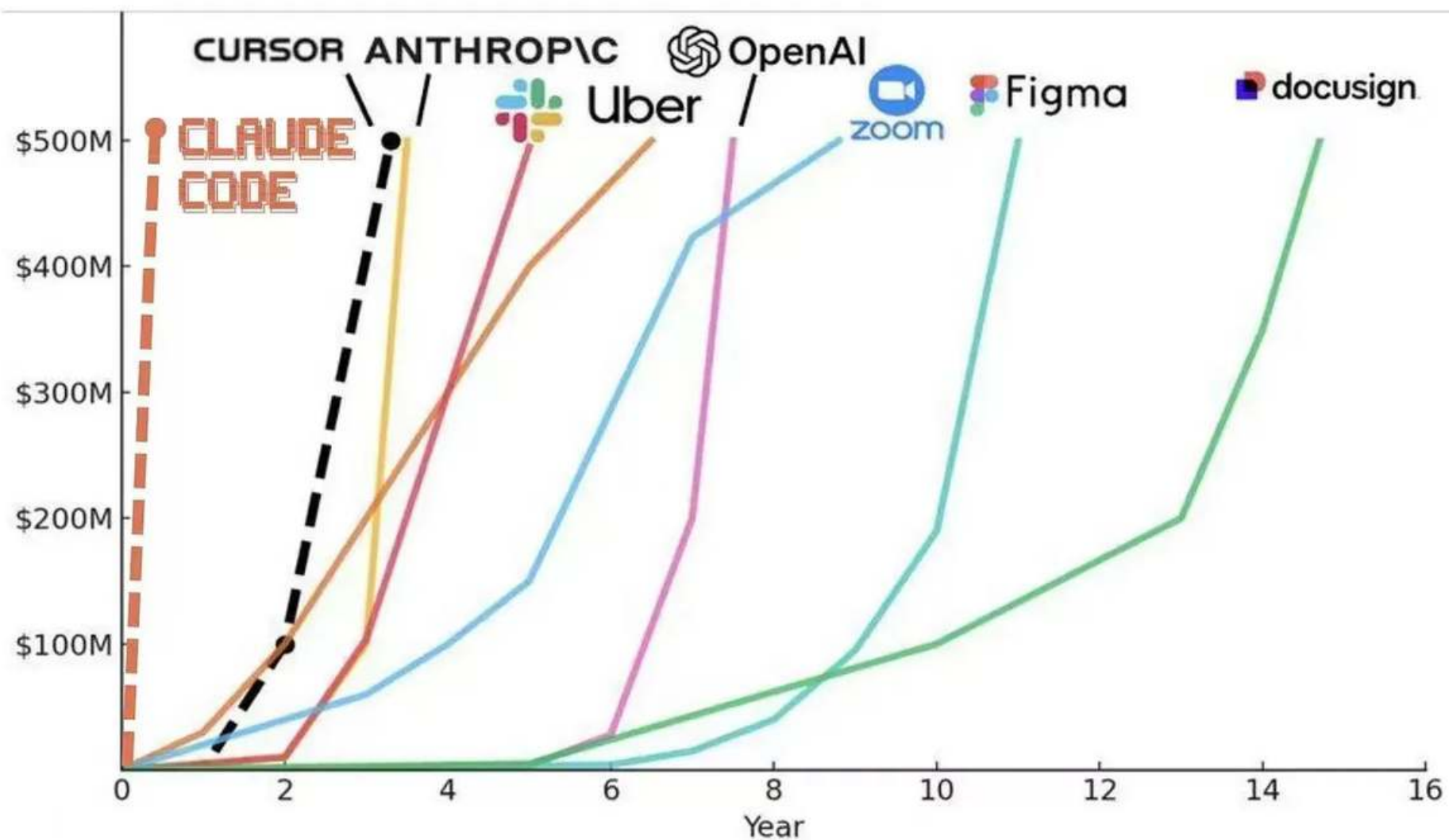


数据下滑，还怎么继续把故事讲下去？

12wk Change	5/23	6/6	6/20	7/4	7/18	8/1	8/15	8/29	9/12	9/26	10/10	10/24
Cursor	55%	25%	3%	-7%	-1%	-14%	-9%	-22%	-29%	-19%	-18%	-11%
Lovable	148%	72%	65%	8%	-41%	-24%	-26%	-35%	-50%	-37%	3%	-15%
Replit	40%	45%	29%	21%	15%	25%	10%	-1%	1%	-9%	-18%	-31%
Base44	950%	29%	13%	172%	506%	647%	757%	1161%	568%	225%	95%	73%
Bolt	12%	21%	33%	61%	34%	-2%	-9%	-23%	-35%	-45%	-31%	-29%
Windsurf	96%	34%	18%	-15%	-28%	-37%	-41%	-33%	-29%	-17%	-30%	-22%
Tabnine	8%	0%	-19%	-24%	-12%	-7%	7%	14%	15%	8%	-20%	-14%
Cognition	-17%	-20%	200%	60%	180%	100%	100%	100%	11%	0%	-36%	-20%
V0	39%	21%	3%	-5%	-13%	-4%	-40%	-90%	-91%	-94%	-95%	-96%
All other	28%	59%	26%	11%	-10%	-1%	-2%	-26%	-8%	-6%	-10%	-15%
DevOps & Code Completion total	66%	41%	29%	9%	-10%	-7%	-10%	-22%	-31%	-25%	-15%	-19%

← → 🔍 Coding当然是大市场，但**最后的赢家**会是谁？

Years from founding to \$500M ARR



通用Agent是不是我们这边的“独有说法”？

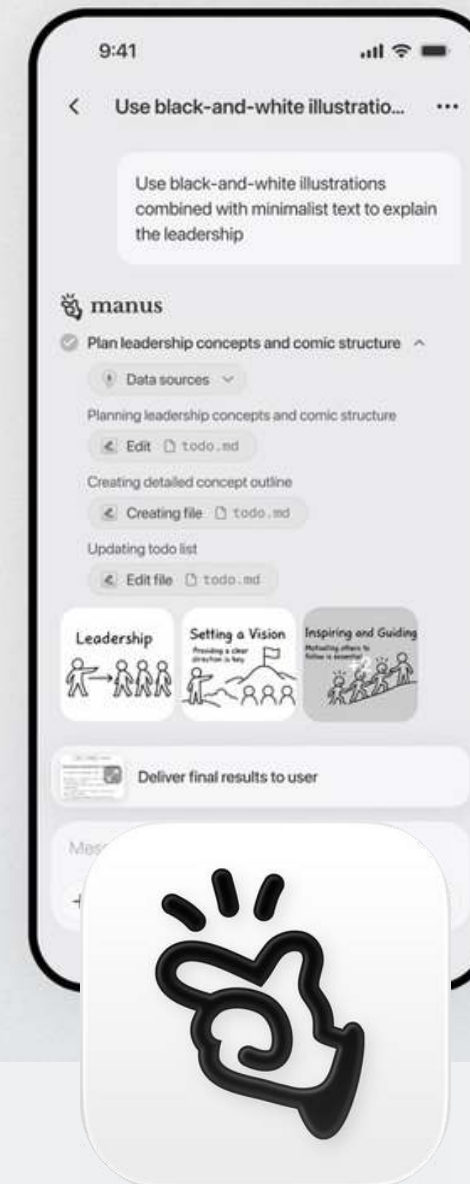
Will's GenAI产品数据 通用Agent



更多

排行	产品名	分类	aiwatch.ai	9月 访问量 (万Visit)	相对8月 变化
1	Manus	通用自动化Agent		1,388	-18%
2	Genspark	通用自动化Agent		876	4%
3	扣子空间	通用自动化Agent		290	8%
4	Skywork	通用自动化Agent		177	65%
5	MGX	通用自动化Agent		119	70%
6	Lindy	通用自动化Agent		86	-19%
7	天工AI	通用自动化Agent		59	-5%
8	MiniMax Agent	通用自动化Agent		52	-35%
9	flowith	通用自动化Agent		42	-9%
10	Activepieces	通用自动化Agent		28	34%

Works For You
Your personal AI Agent that
delivers direct results



Genspark

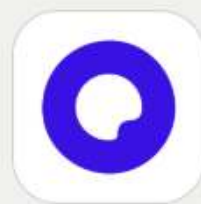
BEST AI
Search
Lightning-fast
Browser
Powerful AI
Agents



← → 🔍 🔍 浏览器战争升级

不知道你们是否想过

浏览器为什么又成了“兵家必争之地”？



那些家伙们

Google是否拆分Chrome可能真的会影响
到Google的AI大业，
因为Safari内置搜索更换已经在影响了.....

那些创业公司们

大模型公司、AI搜索公司、Agent公司、
纯新一代的浏览器公司，似乎都要疯狂的
弄浏览器，图什么呢？

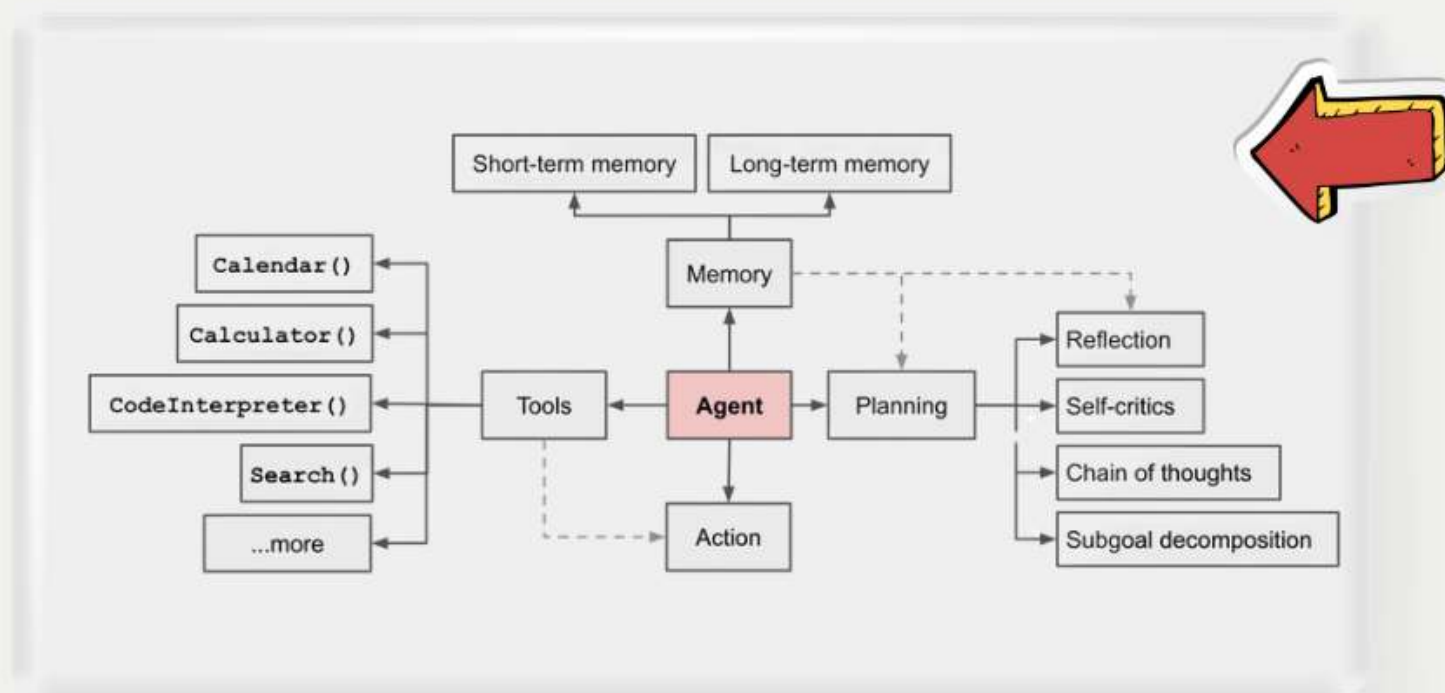
中国的“巨头”们

夸克成了阿里体系内2C-AI应用的“排头兵”
侧面激发了在腾讯内部可能已经几乎处于
躺平状态的QQ浏览器.....

NEXT

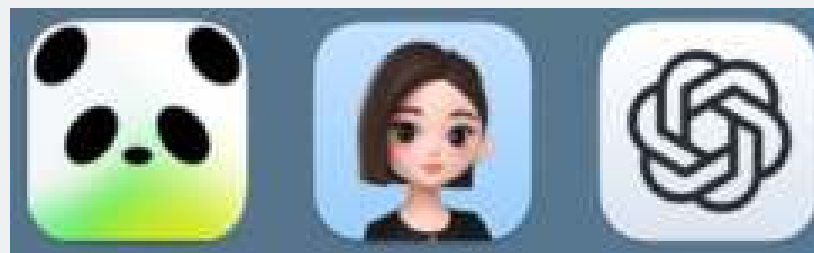
我们似乎需要给AGENT提供一个稳定、可拓展的

运行环境，然后“策马狂奔”



NEXT

这是我的程序坞



那这些呢？



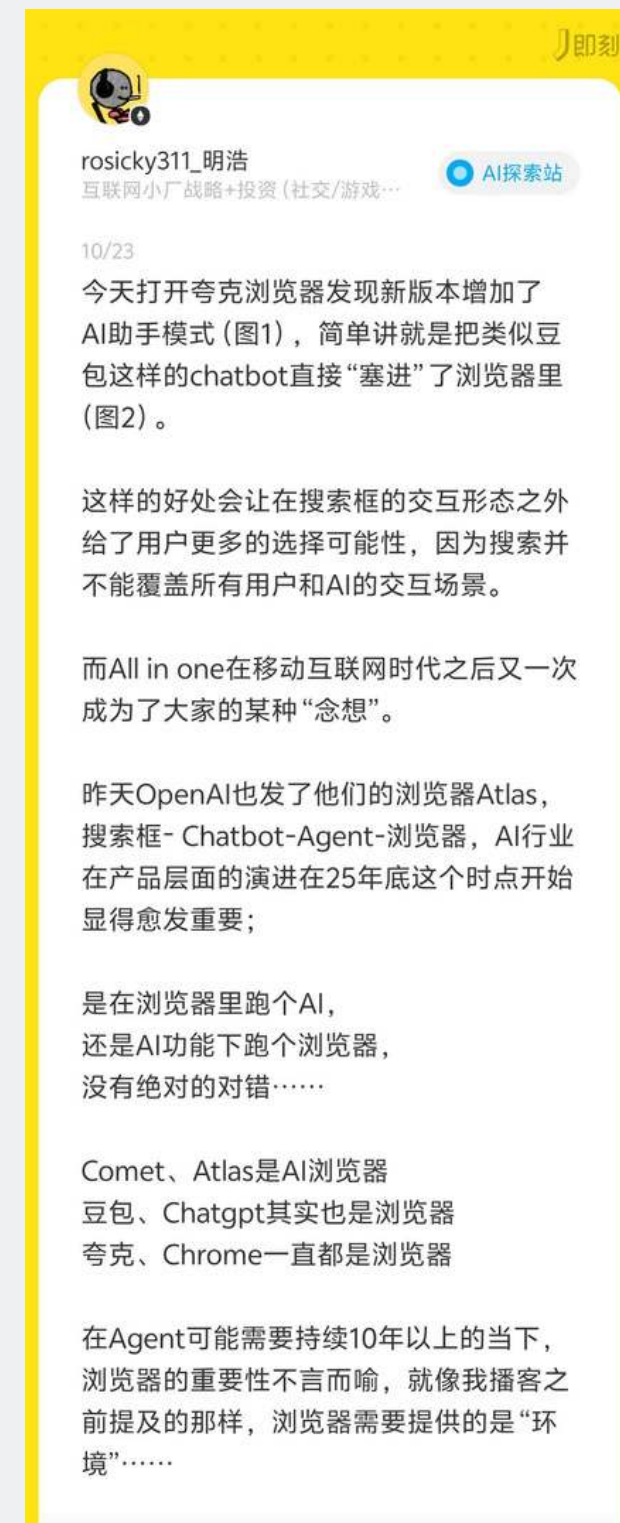
这些肯定是浏览器了

← → 🔍 Chat - 搜索 - Agent - 浏览器，整个就是“大乱斗”……

Agent里面跑个浏览器



浏览器里跑个Agent



← → 🔍 🔍 整体看，行业的用户规模还是在快速增长的，但……

中国和海外AI应用2025年9月MAU（月活跃用户规模）& 季度增长率

海外AI应用月活规模突破14亿，仍保持稳健增长；中国市场增速迅猛，用户量翻倍



MAU

144823.43w

2025年9月海外AI应用的月活跃用户规模已突破14亿。这一庞大的用户基数明确表明，AI技术在海外市场已实现广泛渗透与高度普及。

季度增长率

19.16%

海外AI应用市场展现出卓越的增长潜力：即便在MAU已超14亿的规模下，其第三季度环比增长率仍高达19.16%。



MAU

46688.01w

中国整体AI应用活跃用户数超 4.6 亿，增长平稳。

季度增长率

7.17%

由DeepSeek引爆的AI应用普及潮，推动市场在2025年Q1录得104.41%的爆发性增长。至Q3，增速已常态化至7.17%，显示市场增长动能正由初期的快速渗透转为平稳扩张。

Source: Xsignal数据交互平台AI Holo（AI全息）数据库

时间：2025.7-2025.9

增长率：AI应用2025年9月MAU对比2025年6月MAU的增长百分比。



TOP 50 AI应用季度榜单|Web+APP|海外

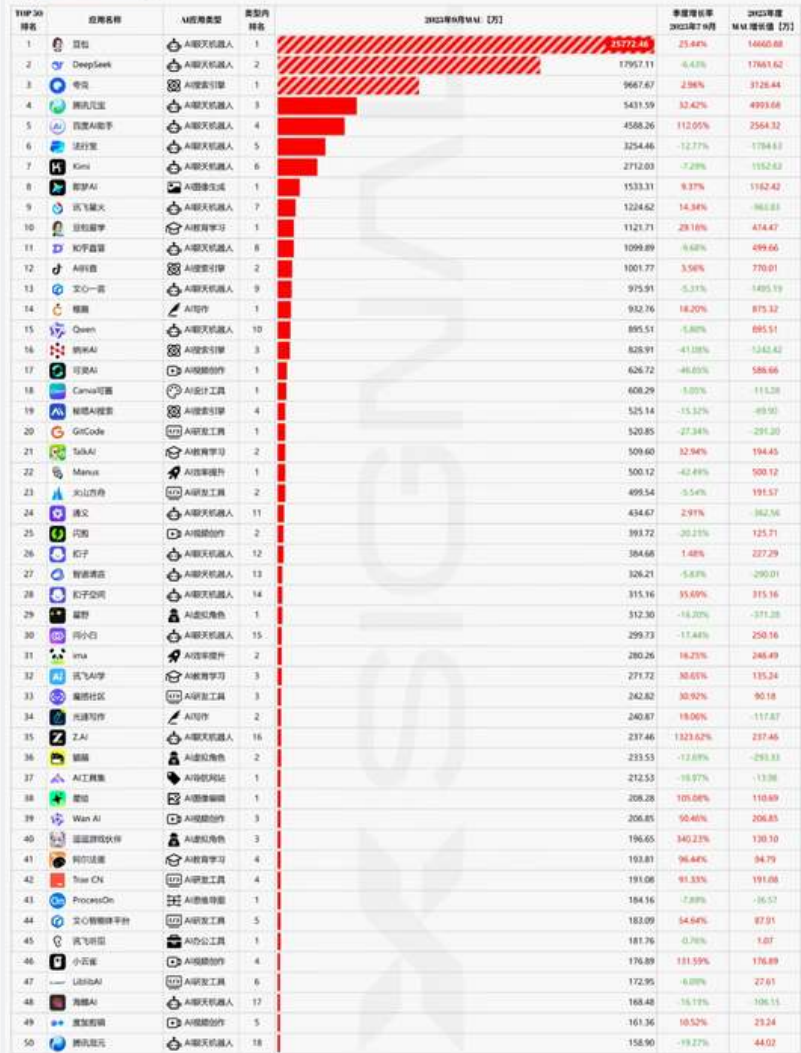
Google Gemini 年增 2.5 亿强势追赶，AI 教育应用 Wayground 季度暴涨 7 倍



Source: Xsignal数据交互平台AI Holo（AI全息）数据库
时间：2025.7-2025.9
*增长率：AI应用2025年9月MAU对比2025年6月MAU的增长百分比。

TOP 50 AI应用季度榜单|Web+APP|中国

近四成应用增长承压，百度AI助手凭112%增速重塑格局



Source: Xsignal数据交互平台AI Holo（AI全息）数据库
时间：2025.7-2025.9
*增长率：AI应用2025年9月MAU对比2025年6月MAU的增长百分比。

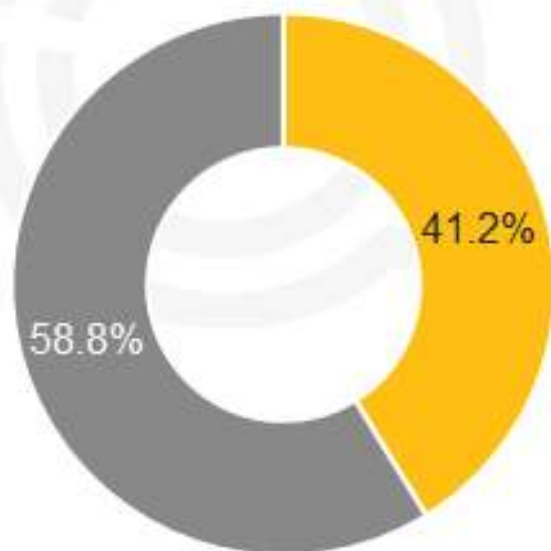
“曾经围绕模型的狂热与想象正逐渐退潮，市场开始用最朴素的标尺：
用户价值与活跃规模，去度量真正的胜出者。”

---Xsignal主编刘震

← → 🔍 谁会甘心做个**插件**呢? ---2.0

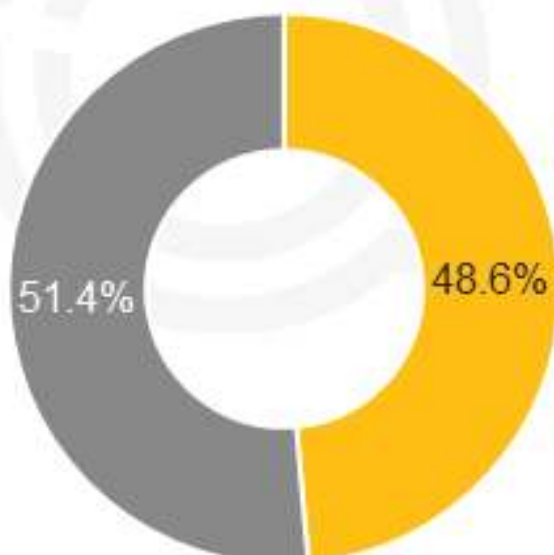
原生App整体

■ 正增长应用数量占比
■ 负增长应用数量占比



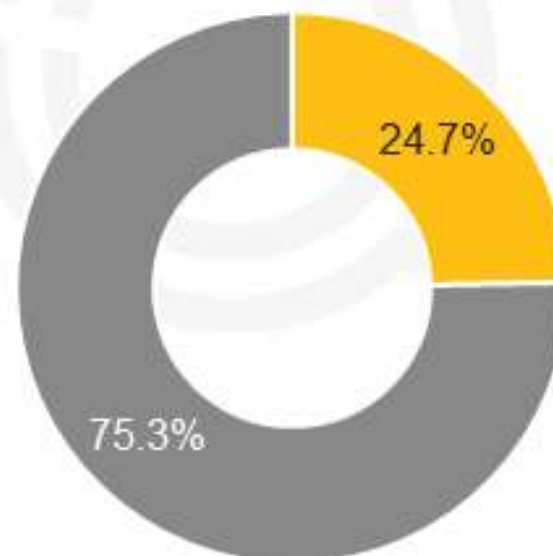
应用插件整体

■ 正增长应用数量占比
■ 负增长应用数量占比



PC网页整体

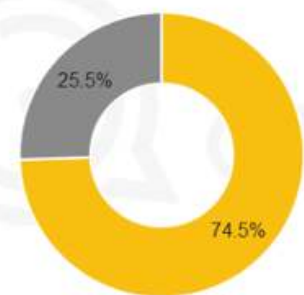
■ 正增长应用数量占比
■ 负增长应用数量占比



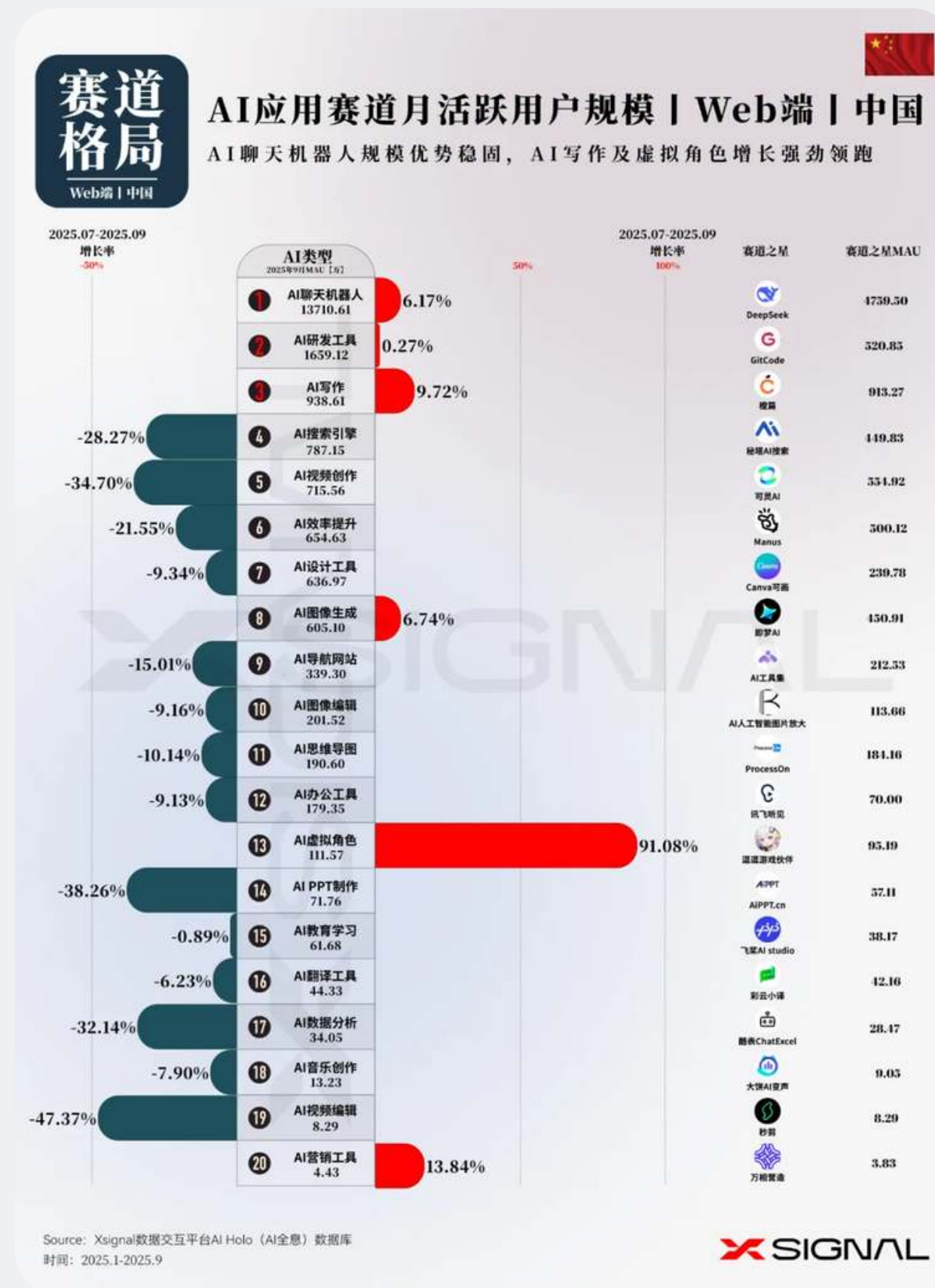
上半年的数据，插件还有3/4是正增长的
这怎么到了Q3就开始溃败了？

应用插件整体复合增长率表现

■ 正增长应用数量占比 ■ 负增长应用数量占比



← → 🔍 🔍 谁会甘心做个**插件**呢？---2.0



Web端的数据
确实有点惨不忍睹

← → 🔍 人均单次Token是一个合适的衡量指标吗？

2025年9月 vs 2024年9月 移动端三类形态AI应用 人均单次Token 变化



注：人均单次token消耗量：在特定时间段内，每个独立用户平均每次与AI应用进行实时、有效交互时，所消耗的由生成式AI大模型在推理过程中产生的输入、输出token总数。

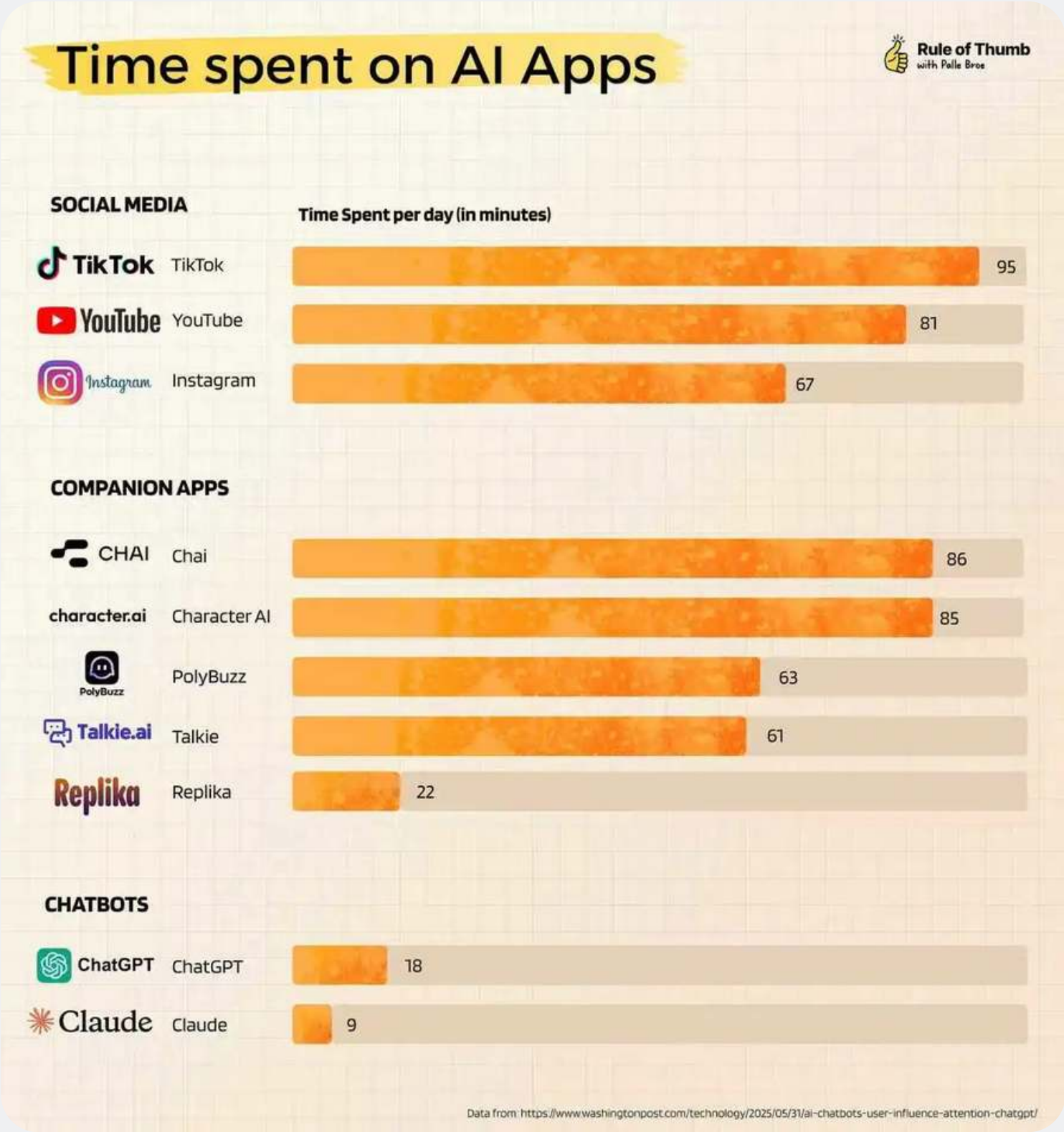
Source: QuestMobile AI产业研究院 2025年10月



人均单次Token的下降
到底是模型效率提升引发的？
还是……

← → 🔍

如果Token不合适，那人均使用时长呢？



#	产品名称	产品分类	月浏览量	使用时长	热门地区
1	讯飞智文 zhiwen.xfyun.cn	PPT	12.98万	00:15:50	中国
2	MasterGo mastergo.com	设计工具	216.08万	00:14:25	中国
3	Pixso AI pixso.cn	设计工具	62.33万	00:12:00	中国
4	即时AI设计 js.design	设计工具	67.81万	00:10:41	中国
5	扣子 coze.cn	智能体	528.72万	00:08:37	中国
6	哩布哩布AI liblib.art	图像生成	273.04万	00:08:36	中国
7	AutoDL算力云 autodl.com	算力平台	102.33万	00:08:06	中国
8	问小白 wenxiaobai.com	AI ChatBots	514.79万	00:06:18	中国
9	同花顺问财 lwencai.com	财经	144.2万	00:05:34	中国
10	智谱AI开放平台 bigmodel.cn	大模型	150.41万	00:05:26	中国

← → 🔍 🔍 还不是因为别的指标东西不好衡量吗.....

产品叙事

OpenAI是家什么公司?

- 一家5000亿美金的初创公司
- 能用1-2个描述就概括了么?
- 技术 or 产品?
- 还是全都要?
-

热门产品方向

- Chatbot
- Coding
- Agent
- 那些数据
-

ARR的叙事

- ARR依然是有效手段吗?
- ARR之外还有什么指标?
- 毛利? 净利? RPO?
- 如果不增长怎么办?
-

← → 🔍

ARR到底是什么？

ARR Formula



AI's \$100 Million ARR Club

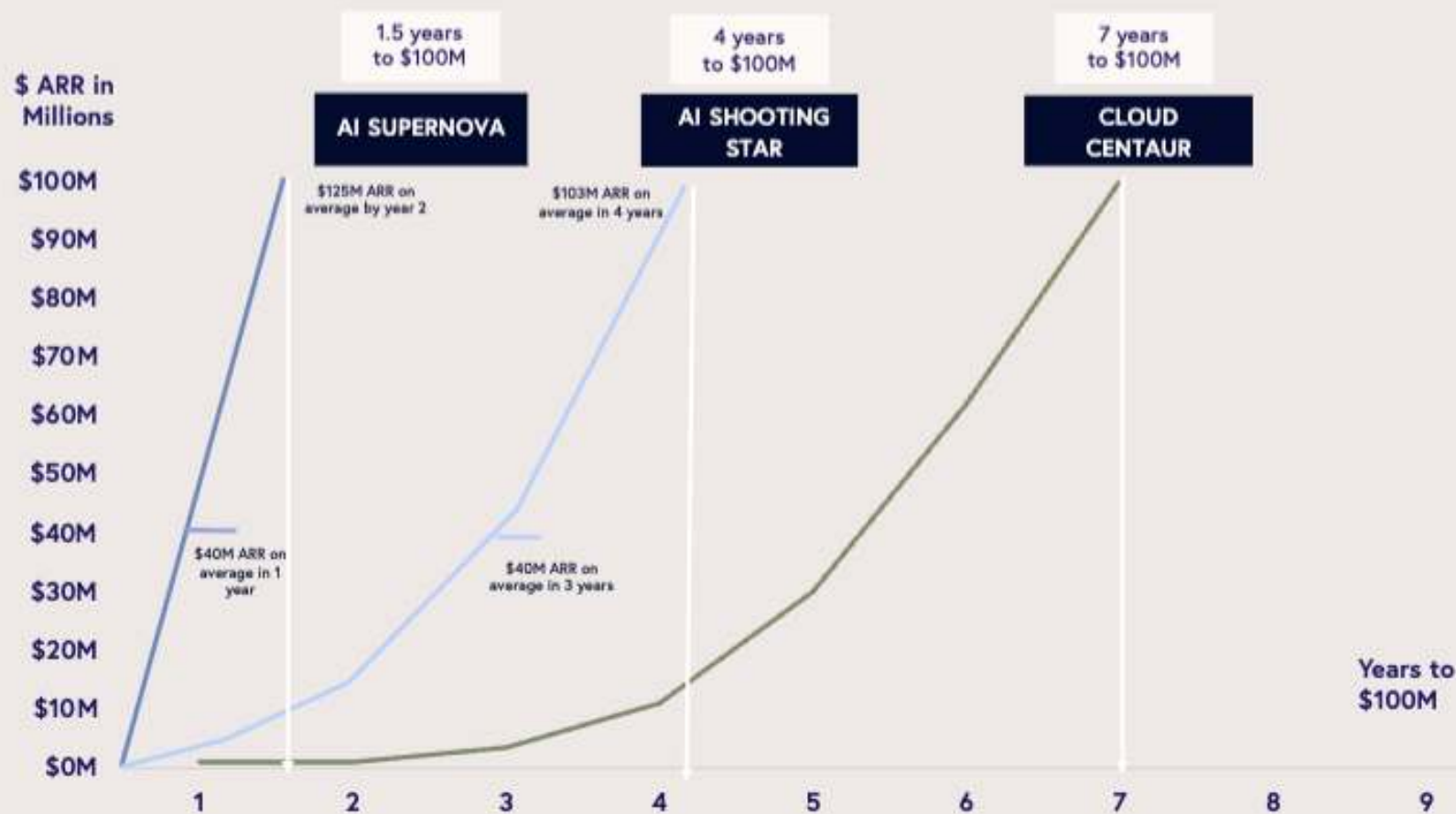
AI startups are rocketing into the \$100 million ARR club faster than ever.

Company	ARR	Total Funding	Valuation
OpenAI	>\$12B	\$30B	\$500B
Anthropic	>\$5B	\$28.6B	\$170B
Anysphere	>\$500M	\$1B	\$9B
ElevenLabs	>\$200M	\$281M	\$6.6B
Midjourney	>\$200M	Bootstrapped	N/A
Perplexity AI	>\$150M	\$1.1B	\$20B
Replit	>\$150M	\$450M	\$3B
Glean	>\$103M	\$770M	\$7.2B
xAI	>\$100M	\$17B	\$113B
Cohere	>\$100M	\$1.6B	\$7B
Abridge	>\$100M	\$762.5M	\$5.3B
EliseAI	>\$100M	\$342M	\$2.2B
Synthesia	>\$100M	\$335.6M	\$2.1B
Lovable	>\$100M	\$222.5M	\$1.8B

Source:The Information reporting

← → 🔍 当ARR成为事实上的**衡量标准**

The new \$100M ARR growth curve for AI startups



Source: Bessemer research from portfolio data

AVERAGE RATES OF GROWTH

SUPERNOVAS

Year 1 - \$40M ARR
Year 2 - \$125M ARR

SHOOTING STARS

Year 1 - \$3M ARR
Year 2 - \$12M ARR
Year 3 - \$40M ARR
Year 4 - \$103M ARR

CLOUD CENTAUR

On average, the top private cloud companies reach \$100M ARR in < 7 years.



EliseAI Joins the

"\$100M" ARR Club

Company	ARR	Total Funding	Valuation
Elise ^{AI}	>\$100M	\$342M	\$2.2B
IIElevenLabs	>\$200M	\$281M	\$6.6B
glean	>\$103M	\$770M	\$7.2B
synthesia	>\$100M	\$335.6M	\$2.1B
Lovable	>\$100M	\$225.5M	\$1.8B
cohere	>\$100M	\$1.6B	\$7B

← → 🔍 🔍 但问题是……



Voltaire

This agglomeration which was called and which still calls itself the Holy Roman Empire was neither holy, nor Roman, nor an empire.

AZ QUOTES

A R R

- 既不**年度**
- 又不**经常性**
- 甚至不是**收入**

← → 🔍 🔍 不过，真金白银花出去的钱，总是能说明一些问题的……

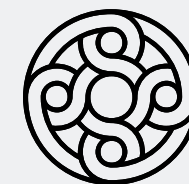
AI Apps 50: Where Startups Spend on AI

1. OpenAI	11. Delve	21. glean	31. descript	41. Otter.ai
2. ANTHROPIC	12. perplexity	22. Photoroom	32. grammarly	42. alma
3. replit	13. instantly	23. Solve Intelligence	33. manus	43. applaud
4. freepik	14. customer.io	24. GAMMA	34. Cognition	44. CapCut
5. ElevenLabs	15. KlingAI	25. clay	35. OpusClip	45. motion
6. CURSOR	16. Retell AI	26. Cluely	36. happyscribe	46. crisp.
7. Fyxr.ai	17. Canva	27. crosby	37. 11x	47. arcads
8. Lorikeet	18. Lovable	28. Midjourney	38. PLAUD	48. emergent
9. micro1.	19. Metaview	29. combinely	39. Serval	49. Read AI
10. Notion	20. ADEPT	30. Merlin	40. ada	50. tavus

Some of the companies mentioned are portfolio companies of a16z. For a full list of portfolio companies, please visit a16z.com/investment-list.



- 泛/通用：垂直=6：4
- 垂类代表：客服、销售、人力
- Coding：4/50（不算通用模型）
- 也出现在流量榜单上的应用12个；
- 中国：可灵、manus、plaud、剪映



MERCURY

← → 🔍 谈到ARR，谈到钱，那么我们就谈的再多点.....



资本趋势

二级市场

- MAG7的叙事延续
- 那些“妖艳异常”的狠角色
- 延展行业
-

一级市场

- AI是事实上的核心叙事
- Big Deal + Big Fund
- 某种程度上的“被拉爆”
-

退出

- IPO
- 并购
- “野路子”
-

← → 🔍 这个系列的图表，我几乎每次都要用.....

S&P 500 MARKET CAP REACHES RECORD \$60.3 TRILLION IN OCTOBER 2025

Total Market Cap
Oct 24, 2025
\$60.3 Trillion



Sources: S&P Global, Econovis

www.econovis.net @econovisuals EV

AI BOOM PROPELS 'MAGNIFICENT SEVEN' MARKET CAP PAST \$21 TRILLION RECORD

Total Market Cap
Oct 24, 2025
\$21.1 Trillion



Sources: NASDAQ, Econovis

www.econovis.net @econovisuals EV

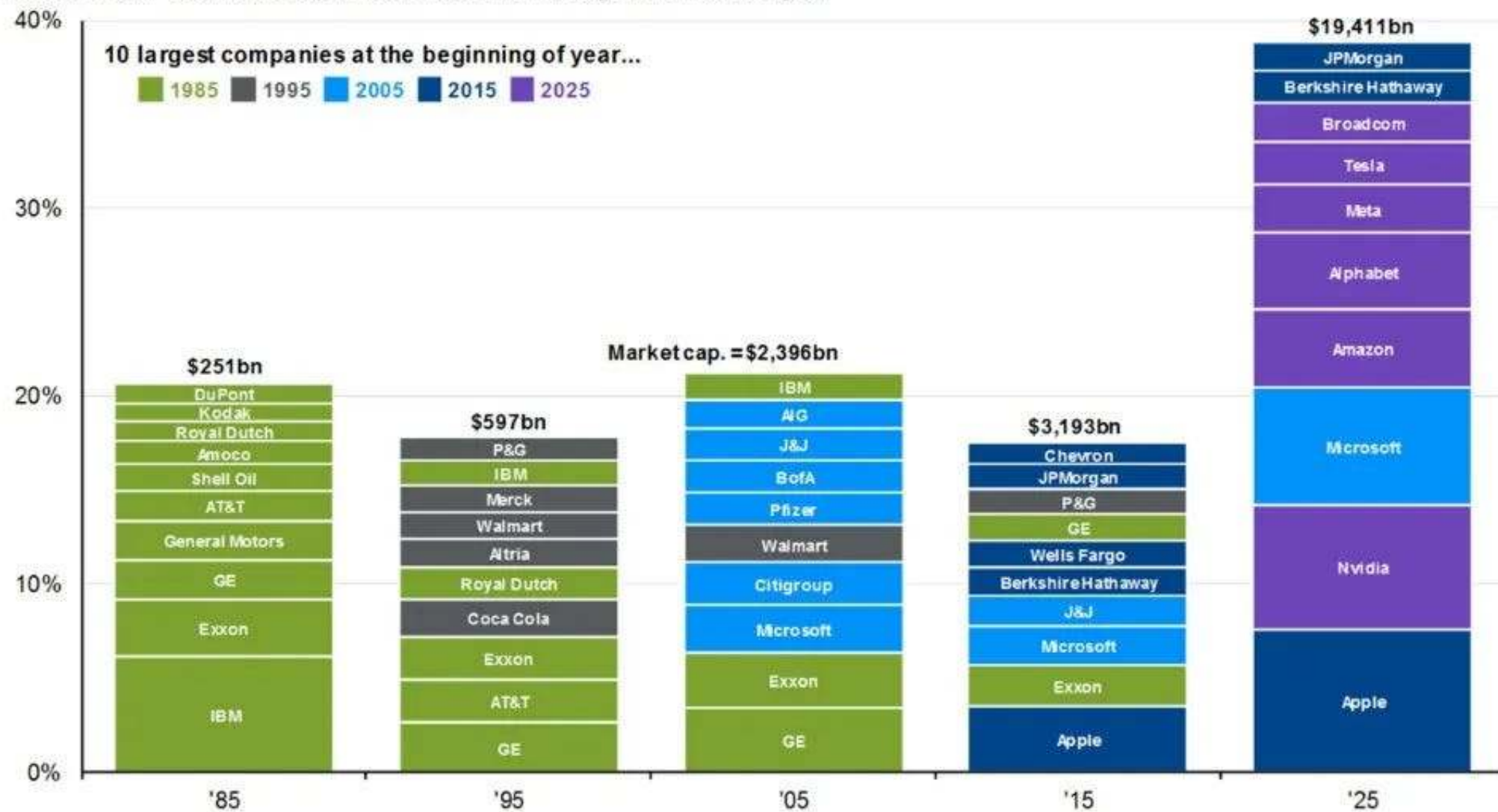
← → 🔍 🔍 拉长周期看，过去的多年整个市场也是科技公司引领的

Top 10 companies by decade

Actions ▾

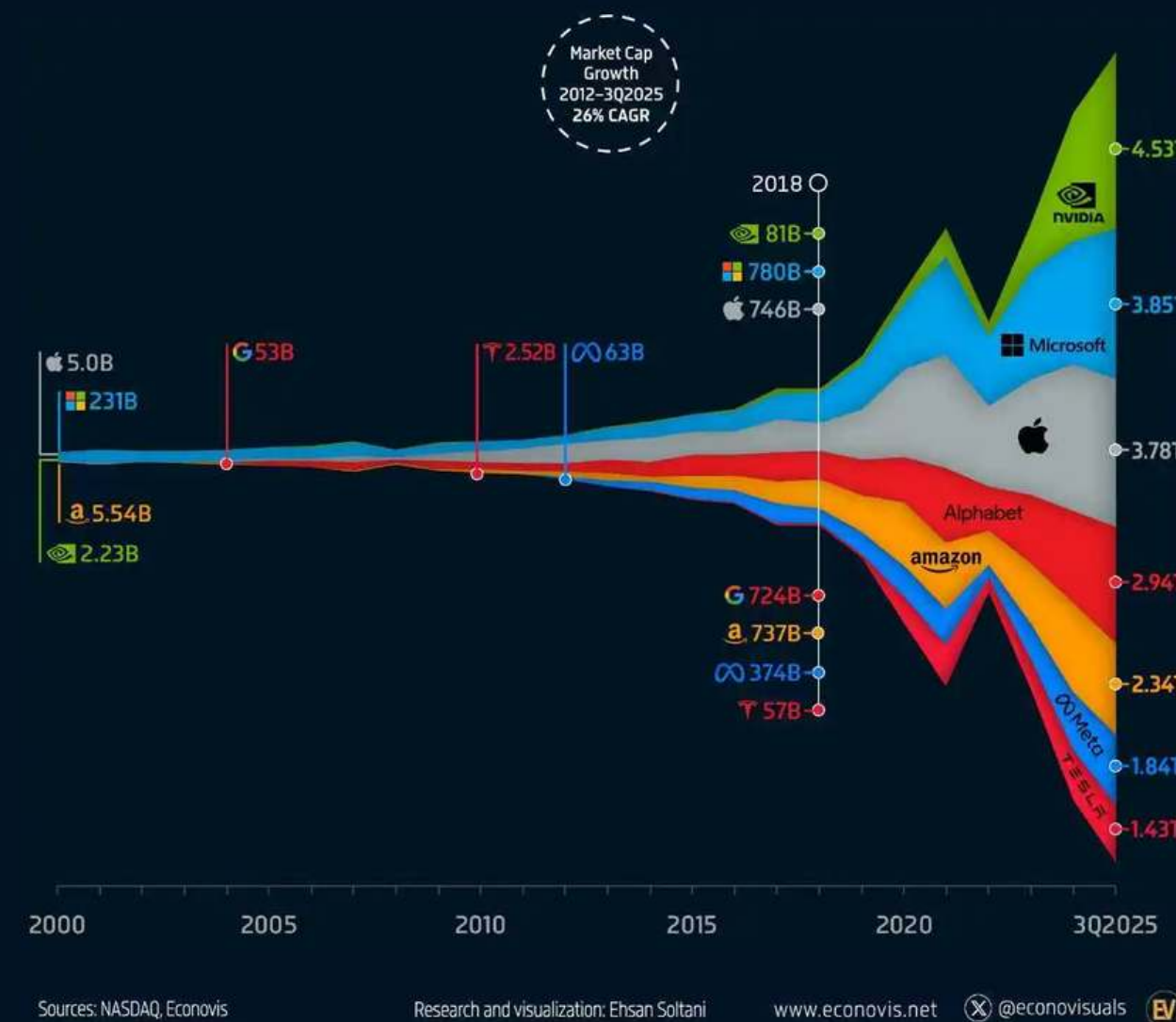
Top 10 S&P 500 companies by market capitalization

Percent of S&P 500 market capitalization as of the first day of the indicated year



U.S. BIG TECH MARKET CAP BOOM (2000–Q3 2025)

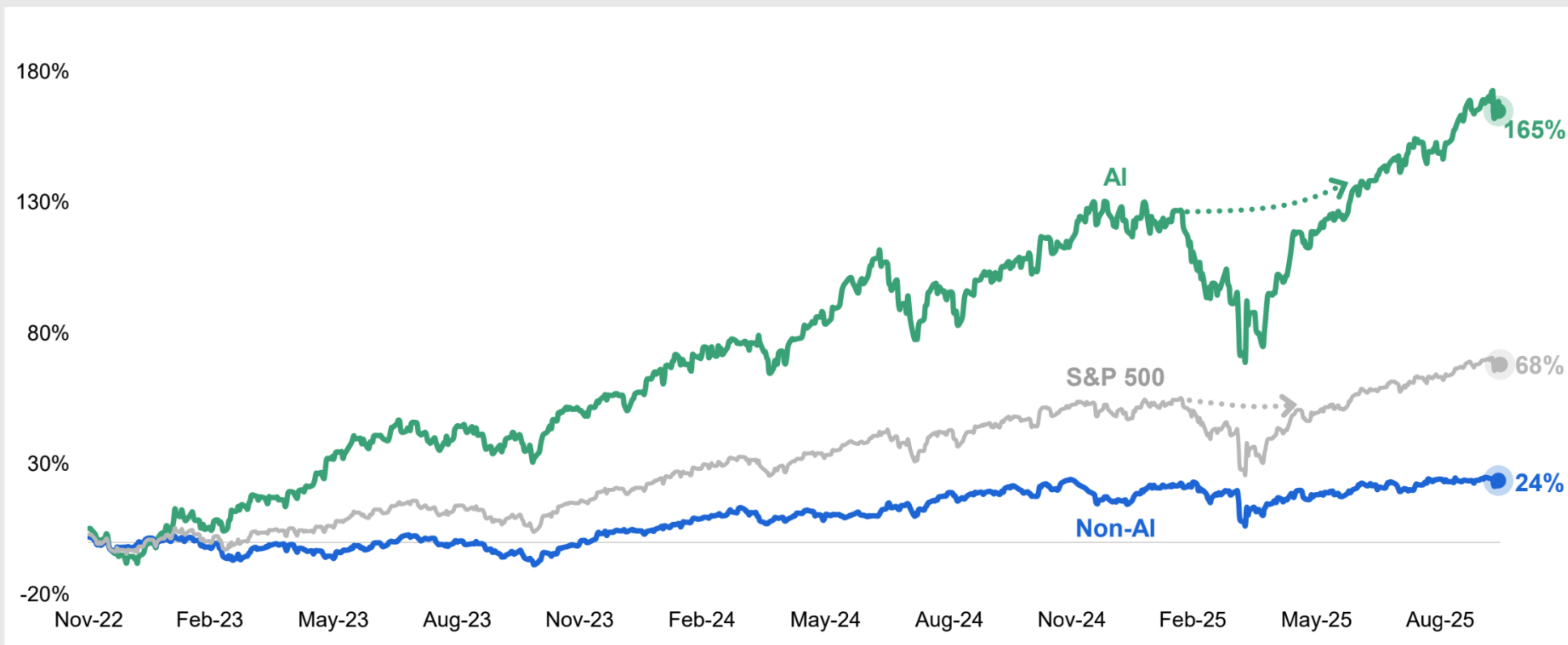
Total Market Cap
Sep 30, 2025
\$20.7 Trillion



← → 🔍 ChatGPT发布马上三年，AI当然引领风潮

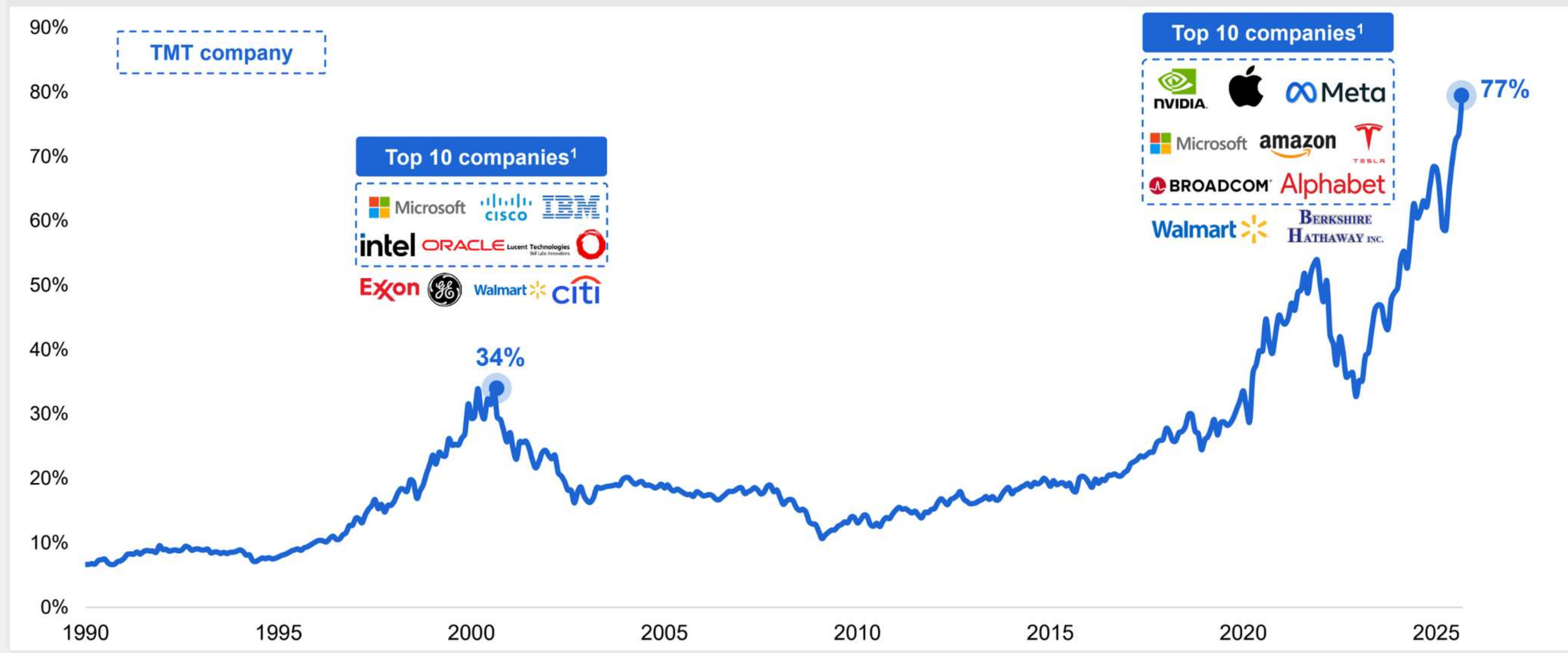
AI continues to drive markets higher

S&P 500 AI vs non-AI returns since launch of ChatGPT (Nov-22)



← → 🔍 科技巨头们引领的这一轮大牛市

Market cap for top 10 companies as % of US GDP

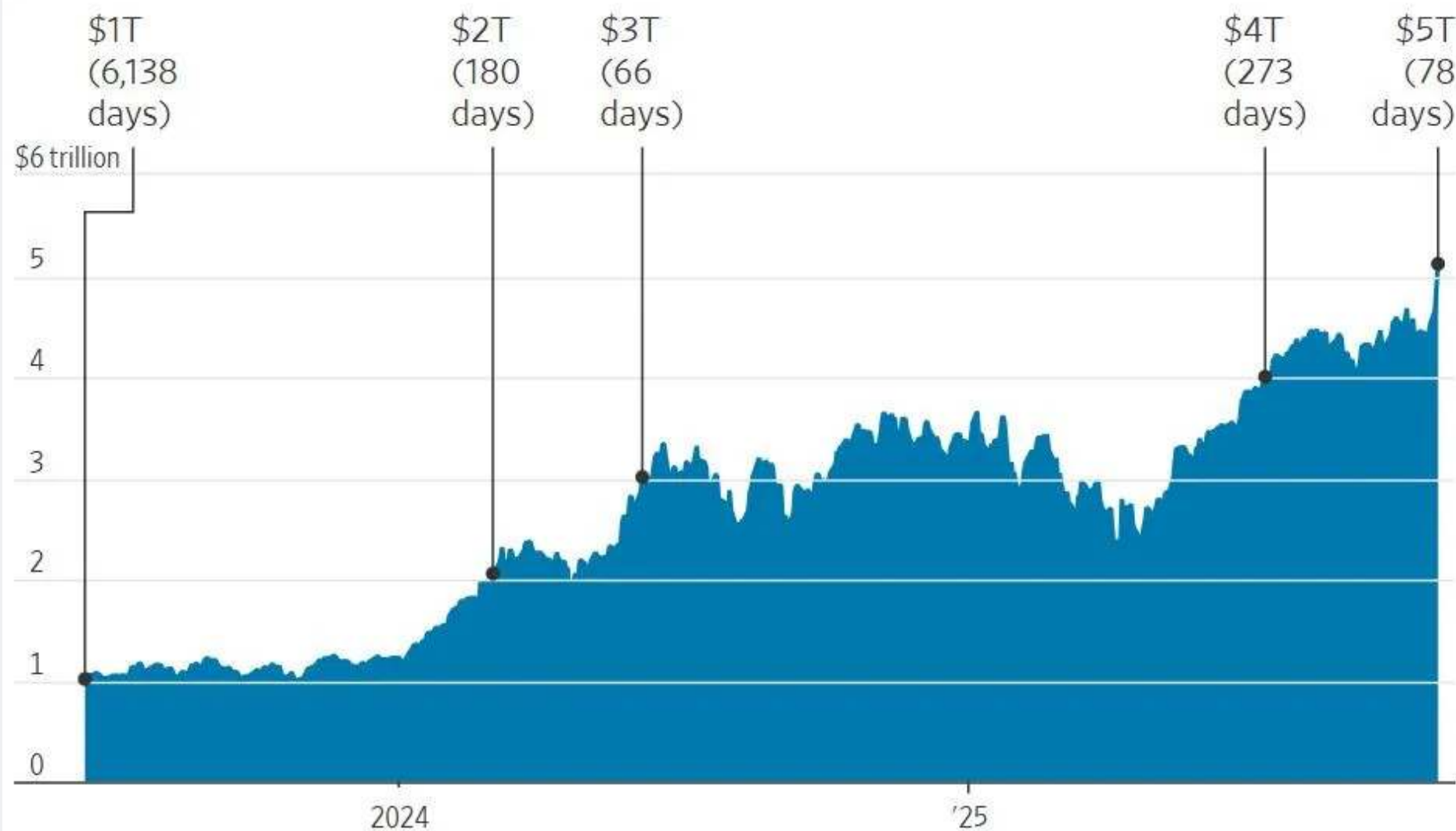


← → 🔍 即便回顾里已经有了，但Nvidia依然需要再增加一些笔墨

Slow, Then Fast

The number of trading days it took Nvidia to reach successive valuation milestones

Nvidia's market cap



Note: Most recent data point shows intraday trading on Oct. 29, 2025. Previous data reflects closing levels.

Sources: LSEG, Dow Jones Market Data

nVIDIA vs Big Pharma

\$4.6T
Market cap



\$4.2T
Combined Market Cap

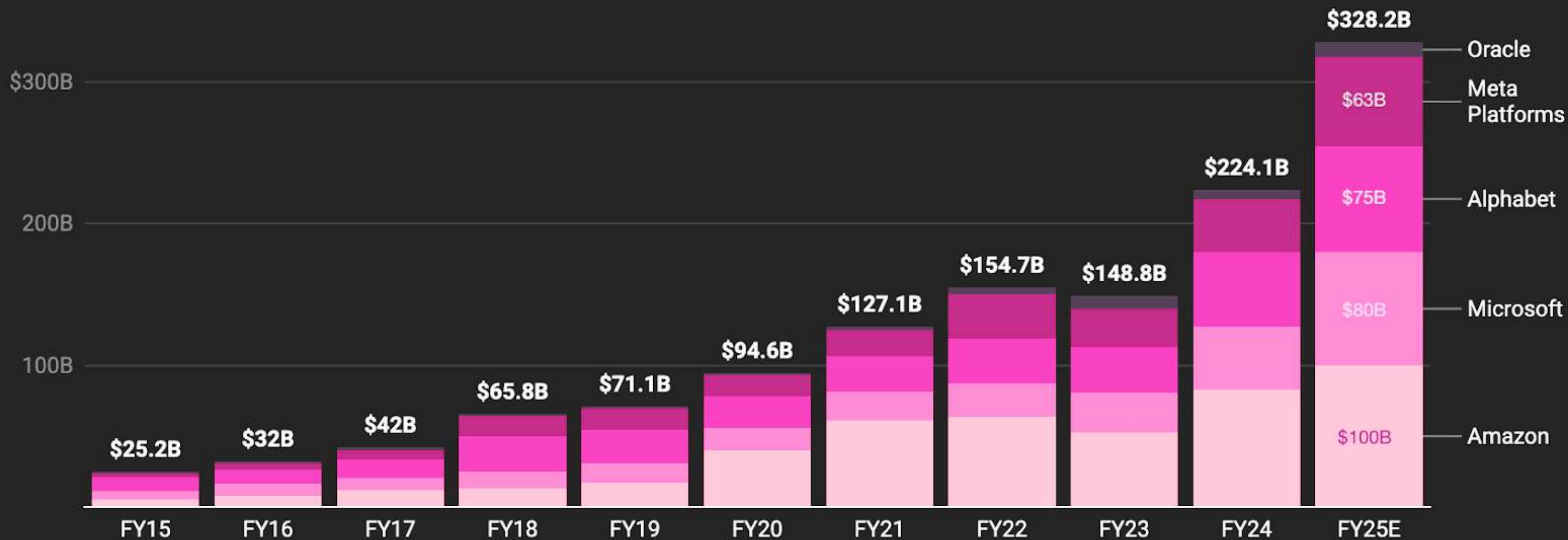


Oct 2025
<https://www.linkedin.com/in/rezazahiri/>

Reza Zahiri

← → 🔍 老生常谈的Capex

Hyperscalers Plan Another Year of Near 50% Growth in AI Capex



FY25 numbers are based on guidance from the managements of the five hyperscalers shown

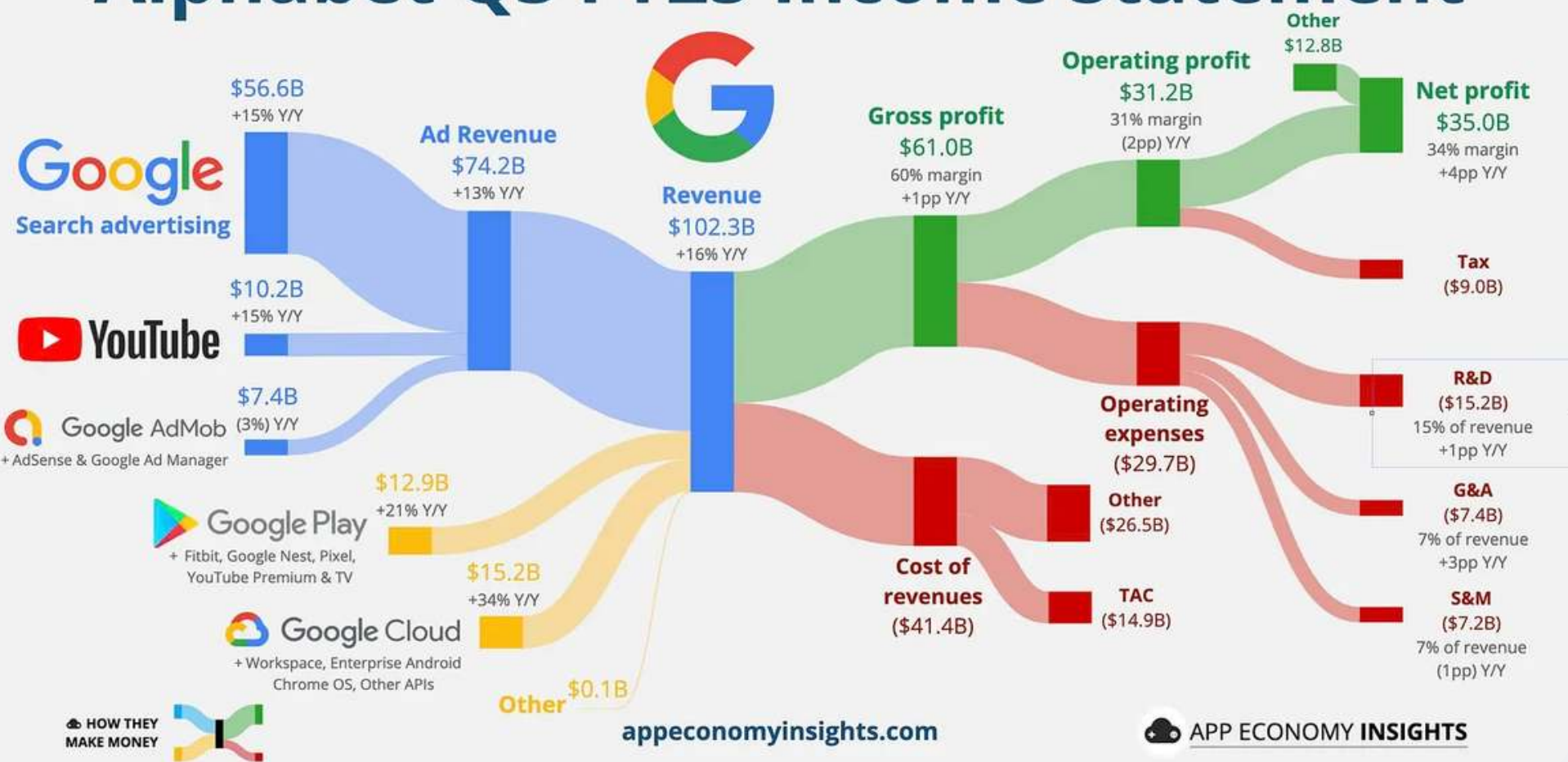
Chart: The Pragmatic Optimist Newsletter • Source: [Company sources](#)

← → 🔍 热乎的Google财报

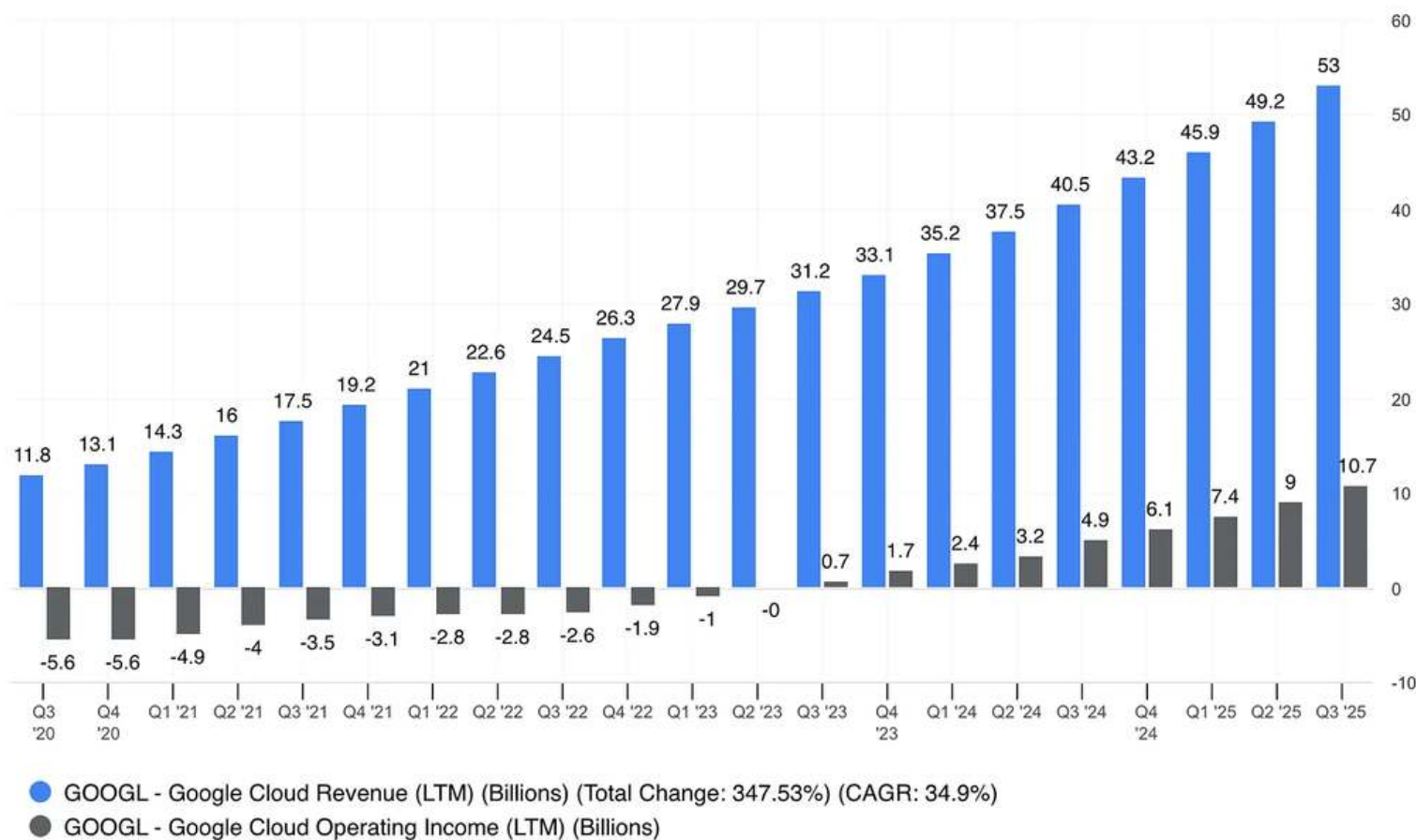
🔍 Google's NVIDIA Moment?

TPUs, Anthropic, and the race to power AI

Alphabet Q3 FY25 Income Statement



Google Cloud Revenue & Profit (TTM)

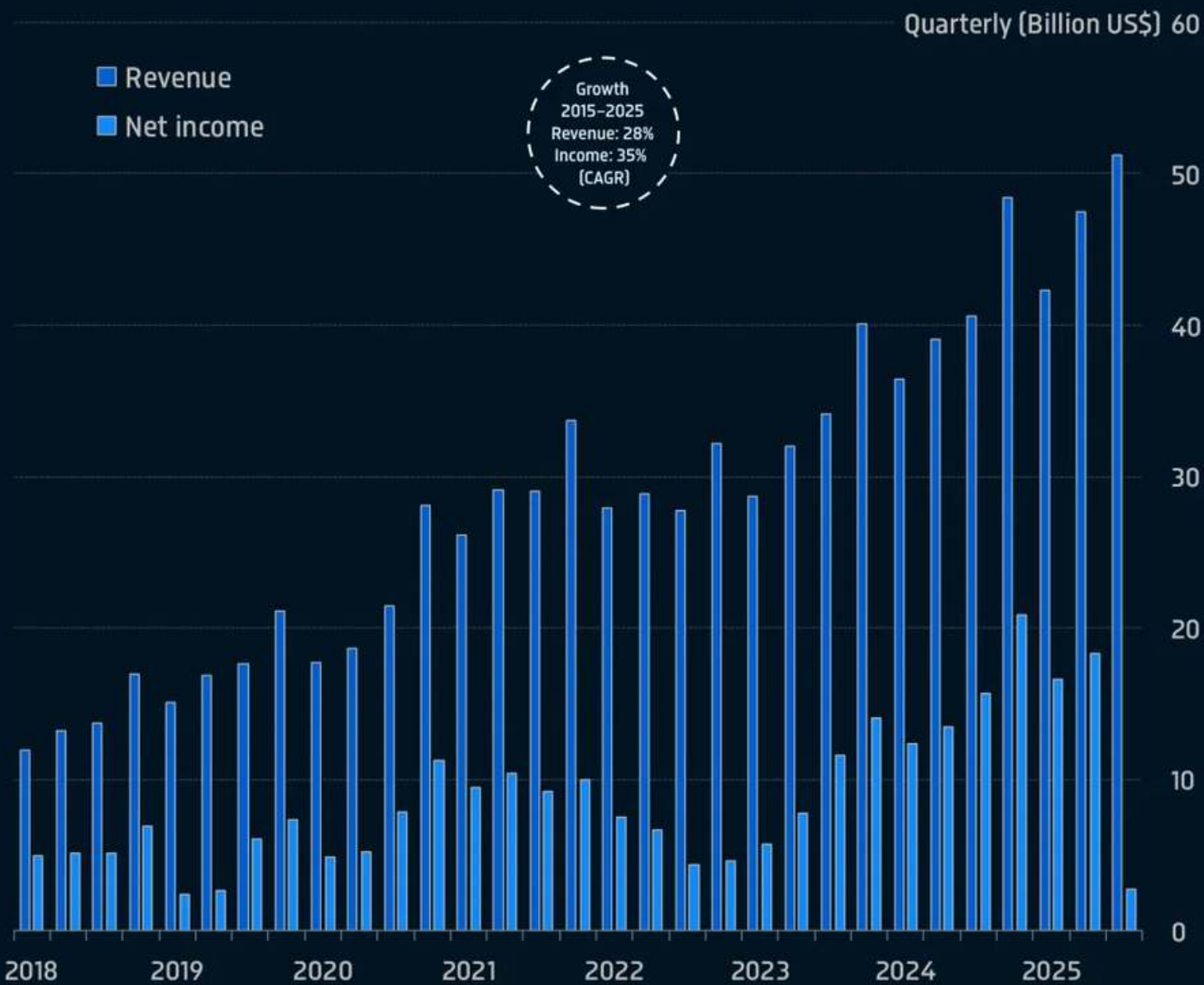


APP ECONOMY INSIGHTS

Powered by Fiscal.ai

← → 🔍 热乎的Meta财报

Meta Q3 2025: Revenue Up 26% to \$51.2B, Net Income Down 83% to \$2.7B



Sources: Meta Platforms Financial Statements

www.econovis.net

✕ @econovisuals



GOOGL

shares



+7.68%

META

shares



-7.31%

Meta(NASDAQ:META)

可融资

可卖空

\$666.47 -85.20 -11.33%

Rani Molla

META FISCAL

13H

Meta is diving because it's not Google or Microsoft

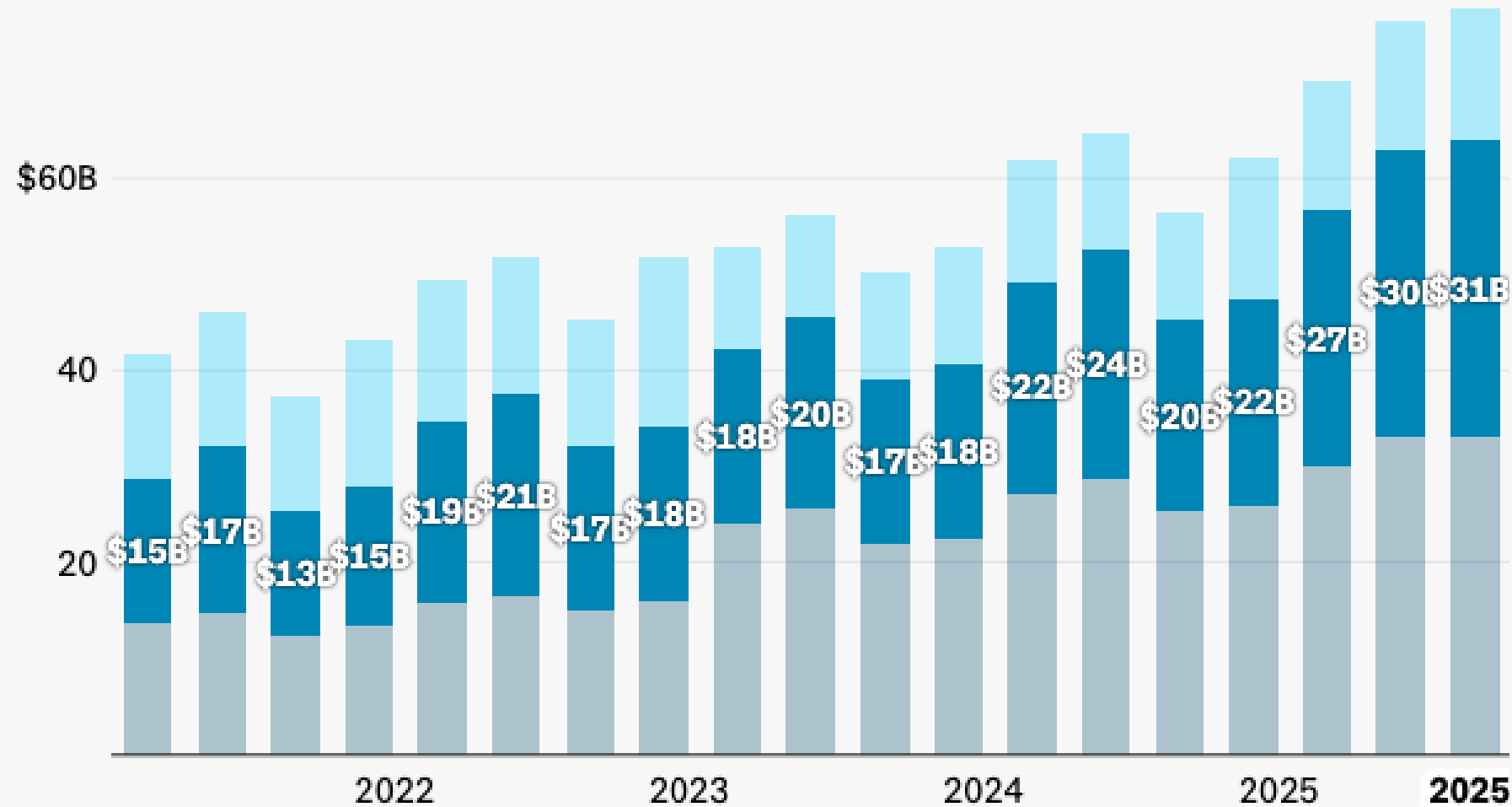
The stock is down sharply as investors question the social media company's attempt to spend like the cloud companies.



← → 🔍 热乎的微软财报

Microsoft quarterly revenue by business unit

- Productivity and Business Processes
- Intelligent Cloud
- More Personal Computing



Source: The company

Sherwood



+17%
Productivity and Business Processes
\$33.0 Billion

+17%
Microsoft 365 Commercial cloud

+26%
Microsoft 365 Consumer cloud

+10%
LinkedIn

+18%
Dynamics 365

+28%
Intelligent Cloud
\$30.9 Billion

+40%
Azure and other cloud services

Windows Server
SQL Server
System Center
GitHub
NUANCE

+4%
Personal Computing
\$13.8 Billion

+6%
Windows and OEM Devices

+1%
Xbox content and services

+16%
Search and news advertising

Q1 2026 revenue growth by business line

www.quartr.com

0015 0025306

← → 🔍 🔍 热乎的亚马逊财报

Sherwood

亚马逊第三季度财报超预期，股价飙升



月亮-加罗

礼来公司股价上涨，此前其盈利和销售额均远超华尔街预期。

J·爱德华·莫雷诺



一点点地消磨掉

特朗普和习近平没有谈到英伟达的“超级无敌”Blackwell芯片。

卢克·卡瓦



元财政

Meta正在走下坡路，因为它既不是谷歌也不是微软。

拉尼·莫拉

跌，因特朗普与习近平会...

社交媒体巨头表示人工智能资本支出...

Core Scientific股东投票反对CoreW...

据STAT报道，Moderna可能面临“收...

0

手机

拉尼·莫拉

最新消息



- Coinbase公布的销售额、利润和交易量均超出预期。
- 该策略盈利超出预期，盘后交易价格上涨。
- 西部数据股价飙升，原因是其第一季度销售额和盈利均超出预期，业绩指引也优于预期，并且提高了股息。
- Reddit股价在盘后交易中上涨，此前公布了第三季度财报。

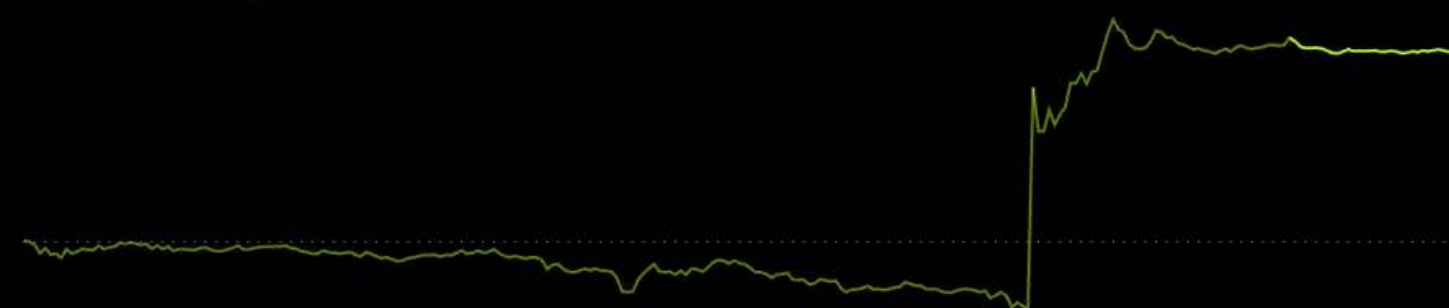


Amazon

\$251.40

-\$7.44 (-3.23%) Today

+\$28.54 (+12.81%) Overnight



“AWS 的增长速度达到了自 2022 年以来的最高水平，同比增长 20.2%。我们持续看到人工智能和核心基础设施的强劲需求，并且我们一直致力于加速提升产能——在过去 12 个月中新增了超过 3.8 吉瓦的产能。”

← → ↺ 🔍 不同的策略也站位，决定了这一年的走势.....

NVIDIA, MICROSOFT, ALPHABET, AND META POWER \$3.3 TRILLION SURGE IN MAGNIFICENT SEVEN MARKET CAP IN 2025



Sources: NASDAQ, Econovis

www.econovis.net

✕ @econovisuals

EV

当然，以上所有的分析

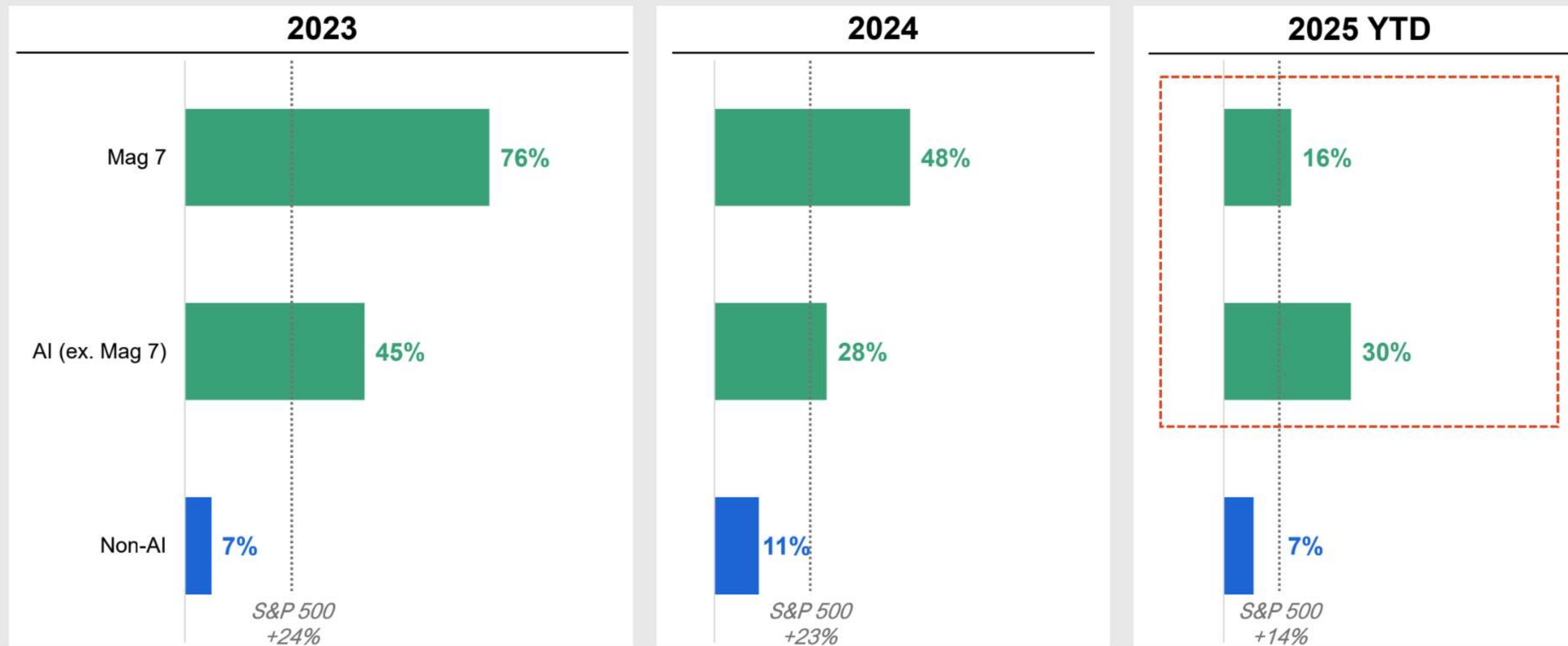
本质上全部都是

马后炮

← → 🔍 Q MAG7的叙事已经3年了，还能持续引领市场吗？

Mag 7 is no longer outperforming

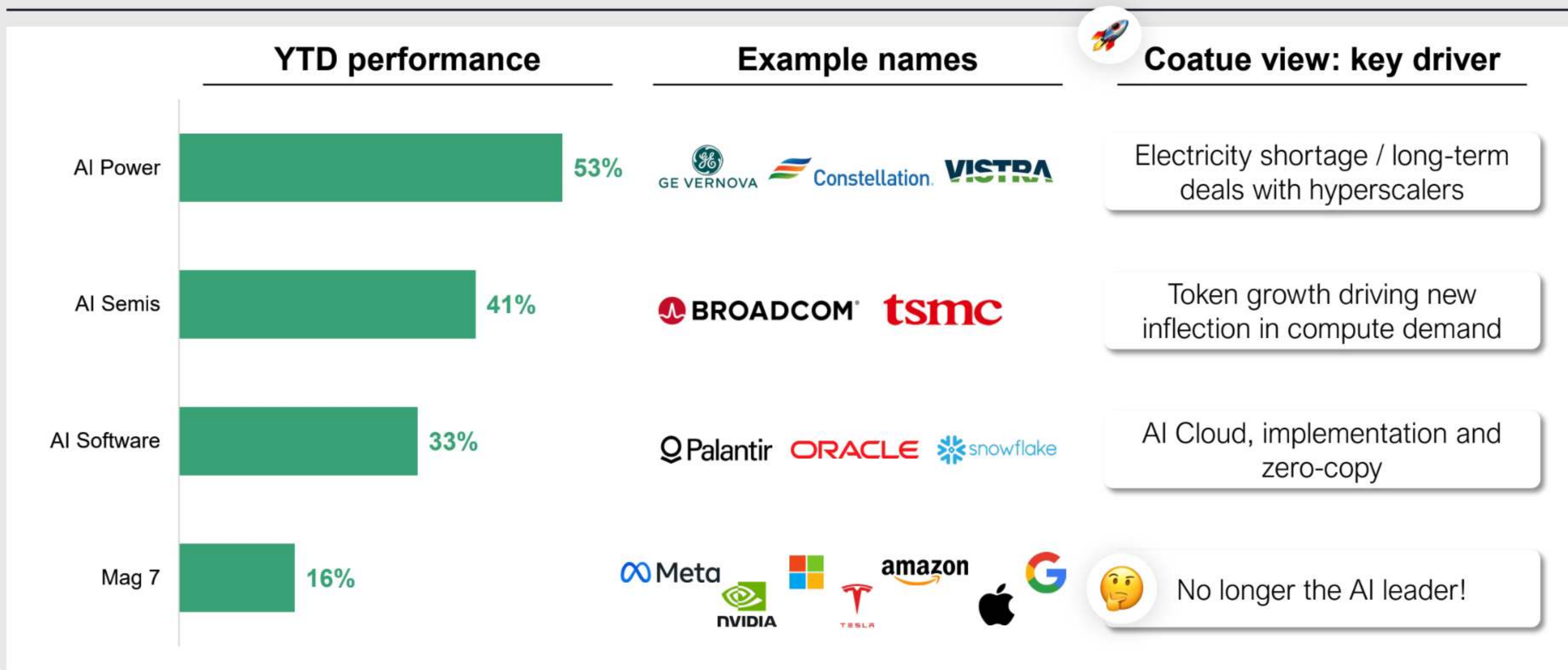
Cohort returns over time



← → 🔍 市场当然需要“新叙事”

The market is welcoming new AI leadership...

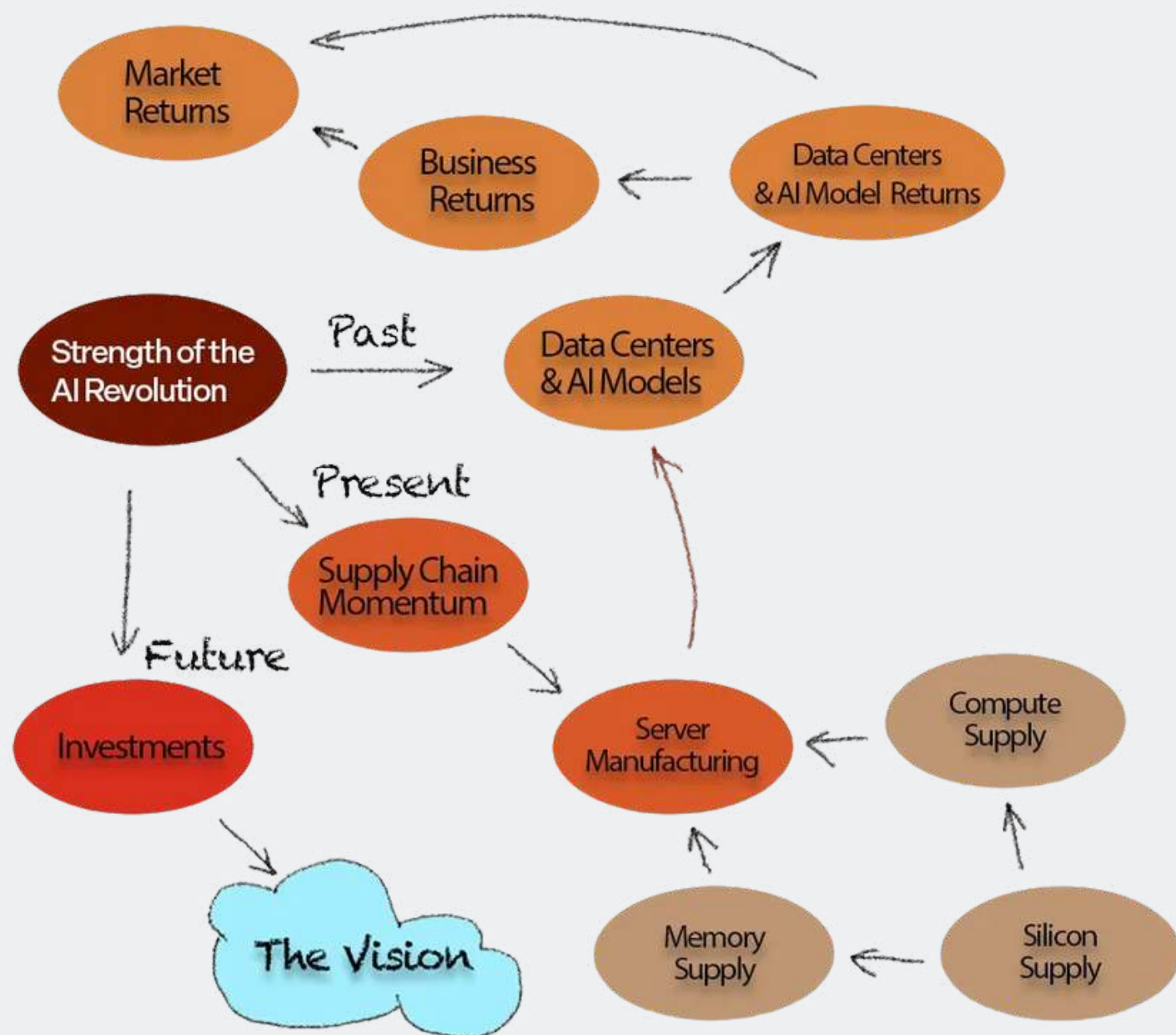
YTD returns of select sectors



← → 🔍 AI这个超级大的“故事”到底走到了哪里？

The Current Strength of the AI Revolution

The Past, Present and the Future of Artificial Intelligence

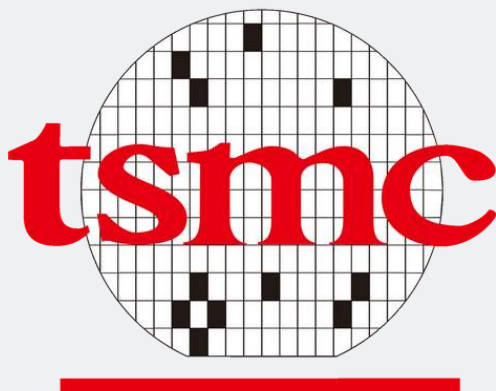
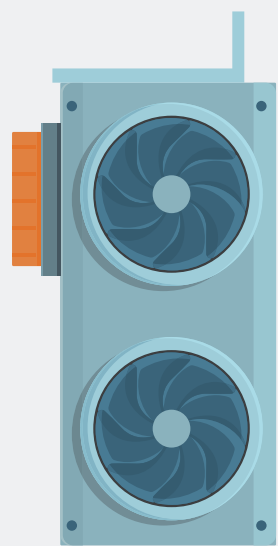


我们看到了众多2025年

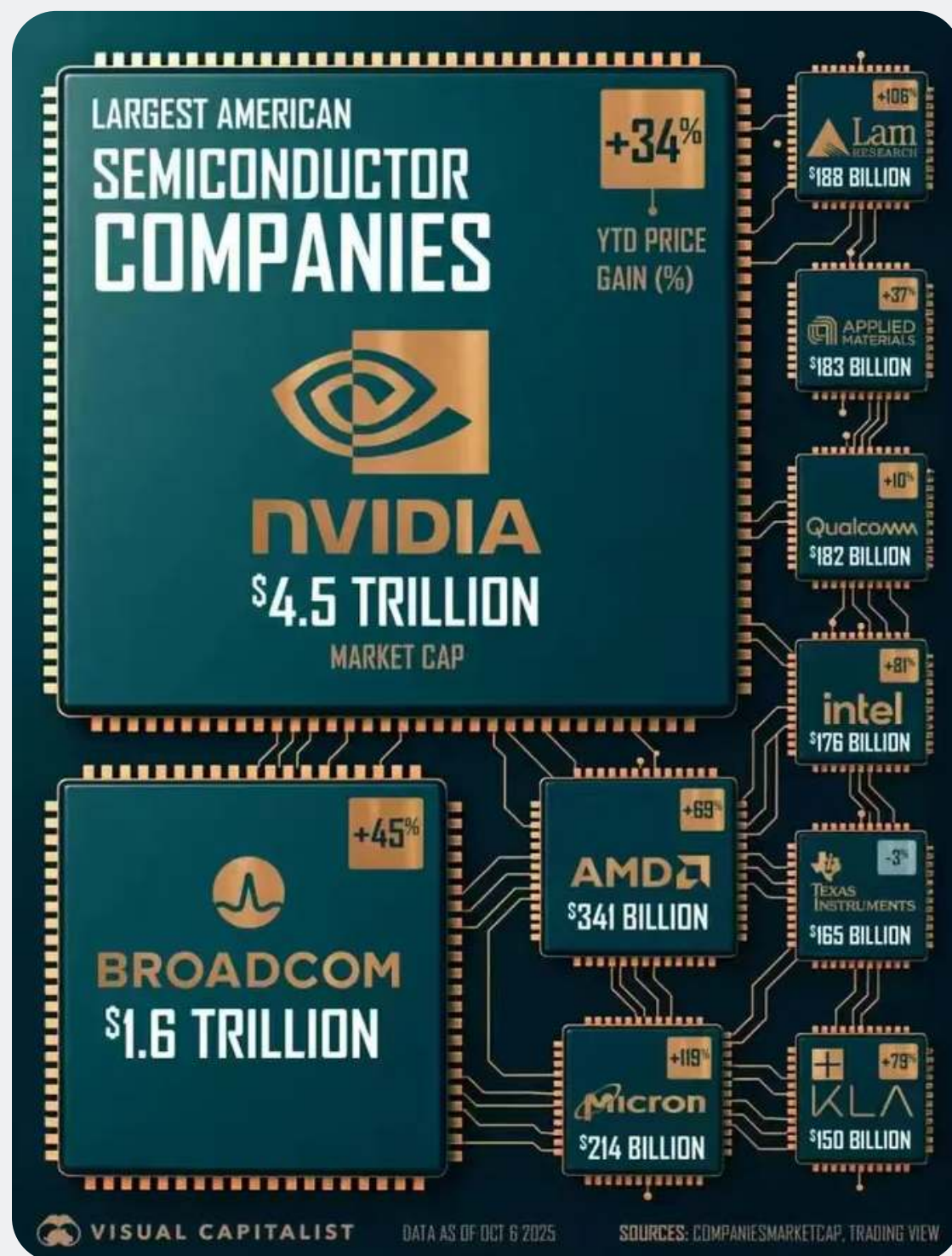
大A疯狂炒作过的“主题”：

- 大模型
- 数据中心
- 服务器
- 算力
- 存储
- 硅供应

← → 🔍 这些行业得到了“**边际**”更多的关注

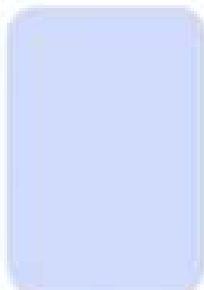
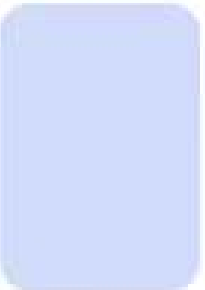


⬅ ➡ 🔁 🔍 **芯片行业**——绝对中心，以及国产GPU们为什么上市都给绿灯



Key players building accelerators for AI training and inference

Based on publicly available data of chips yet to be released and/or available for use

	Major Chipmakers				Cloud Hyperscalers		Challengers		
	NVIDIA	AMD	Intel ¹	Qualcomm	Google	Amazon	SambaNova	Groq	Cerebras
									
Training									
Inference (incl. inference within training)									

← → 🔍 🔍 存储---谁TM能想到，炒到内存和硬盘不够呢？

Exhibit 1: Who Makes the Data Center - 2025
Estimated capex costs of \$39mn per megawatt

Who Makes the Data Center - 2025

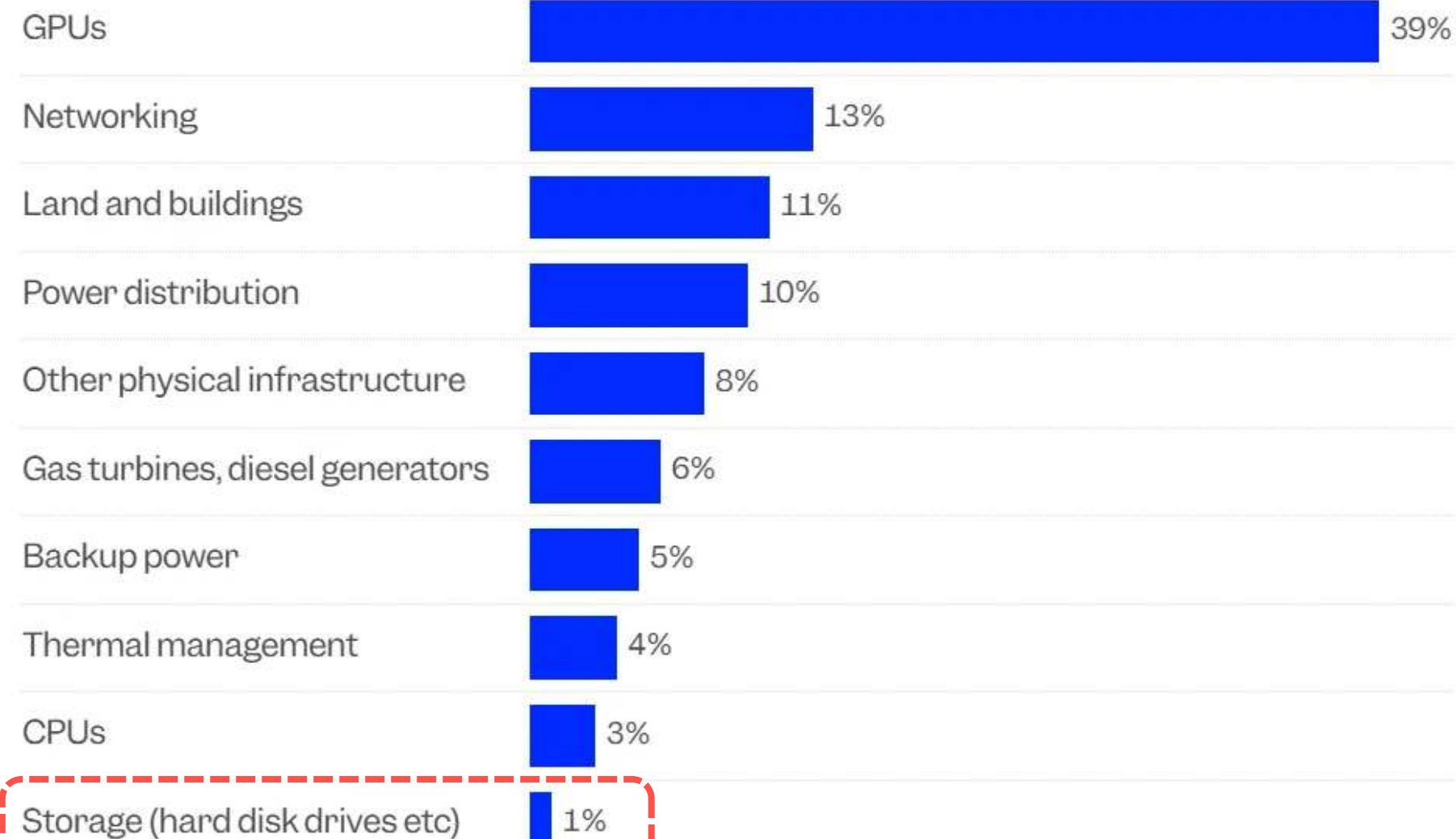
Key equipment, content per megawatt, and global vendors



Source: BofA Global Research

BoFA GLOBAL RESEARCH

Cost breakdown of a 1 GW data center



Note: Estimates based on \$35 billion cost for a 1GW AI data center, using Nvidia GB200 GPUs and the company's NVL72 server. • Source: Bernstein Research

← → 🔍 🔍 存储---谁TM能想到，炒到内存和机械硬盘不够呢？

Ramping up

Year-to-date percentage change, daily.

— Seagate Technology — Western Digital - - Nasdaq Composite



Source: FactSet

Sherwood



rosicky311_明浩

互联网小厂战略+投资 (社交/游戏...

AI探索站

09/23

标普500，今年涨幅TOP5

希捷

西部数据

palantir

纽蒙特矿业

甲骨文

两家做硬盘领跑

一家做ai实施

一家金矿

一家ai云



微博轻享版

雷军

小米创办人，董事长兼CEO；金山软件...

关注 1593 粉丝 2694.5万



2025-10-24 09:07

#k90# 12+512，首销月直降300元，仅售2899元。最近内存涨价实在太多，希望大家能理解我们这份诚意。

REDMI K90 12GB+512GB 首销月直降300

配置	首销月售价
12GB+256GB	¥2599
16GB+256GB	¥2699
12GB+512GB	¥2899
16GB+512GB	¥3499
16GB+1TB	¥3999

1.4万

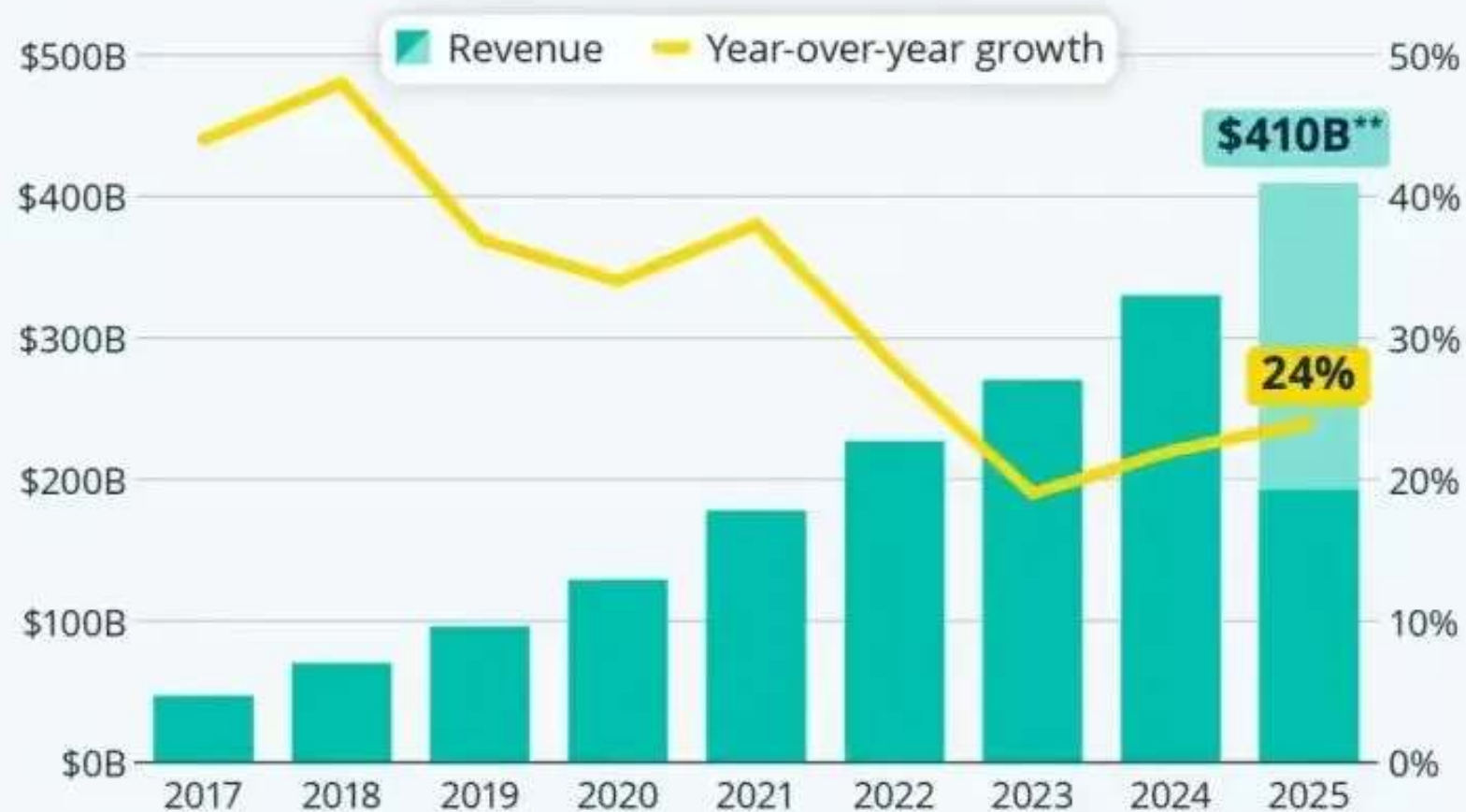
1873

211

← → 🔍 🔍 云，这一轮最直接AI的体现

AI Boom: Global Cloud Market to Exceed \$400 Billion in 2025

Worldwide cloud infrastructure service revenues*



* Includes platform as a service (PaaS) and infrastructure as a service (IaaS) as well as hosted private cloud services

** Projected assuming a growth rate of 24% for Q3 and Q4 2025

Source: Synergy Research Group



statista

Cloud revenue keeps going up...

Quarterly cloud unit revenues, in billions.

Amazon

\$30B

20

10

0

Jan
2021

Jul
2025

Microsoft

Google

Amazon AWS, Microsoft "Intelligent Cloud," Google Cloud.

Source: The companies.

...But the pace of growth is slowing

Quarterly cloud revenue growth, year on year.

Amazon

40%

20

0

Q1
2021

Q3
2025

Microsoft

Google

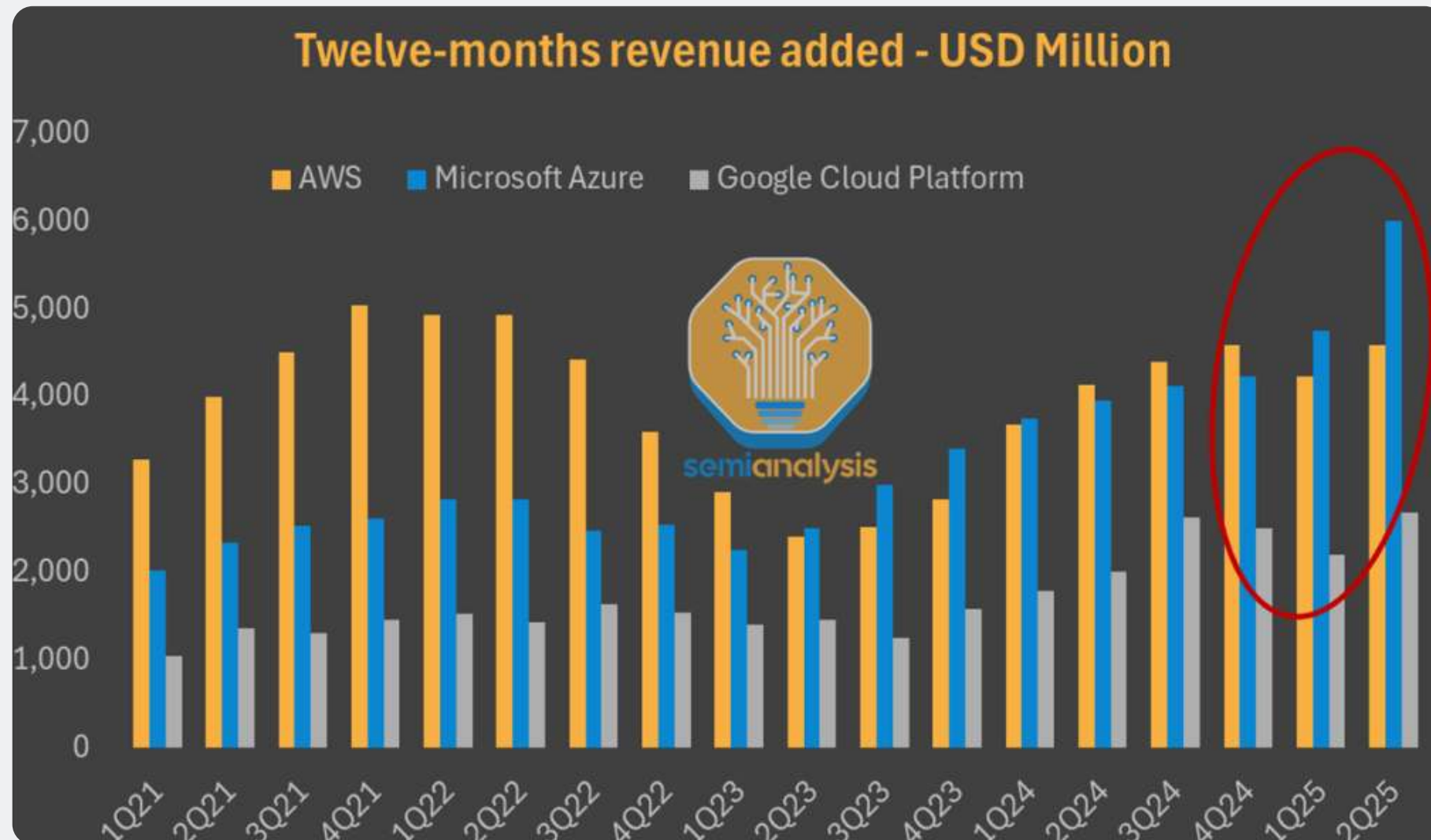
Amazon AWS, Microsoft "Intelligent Cloud," Google Cloud.

Source: The companies.

Sherwood

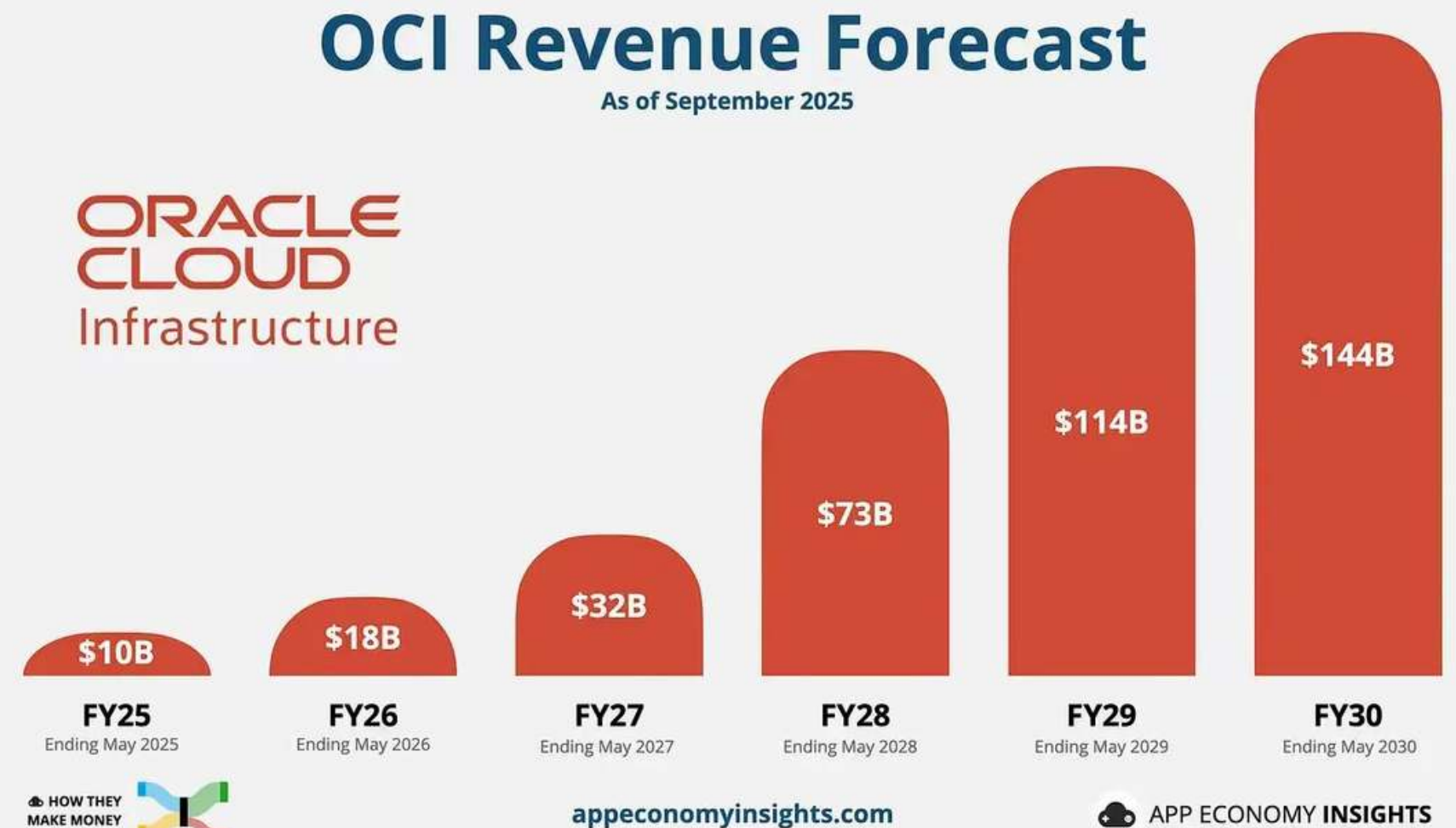
Sherwood

← → 🔍 云已经是稳定行业，那“增量”变的更重要了



Azure的季度增长绝对值已经超过了AWS

靠的就是OpenAI

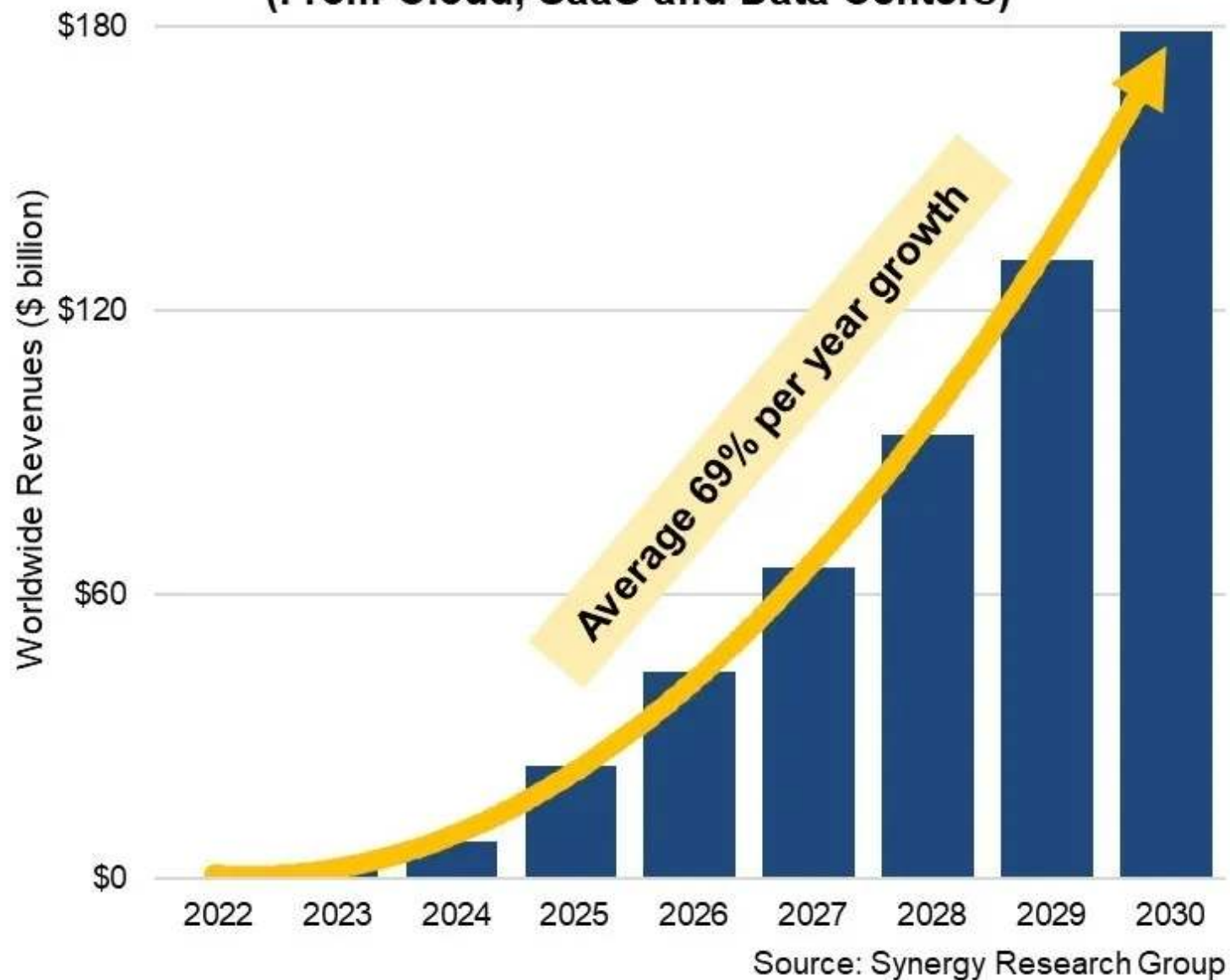


ORACLE云的收入增长曲线过于……

靠的当然也是OpenAI

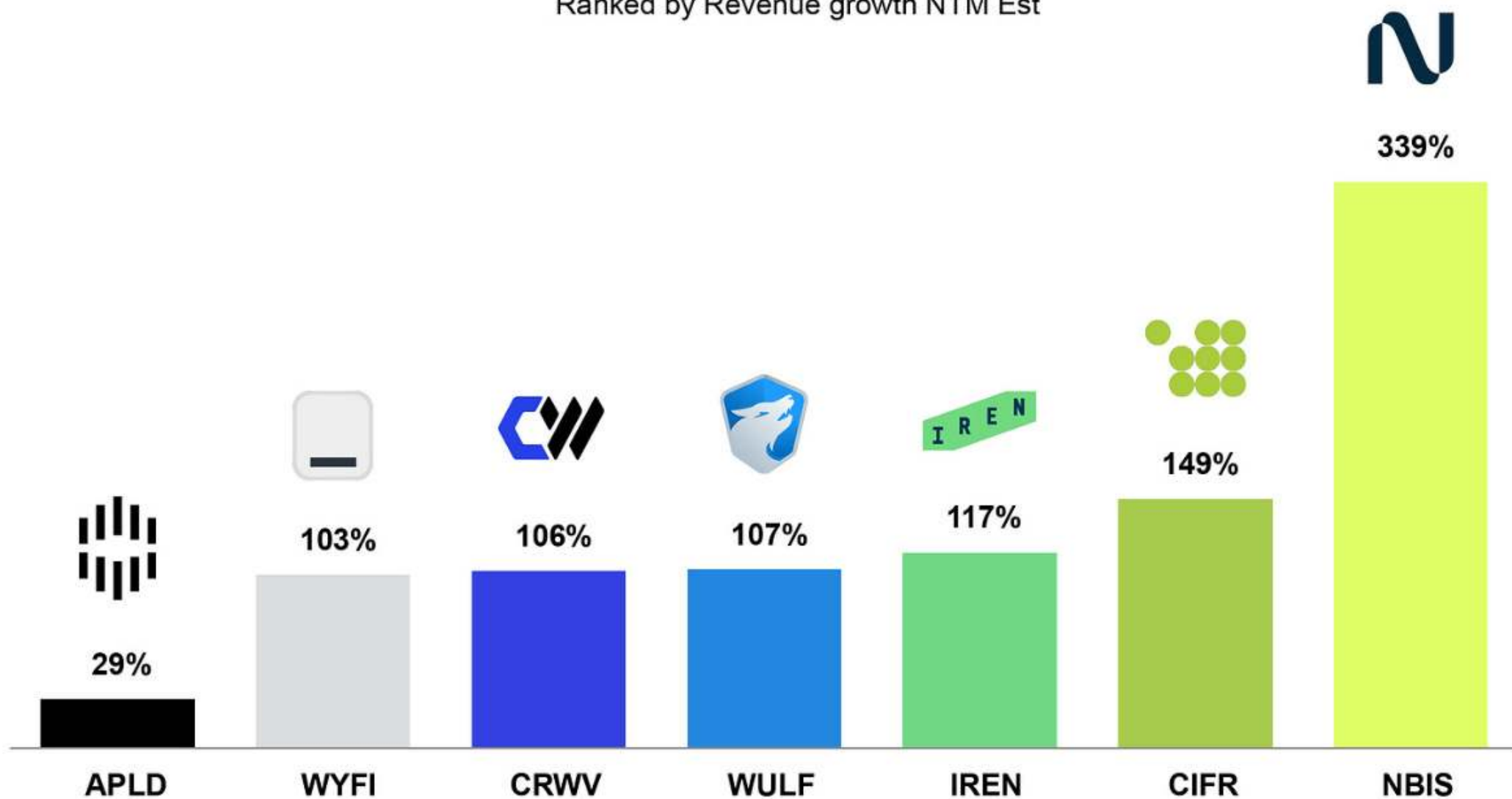
← → 🔍 已经存在的巨头增量之外，其实多了个“新”战场 - **Neocloud**

Neocloud Revenue Forecast
(From Cloud, SaaS and Data Centers)



Neo-Cloud Companies

Ranked by Revenue growth NTM Est



← → 🔍 Q 巨头为什么会允许这样一个市场出现？甚至也成为客户？

SemiAnalysis GPU Cloud ClusterMAX™ Rating March 2025				
Ranking	Neocloud			
				
		together.ai	NEBIUS	
			Azure	
	aws			
		TENSORWAVE		
				
	HOT AISLE			
Underperforming		IrisEnergy	GMI Cloud	
				

The Rise of AI Neoclouds



← → 🔍 所谓“术业有专攻”



按需 GPU 实例

如果您不需要长期容量承诺，可以使用按需 GPU 实例来处理更多工作负载。您可以快速按需启动突发容量，并享受灵活的按需定价。

GPU型号	GPU数量	显存 (GB)	虚拟 CPU	系统内存 (GB)	本地存储 (TB)	实例价格 (每小时)
NVIDIA GB300 NVL72	4 ¹	279	144	960	61.44	联系销售人员
NVIDIA GB200 NVL72	4 ¹	186	144	960	30.72	42.00 美元
NVIDIA B200	8	180	128	2,048	61.44	68.80 美元
NVIDIA RTX PRO 6000 Blackwell 服务器版	8	96	128	1,024	7.68	20.00 美元
NVIDIA HGX H100	8	80	128	2,048	61.44	49.24 美元



← → 🔍 Neocloud战场有很多Nvidia的亲儿子和干儿子们.....

NVIDIA Q2 HOLDINGS BREAKDOWN: OVER 90% BET ON COREWEAVE



CRWV
CoreWeave
Data Center Solutions
Holding Value: **\$3.96B**
Holding Percentage: **91.36%**
Nvidia added **95,100** shares, boosting its stake from 78.17% to **91.36%**.

ARM
Arm Holdings
Semiconductor
Holding Value: **\$178M**
Holding Percentage: **4.11%**

APLD
Applied Digital
Data Center Solutions
Holding Value: **\$77.70M**
Holding Percentage: **1.79%**

NBIS
NEBIUS
Data Center Solutions
Holding Value: **65.87M**
Holding Percentage: **1.52%**

RXRX
Recursion Pharmaceuticals
AI Healthcare
Holding Value: **\$38.99M**
Holding Percentage: **0.90%**

WRD
WeRide
Autonomous Driving
Holding Value: **13.70M**
Holding Percentage: **0.32%**

CoreWeave
\$138.94

+\$98.94 (+247.35%) Year to date



Nebius Group
\$126.19

+\$106.19 (+530.95%) Past year

NEBIUS



← → 🔍 虽然可能是我最熟悉的板块，但已经越来越看不懂了



资本趋势

二级市场

- MAG7的叙事延续
- 那些“妖艳异常”的狠角色
- 延展行业
-

一级市场

- AI是事实上的核心叙事
- Big Deal + Big Fund
- 某种程度上的“被拉爆”
-

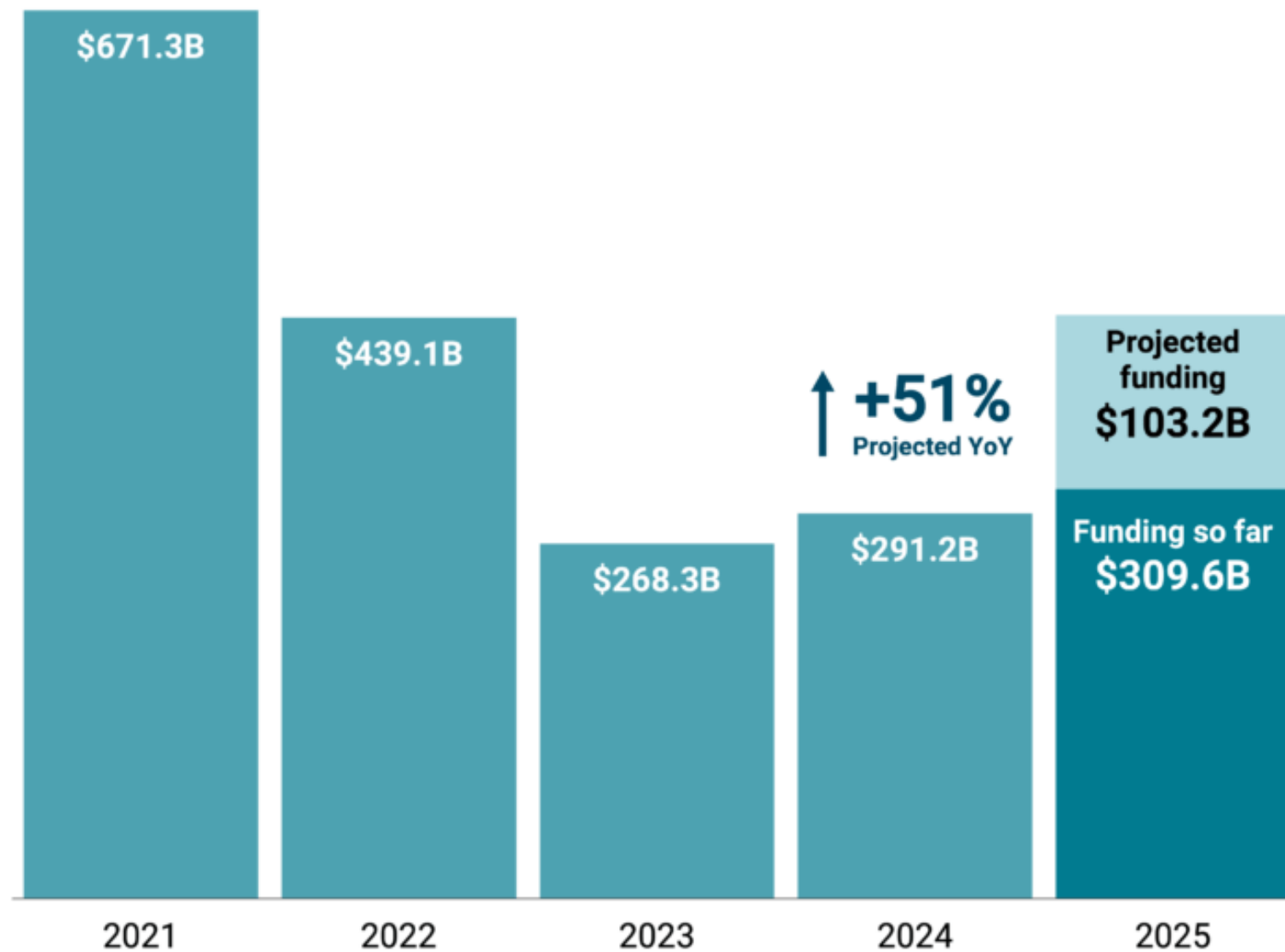
退出

- IPO
- 并购
- “野路子”
-

← → 🔍 市场整体而言，当然算得上是一个“好年份”，但.....

Annual funding reaches \$310B, a 3-year high

Annual equity funding

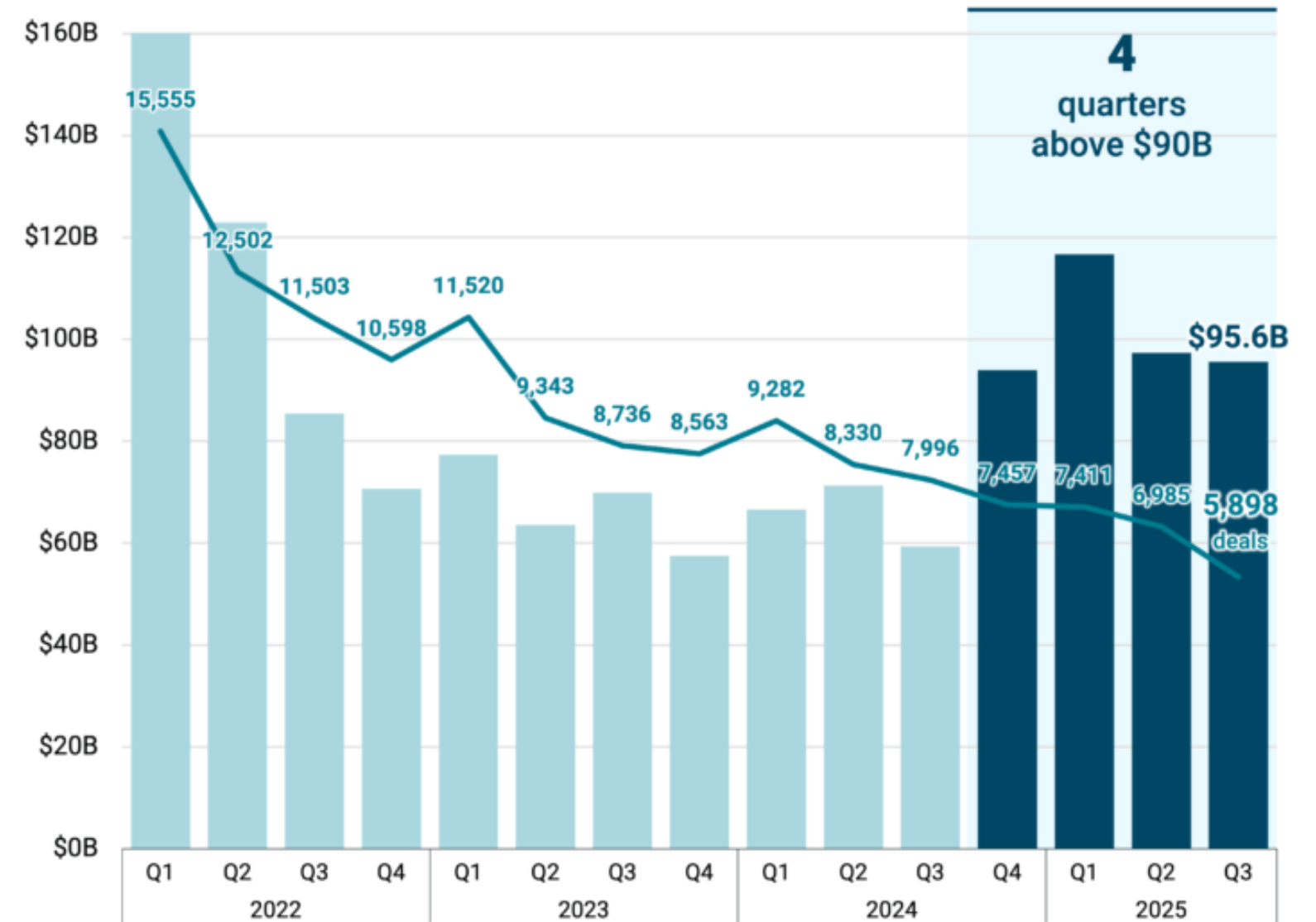


State of Venture Q3'25
© 2025 CB Insights. Equity deals only.

CBINSIGHTS

Funding rebounds in 2025: Q3 marks 4th consecutive quarter above \$90B

Quarterly equity funding & deals



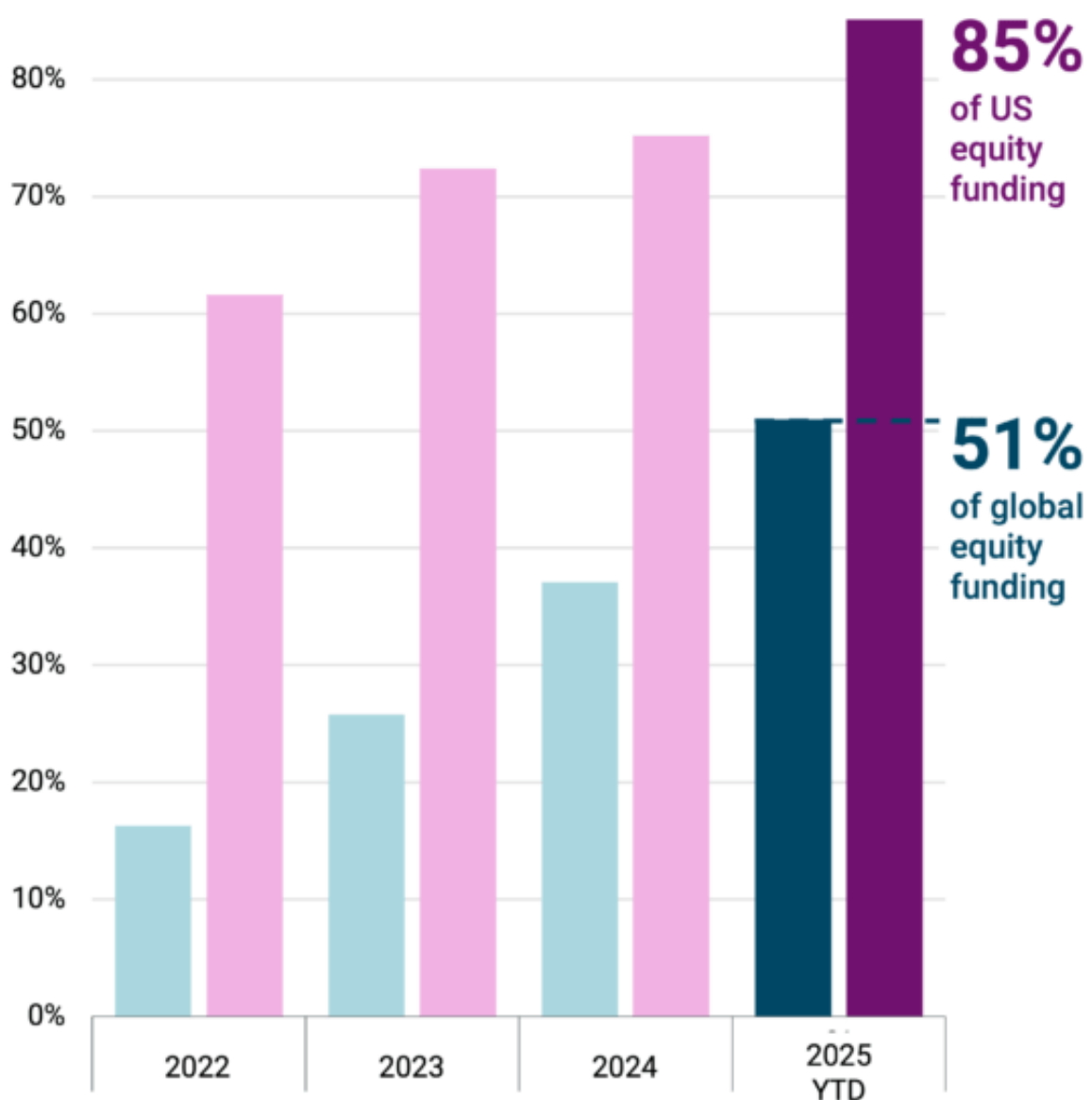
State of Venture Q3'25
© 2025 CB Insights. Equity deals only.

CBINSIGHTS

← → 🔍 AI当然是绝对意义上的**核心主题**

AI is eating venture capital

AI startups' annual share of global equity funding and US funding



State of Venture Q3'25
© 2025 CB Insights. Equity deals only.

Top AI deals in Q3'25

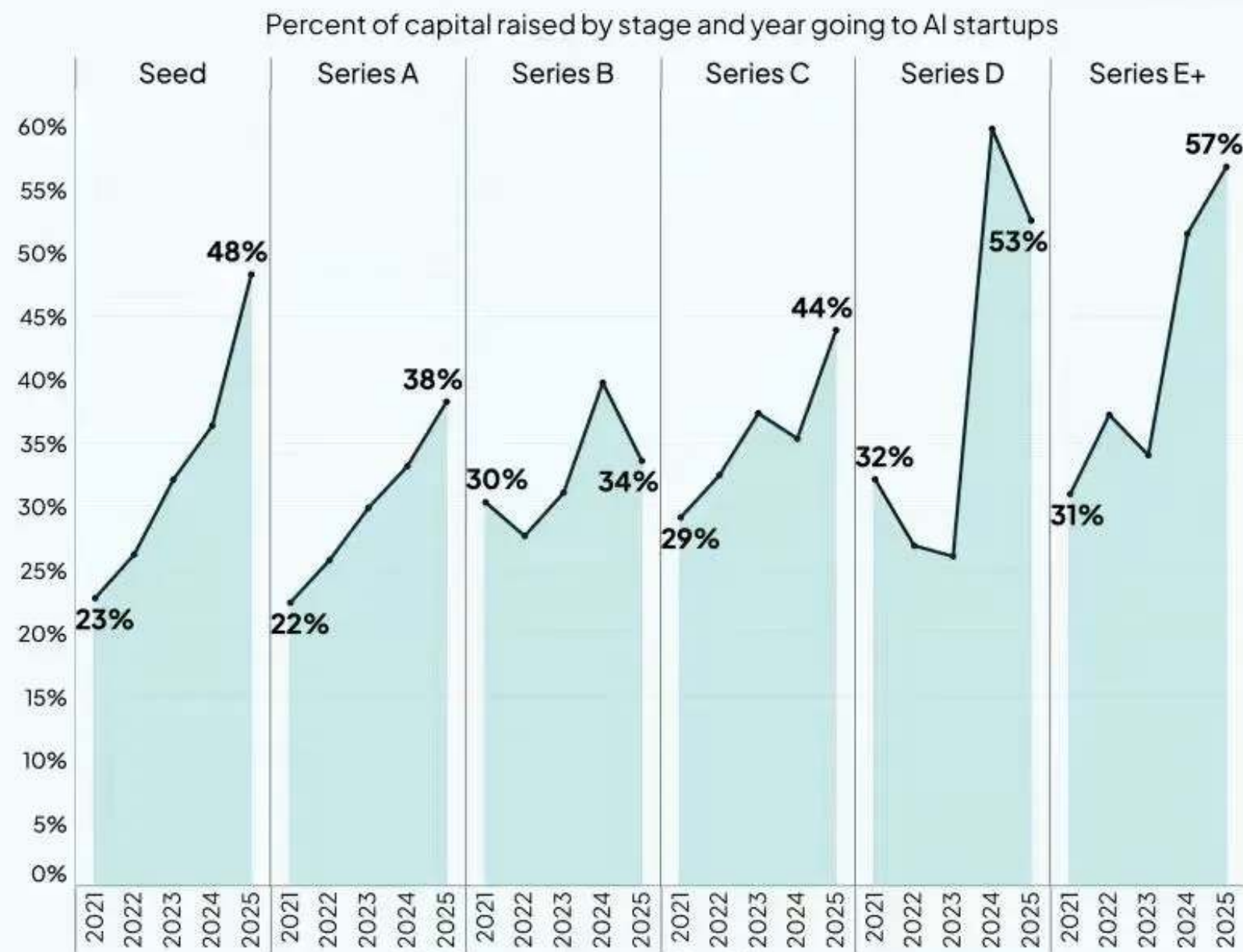
ANTHROPIC
\$13.0B
Mosaic: 966
OpenAI
\$8.3B
Mosaic: 909
Mistral AI
\$1.5B
Mosaic: 930
Princeton Digital Group
\$1.3B
Mosaic: 849
NSCALE
\$1.1B
Mosaic: 899
databricks
\$1.0B
Mosaic: 958
FIGURE
\$1.0B
Mosaic: 913

CBINSIGHTS

How many more AI startups can we fund?

The share of capital devoted to AI companies is rising relentlessly

Data: \$593 billion invested into US startups that use Carta | Q1 2021 – Q2 2025



©2025 eShares Inc., d/b/a Carta Inc. ("Carta"). All rights reserved.

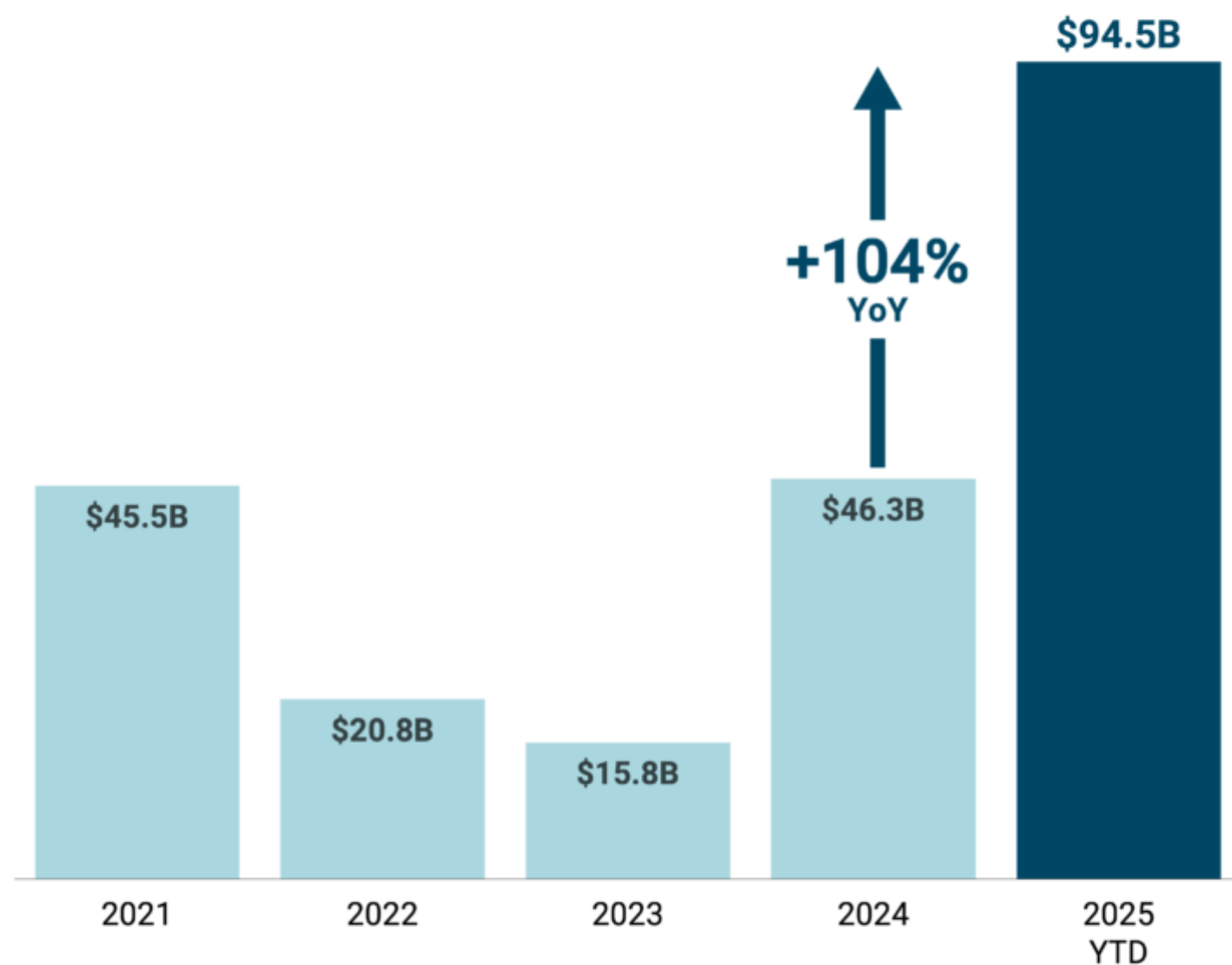
carta

Subscribe for more startup data: carta.com/data

← → 🔍 后期的偏好，造成的是big deal横行

Rise of the decacorn era

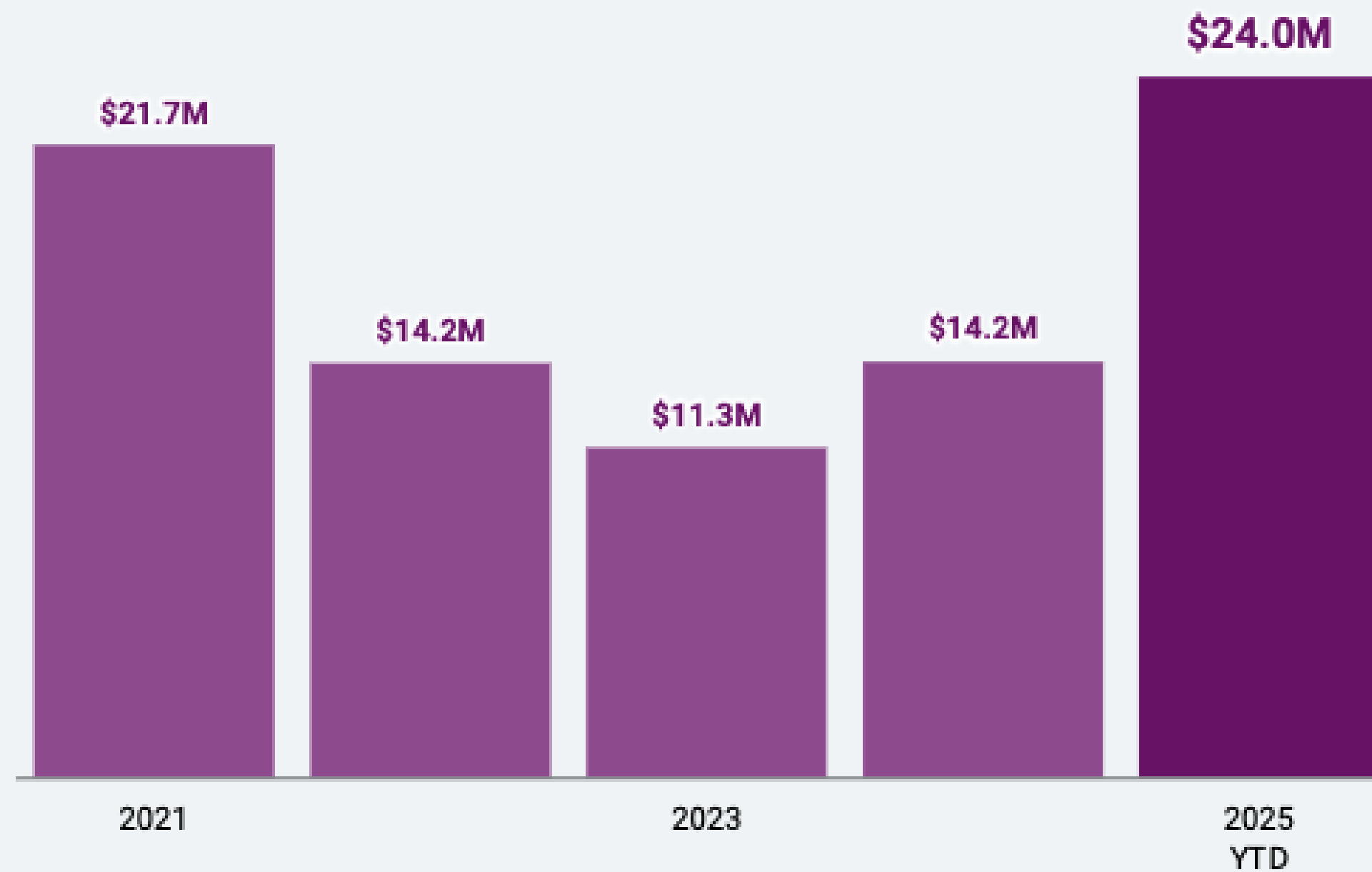
Record funding is flowing to the biggest players



State of Venture Q3'25
© 2025 CB Insights. Equity deals only.

CBINSIGHTS

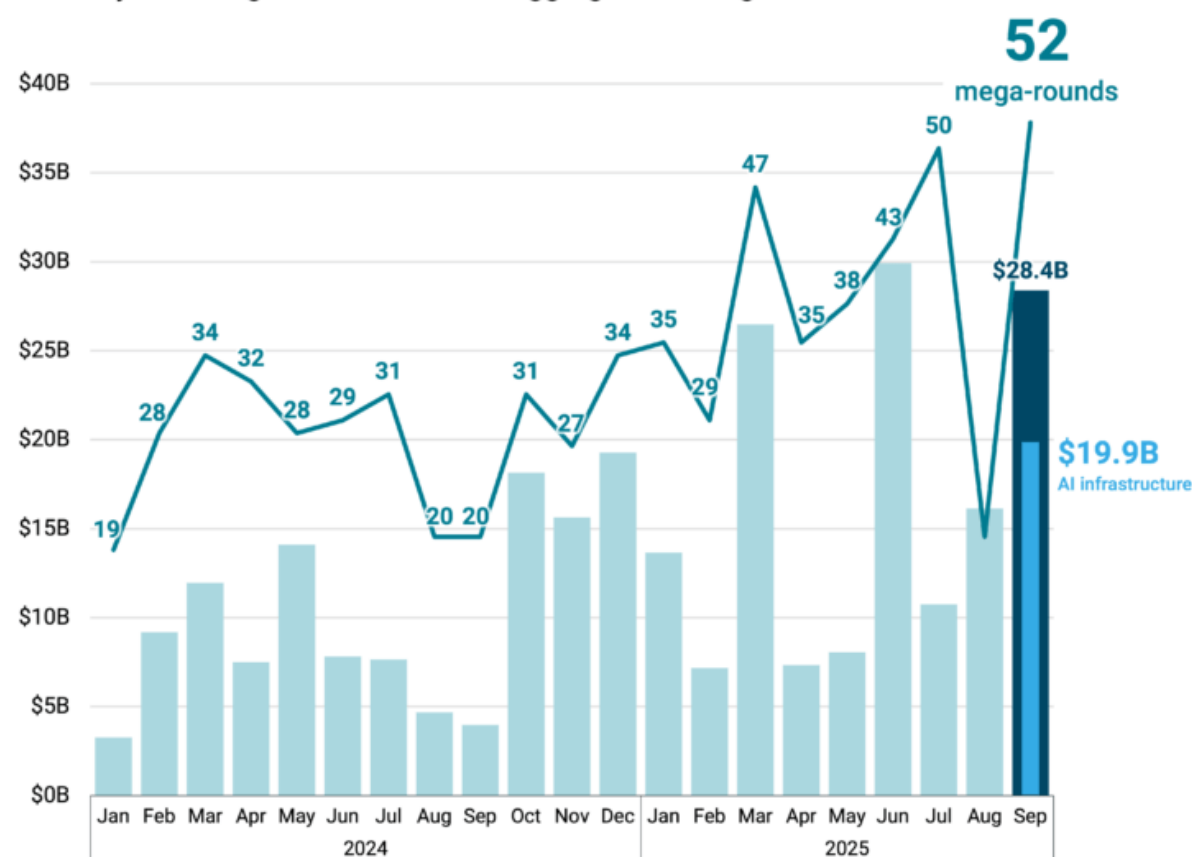
Average Deal Size



← → 🔍 🔍 细分行业角度看，AI Infra可能是最多的.....

AI infrastructure drives record mega-round deals

Monthly tech mega-round deals and aggregate funding



State of Venture Q3'25
© 2025 CB Insights. Data as of 9/30/2025. Note: tech mega-rounds defined as equity deals of \$100M or more raised by tech companies.

CBINSIGHTS

State of AI infrastructure tooling.

Data Labeling / Management	Evaluations / Observability	Inference	
scale Snorkel	MERCOR TURING braintrust Patronus AI Galileo Haize Labs arize	Modal Fireworks AI	baseten together.ai
Browser Infrastructure	AI Agent DevOps	AI Agent Security	AI Agent Orchestration
Browserbase Browser Use parallel Firecrawl	Coder Gitpod E2B Daytona	Arcade descope clerk KEYCARD	LangChain LlamaIndex Pydantic openai
Vector Databases	Model Security	Fine-Tuning	
chroma Pinecone turbopuffer	TROJ.AI CALYPSO AI Prompt:	Adaptive ^{ML} LAMINI	clarifai replicate

← → 🔍 ChatGPT发布之后，一级市场已经**新投出100家独角兽**

The 100 AI unicorns since ChatGPT launched in numbers

CBINSIGHTS

A deep dive on the cohort using CB Insights data

LLM developers count the most unicorns



Note: Companies may appear in multiple markets.

2 China-based startups have largest Mosaic jump

6-month change in Mosaic score



Most likely acquisition targets span industries

Top AI unicorns by M&A Probability



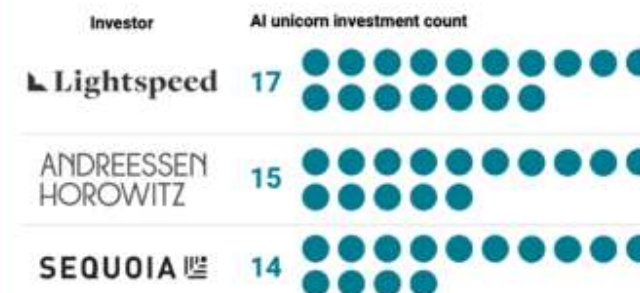
M&A Probability predicts the likelihood of a private company achieving an M&A exit within 2 years.

3 unicorns have exited



Lightspeed is the most active VC investor

All-time investment count in the 100 AI unicorns by VC



Big tech has backed over a third of the list



By unique companies backed



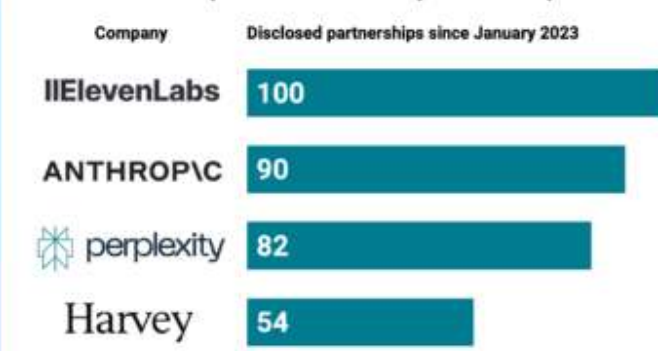
Note: Includes venture arm investments.

Frontier AI labs are most well-funded

Top AI unicorns by total funding



ElevenLabs captures the most partnerships



Skild AI has the highest valuation per employee



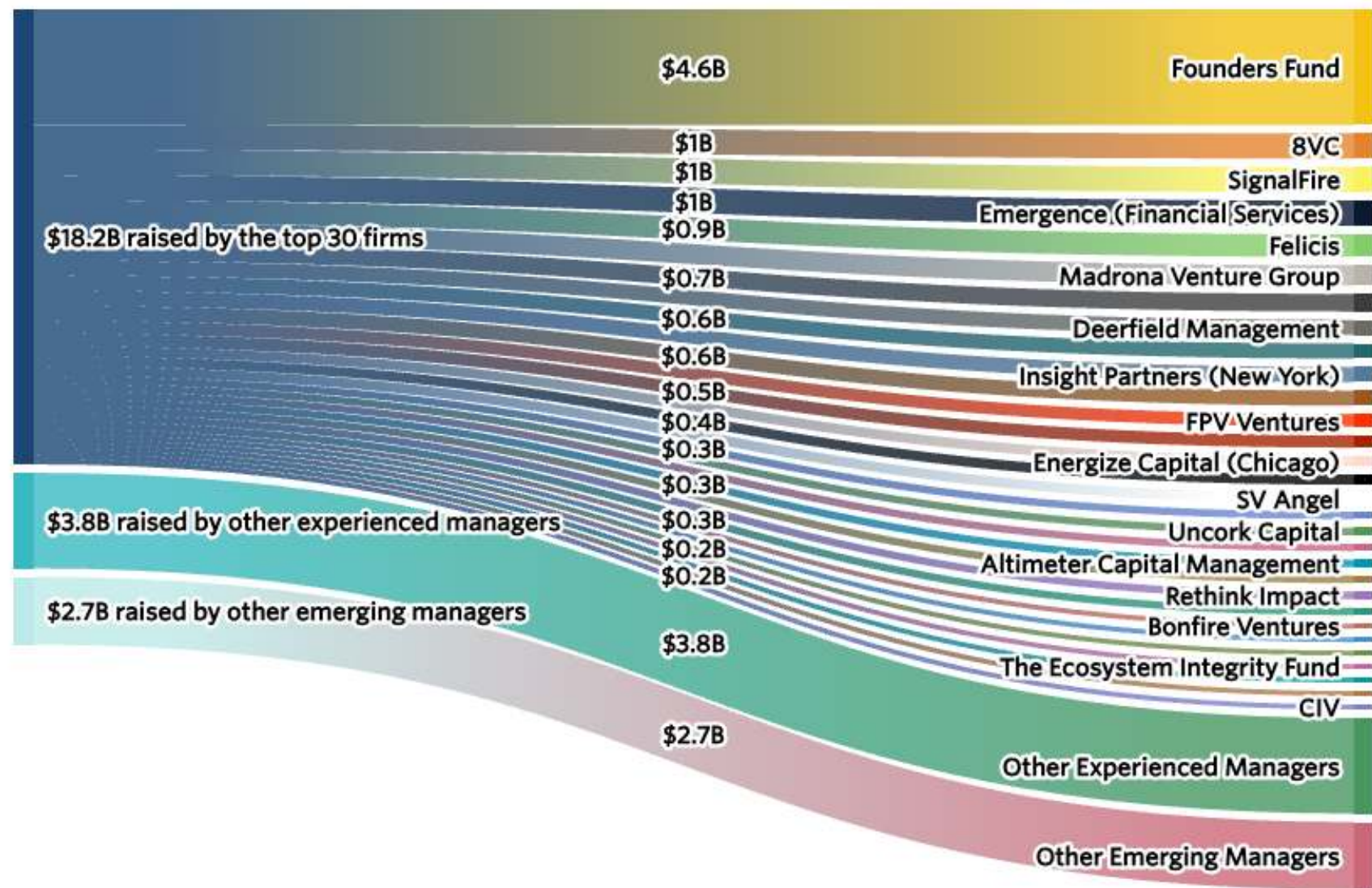
US companies make up two-thirds of the list

Geographic breakdown of the 100 AI unicorns since ChatGPT launched



← → 🔍 🔍 Big fund的图其实是上次的，那代表着.....

Founders Fund reigns in H1 2025

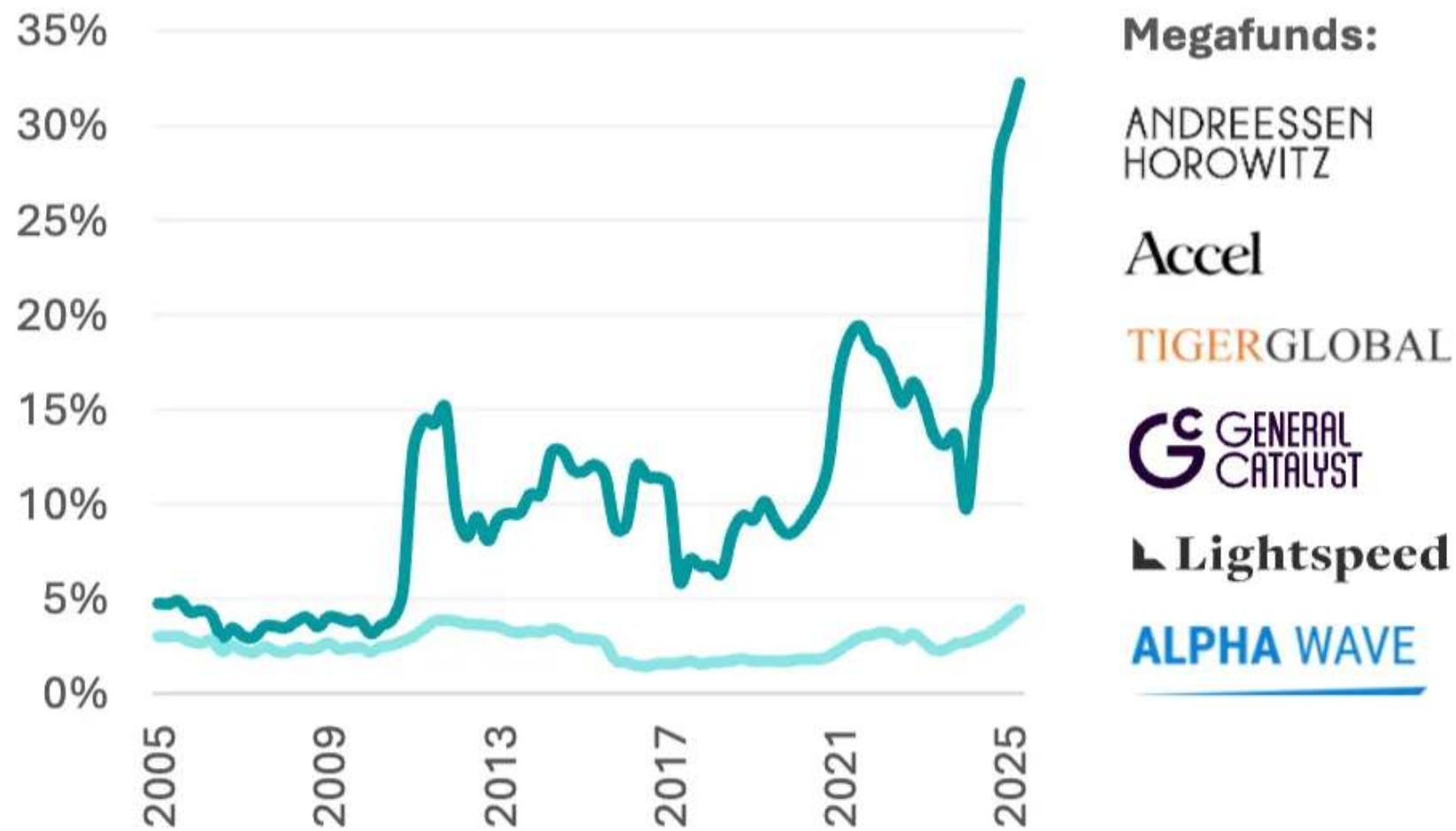


Source: PitchBook • Geography: US • As of June 30, 2025

Mega Managers Are Eating the World

Share of US VC Investment With Mega Manager³ Participation

Share of VC Deals by: — Deal Dollars — Deal Count



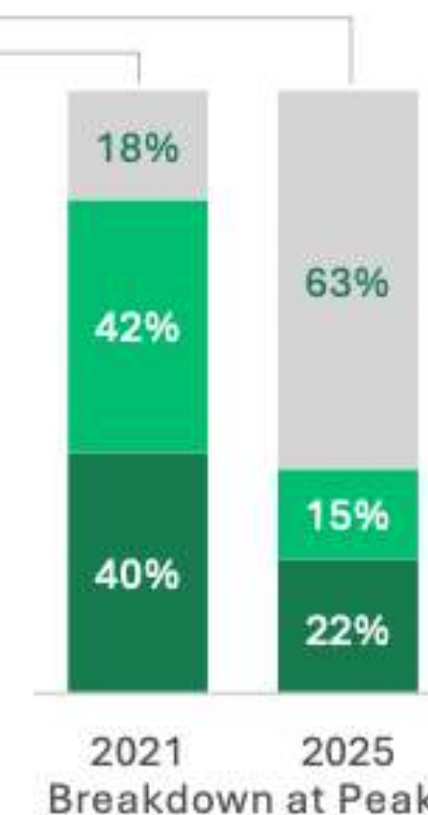
← → 🔍 5亿美金+的deal占这么多，那代表着什么？

Kind of a Big Deal: Mega Rounds Dominate VC Investment

Annualized US VC Investment by Deal Size (Trailing Three-Month Total x 4)¹

Deal Size: ■ Under \$100M ■ \$100M-\$500M ■ Over \$500M

\$450B
\$400B
\$350B
\$300B
\$250B
\$200B
\$150B
\$100B
\$50B
\$0B



小问题+1：
你们知道一个常规的VC fund有多大么？

← → 🔍 虽然可能是我最熟悉的板块，但已经越来越看不懂了



资本趋势

二级市场

- MAG7的叙事延续
- 那些“妖艳异常”的狠角色
- 延展行业
-

一级市场

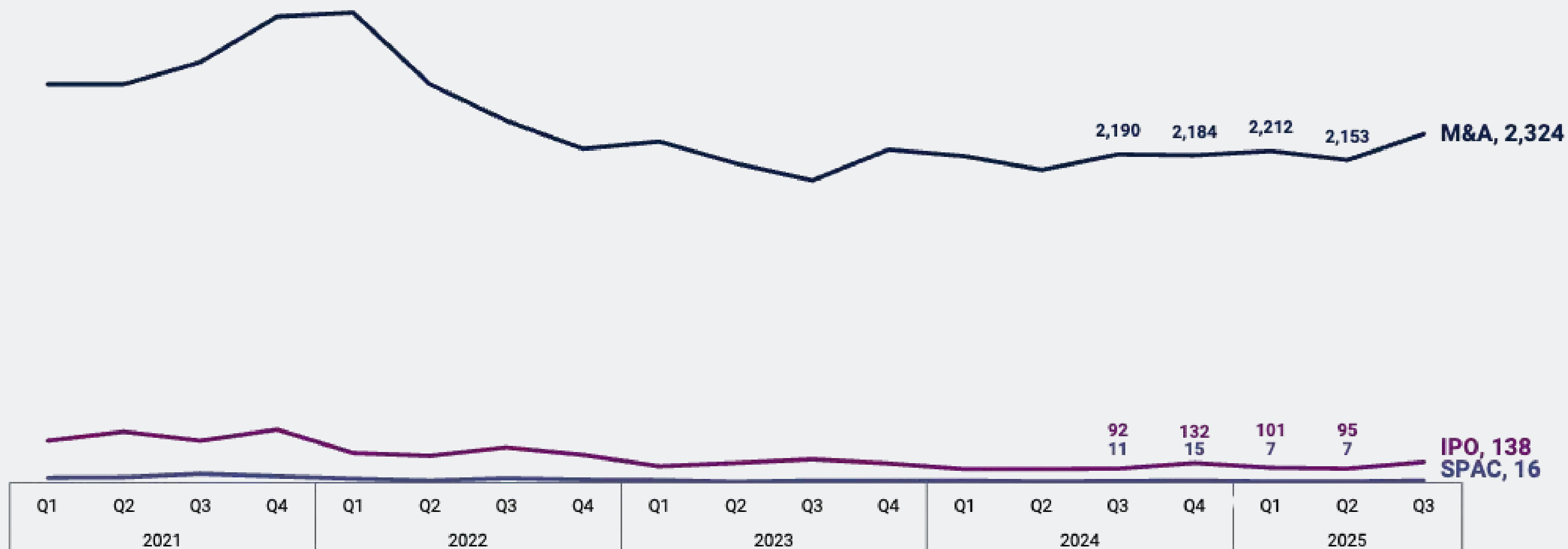
- AI是事实上的核心叙事
- Big Deal + Big Fund
- 某种程度上的“被拉爆”
-

退出

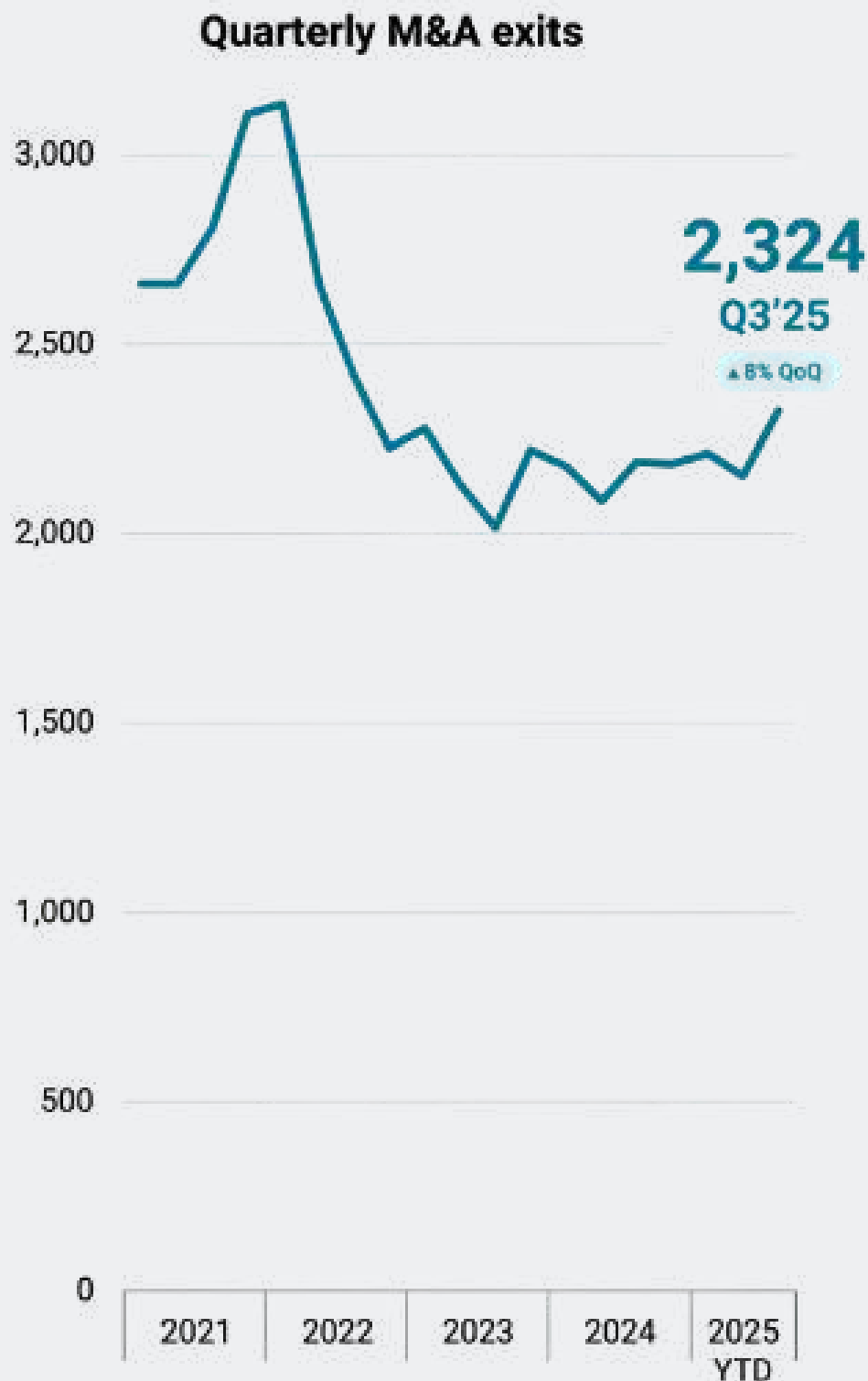
- IPO
- 并购
- “野路子”
-

← → Q Q 当然了，有流动性总比没有强.....尤其在退出上

Quarterly exits



← → 🔍 🔍 用过好几次的标题了——绝对意义上的**并购大年**



Acquired Taste: Startups Buy More Startups

VC-Backed M&A Deals (Excluding LBOs) and the Percentage with a VC-Backed Buyer¹



M&A All Day: AI Unicorns Are Buying to Scale

Notable VC-Backed M&A Buyers Since 2023²

Buyer	TEV	M&A Deals	Focus
OpenAI	\$300B	5	Horizontal AI Platform
stripe	\$92B	6	Embedded Payments
databricks	\$62B	17	Infrastructure Tools for AI
INFINITE REALITY	\$16B	5	AI for Social Media
Figma	\$13B	6	AI for Design
GrubMarket	\$3.5B	17	AI for Food Distribution

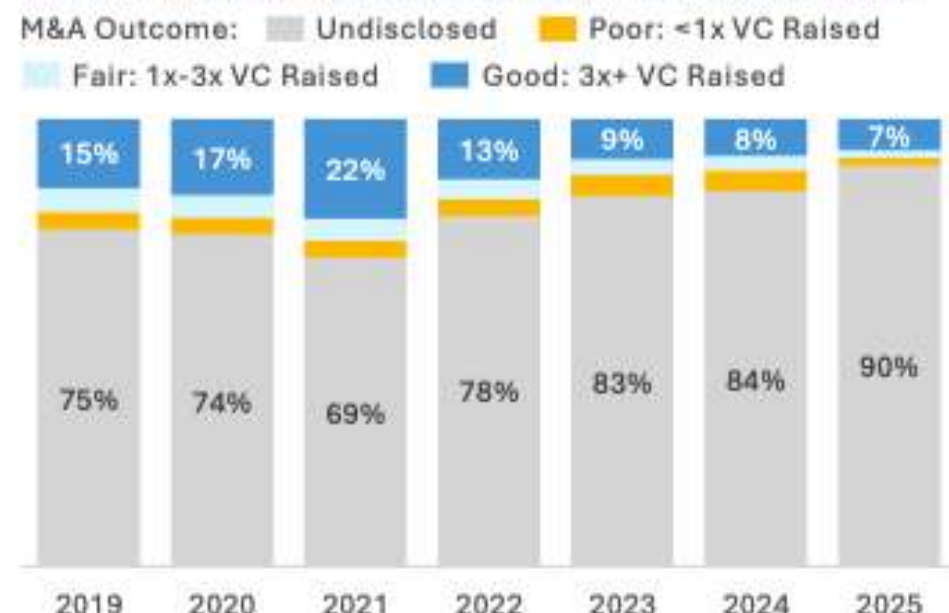
M&A Disappeared; Now It's Coming Back...

Number of US VC-Backed M&A Deals per 100 VC Deals



...But Fewer Deals Have Favorable Outcomes

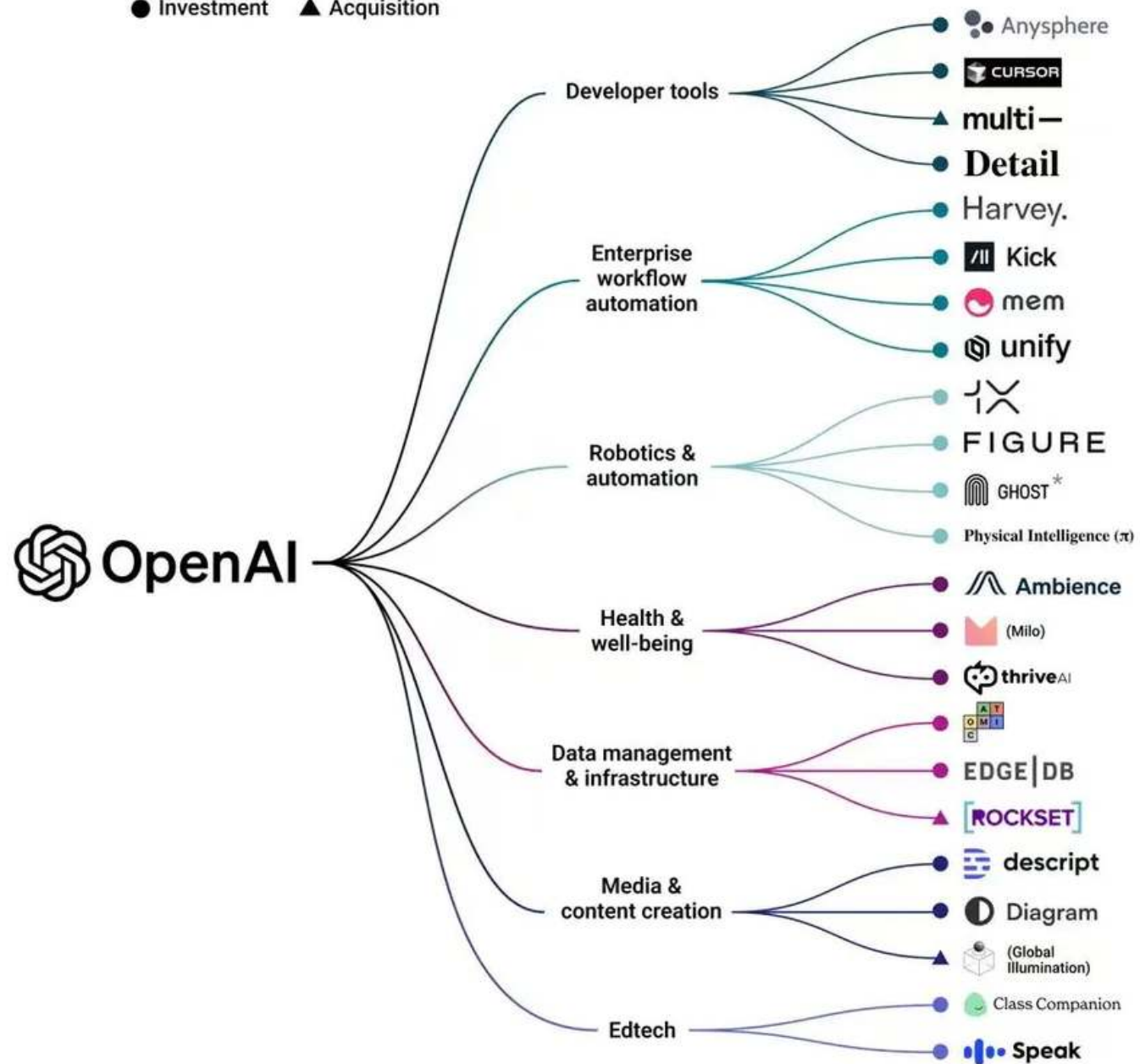
Share of M&A Deals by Outcome (Deal Price/Total VC Raised)



← → 🔍 🔍 OpenAI某种程度上也为市场提供了一定的流动性

OpenAI's startup bets point to new tech capabilities and use cases

● Investment ▲ Acquisition



Source: CB Insights. Based on OpenAI's investments and acquisitions since 2022. Data as of July 2024. Categories are not mutually exclusive, and companies are organized by primary application. Commercial Maturity measures a private company's current ability to compete for customers or serve as a partner. If a company lacks a Commercial Maturity score, this means that there isn't enough available information to measure its maturity, or that it has exited or shut down. *Ghost Autonomy shut down in April 2024.

CBINSIGHTS



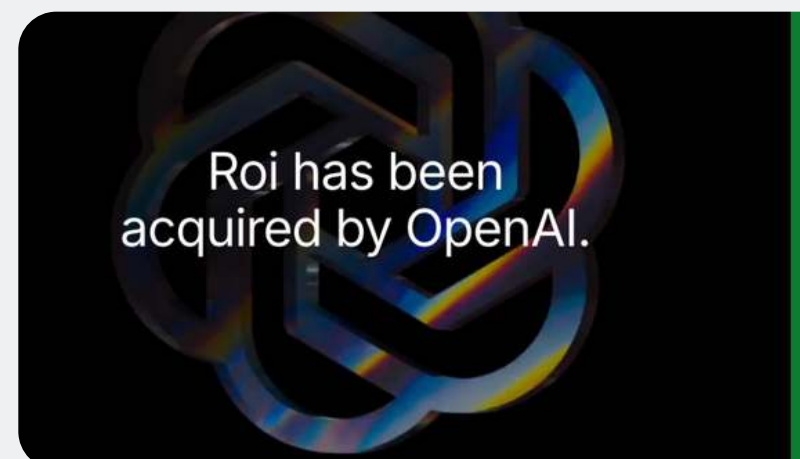
OpenAI acquires product testing startup Statsig and shakes up its leadership team

Maxwell Zeff — 12:04 PM PDT · September 2, 2025



OpenAI buys Sky, an AI interface for Mac

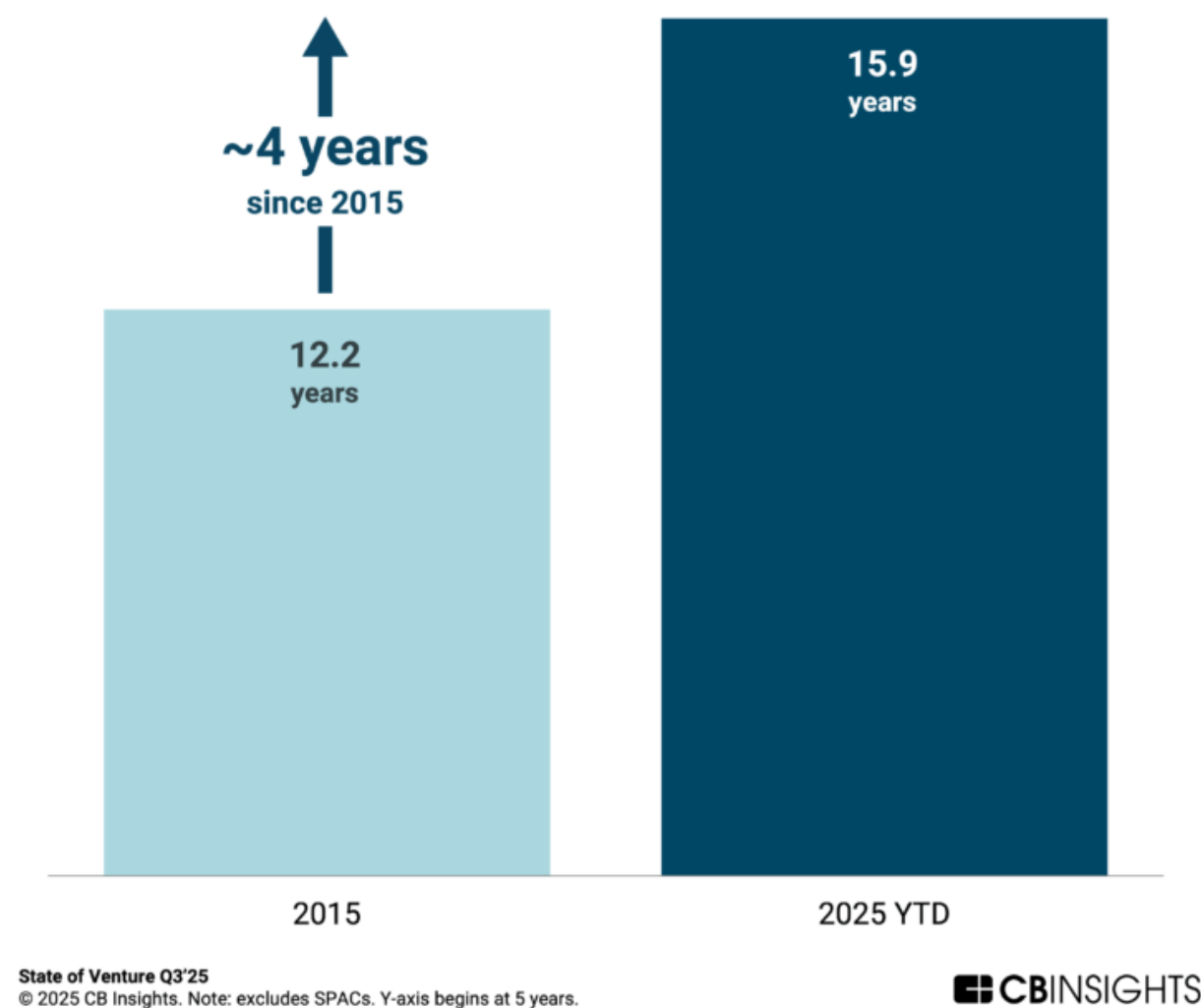
Sarah Perez — 1:53 PM PDT · October 23, 2025



With its latest acqui-hire, OpenAI is doubling down on personalized consumer AI

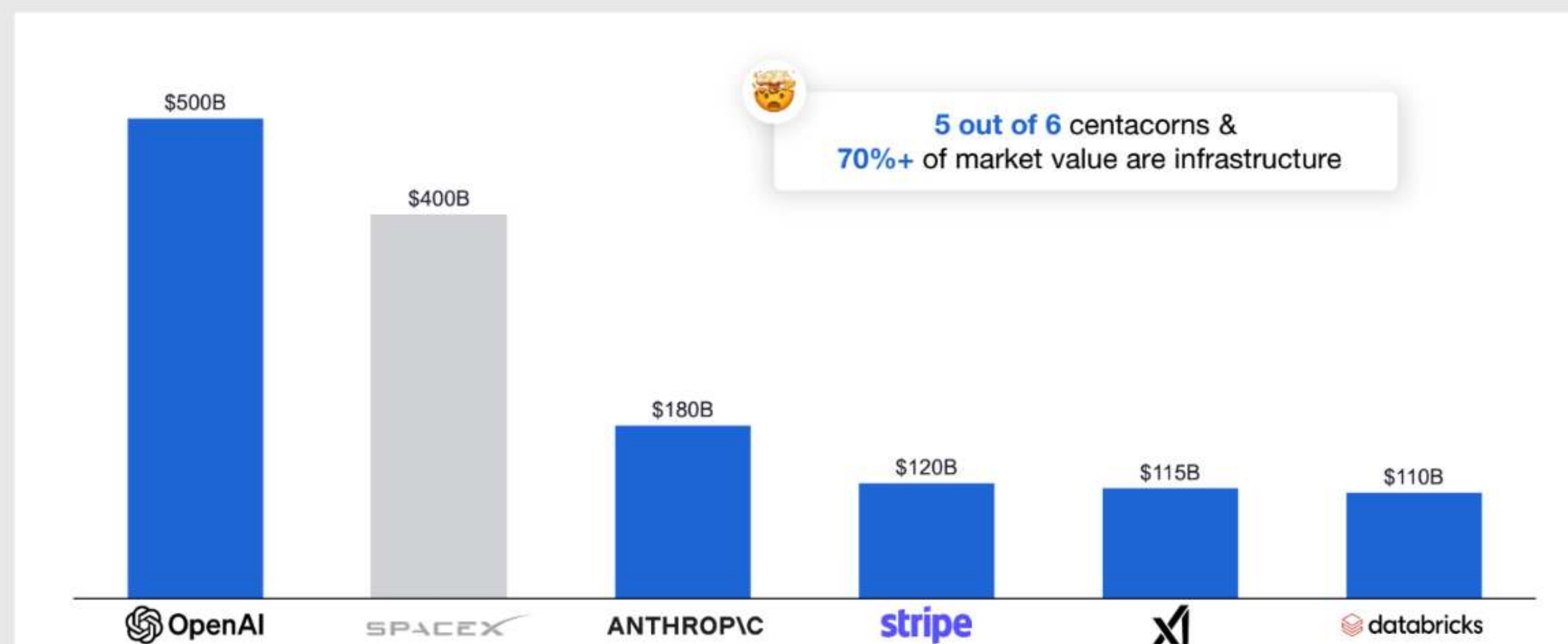
← → 🔍 流动性好，就会有大家伙选择暂时不上市

Tech companies are staying private longer



...and constitutes the lion's share of private centacorns

US private technology companies with \$100B+ valuations



← → 🔍 🔍 那么就必然会有“野路子”出现

	Company	90D Return △	90D % △	Implied Mkt Cap*	90D Implied Mkt △	90D Bid/Ask Vol
	Kraken		+50.44%	\$11.00B	+\$3.67B	
	SpaceX		+5.29%	\$411.00B	+\$20.00B	
	xAI		+12.39%	\$179.00B	+\$19.00B	
	Stripe		+22.56%	\$118.00B	+\$21.70B	
	Groq		+30.25%	\$7.91B	+\$2.98B	
	Anthropic		+127.43%	\$207.00B	+\$131.90B	
	Perplexity AI		+60.15%	\$20.40B	+\$7.70B	
	Anduril		+31.13%	\$54.90B	+\$13.10B	
	OpenAI		+70.77%	\$523.00B	+\$217.00B	
	ByteDance		+11.48%	\$371.00B	+\$38.00B	



Q3二手份额交易



Robinhood 



← → 🔍 我就是为了**这碗醋**包的这顿饺子.....



Sam和老黄到底在搞什么？



我们怎么一步步走到了这？



所谓“泡沫”到底是什么？



← → 🔍 Sam的最近几个月的工作就是：既往不恋-纵情向前



世人似乎都忘记
我之前是干嘛的了...



← → 🔍 命运的齿轮可能转动很久了，我们从**24年2月**开始看

TECH

OpenAI CEO Sam Altman
seeks as much as **\$7 trillion**
for new AI chip project:
Report

PUBLISHED FRI, FEB 9 2024 8:55 PM EST
UPDATED 6 HOURS AGO



24年2月，
我就说了一个数字：
7万亿美金

当时的OpenAI什么状态？

- 公司估值**860亿美金**
- 23年收入**16亿美金**
- 周活跃用户**1亿**

← → 🔍

我们把当时Sam要解的题列出来



AI
OpenAI says GPT-5 stacks up to humans in a wide range of jobs
Maxwell Zeff - Sep 25, 2025

AI
Sam Altman says ChatGPT has hit 800M weekly active users
Rebecca Bellan - Oct 6, 2025

SoftBank Greenlights Remaining \$22.5 Billion of OpenAI Investment
By Sri Muppidi

OpenAI completes for-profit restructuring and grants Microsoft a 27% stake in the company

25年年底这个时点
Sam把这些问题几乎全解了

← → 🔍 但即便是这样，Sam发现这依然远远不够，然后25年初.....



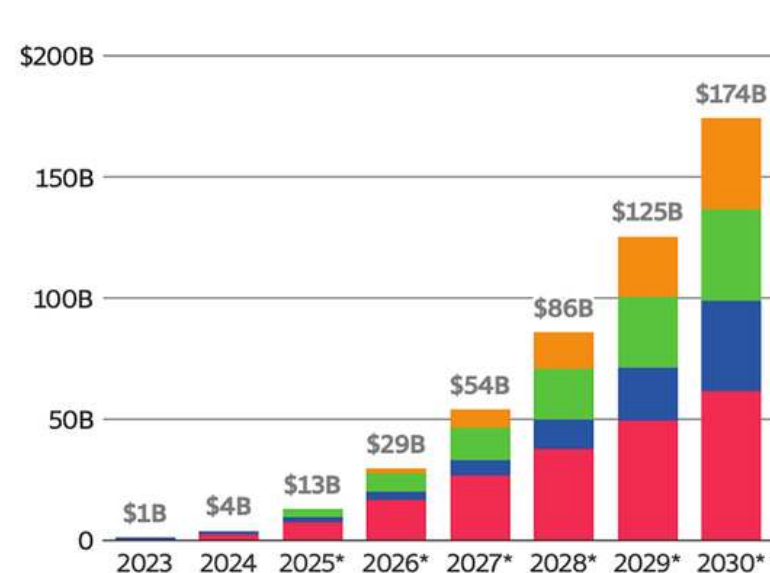
← → 🔍 Sam在某天应该是突然顿悟了，大饼可以画到2030

Rising Revenues

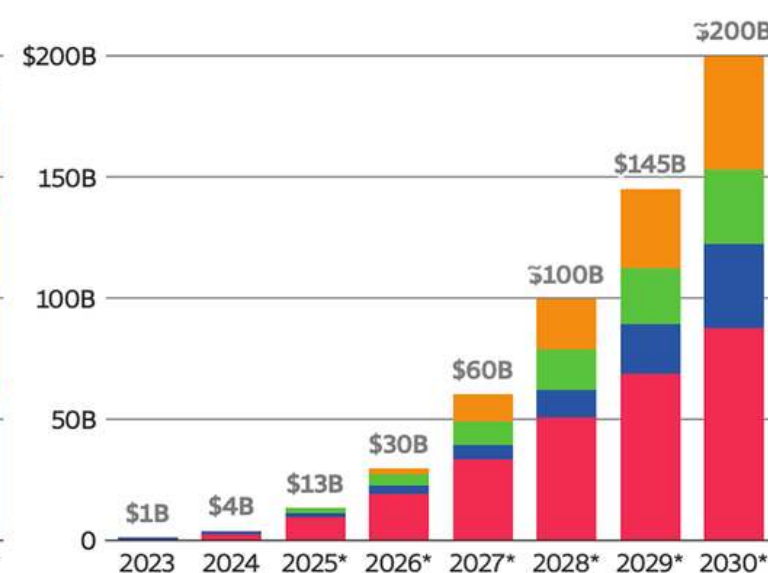
OpenAI has increased its revenue projections, anticipating more from ChatGPT and new products, including monetizing people who use ChatGPT for free.

● ChatGPT ● API ● Agents ● New products (including free user monetization)

Projection as of Q1 2025



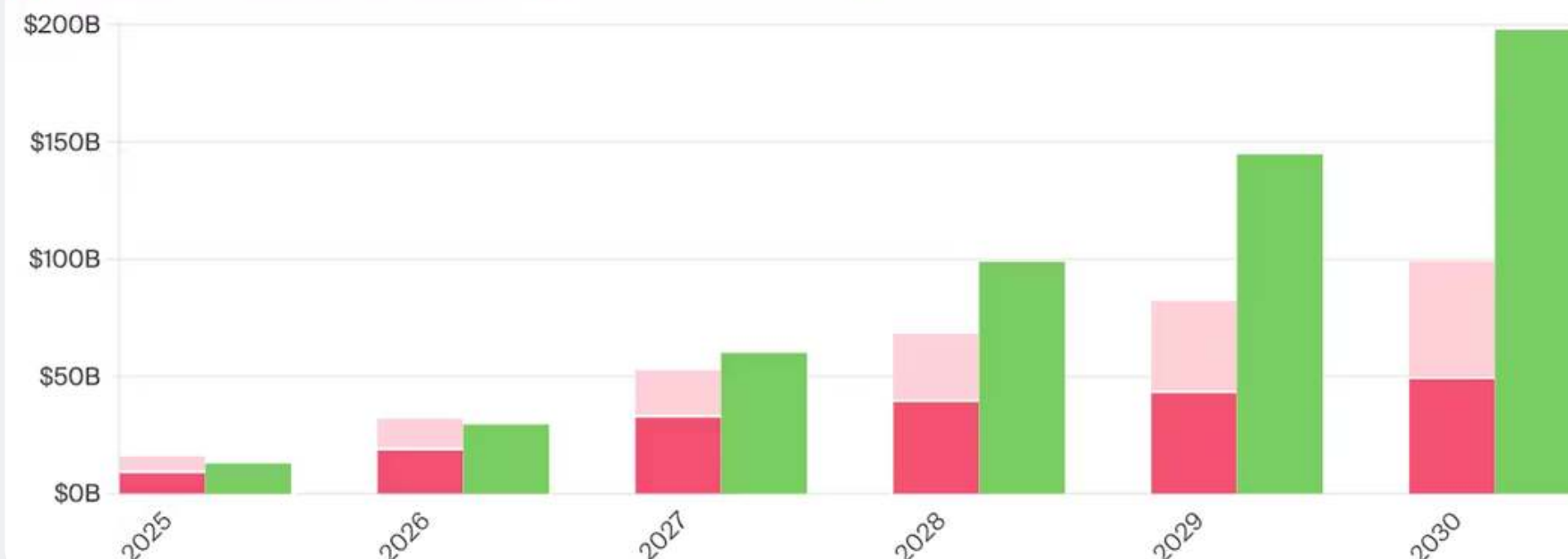
Projection as of Q3 2025



Compute Crunch

Even as OpenAI matures, it projects to spend half of its revenue on cloud computing costs—higher than any other company of its scale.

■ Research and Development Compute Cost* ■ Inference Compute Cost ■ Total Revenue



● Free cash flow projections as of Q1 2025 ● Projections as of Q3 2025



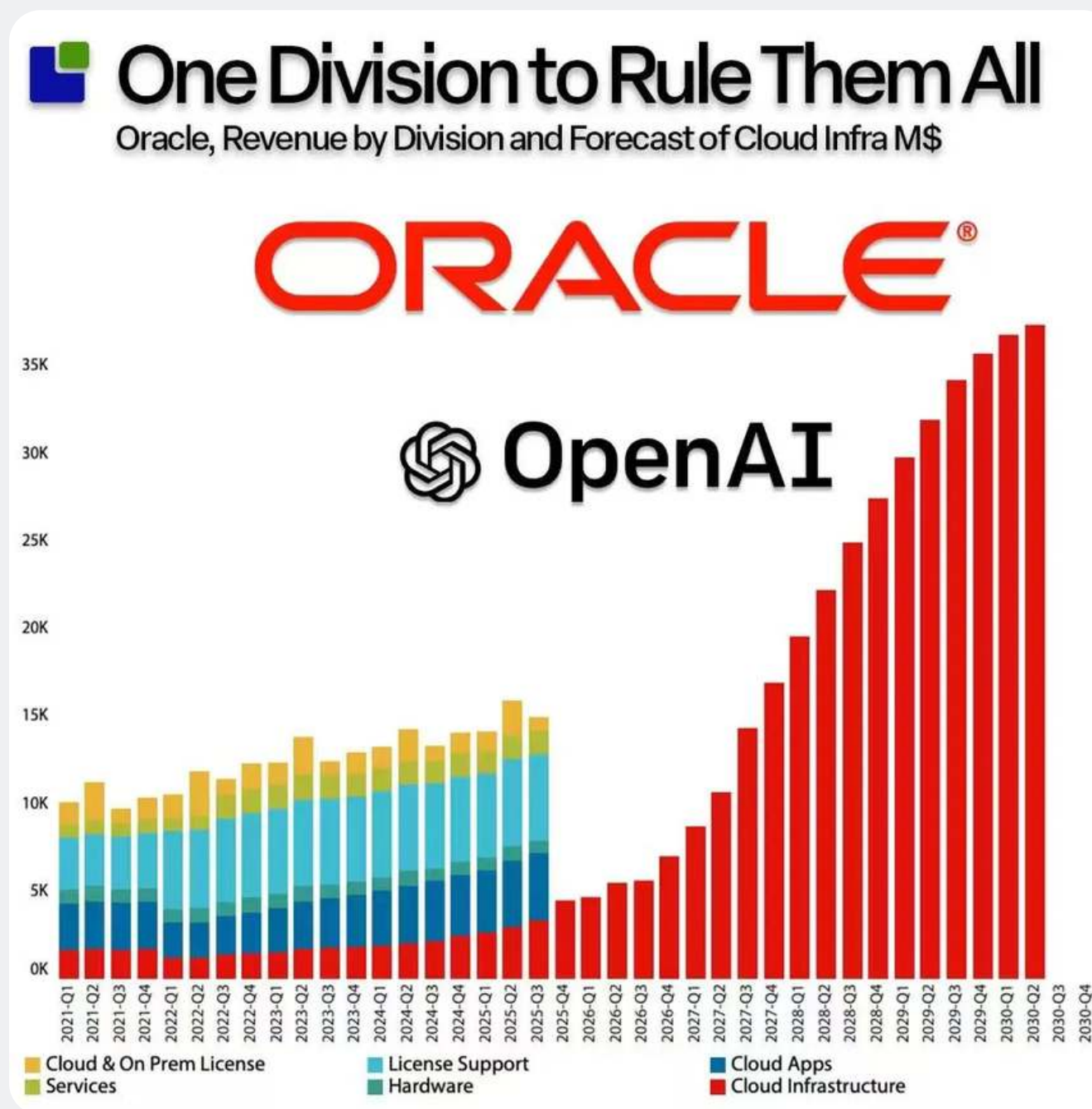
OpenAI expects business to burn \$115 billion through 2029, The Information reports

PUBLISHED SAT, SEP 6 2025 5:19 PM EDT

REUTERS

SHARE [f](#) [X](#) [in](#) [✉](#)

← → 🔍 既然我的预期已经打满到5年后，那么.....



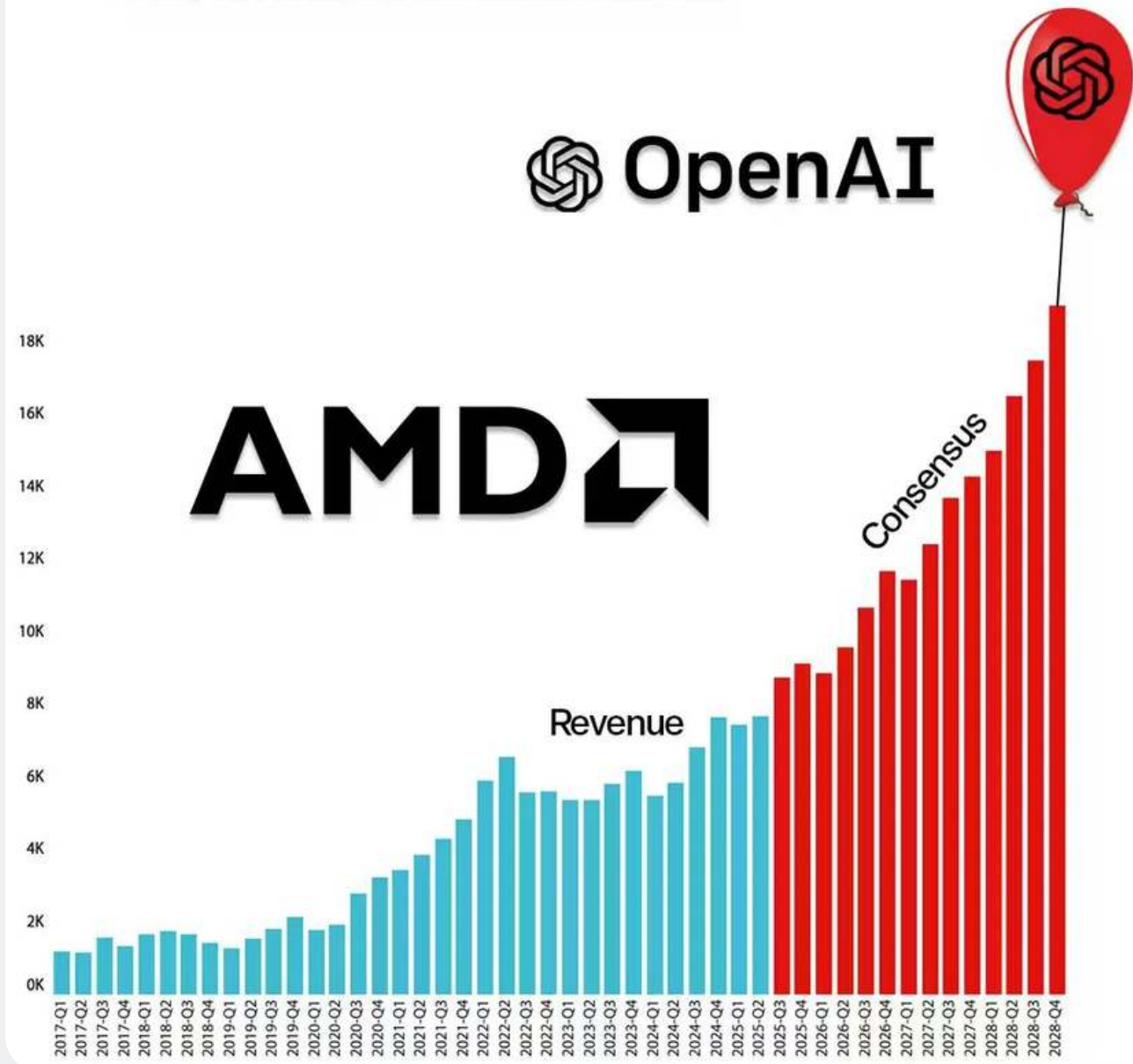
← → 🔍 我是金手指啊.....指谁谁起飞!

October 6, 2025 Company

AMD and OpenAI announce strategic partnership to deploy 6 gigawatts of AMD GPUs



Revenue for Equity: The New Circular Economy of AI
AMD Revenue and Consensus Revenue



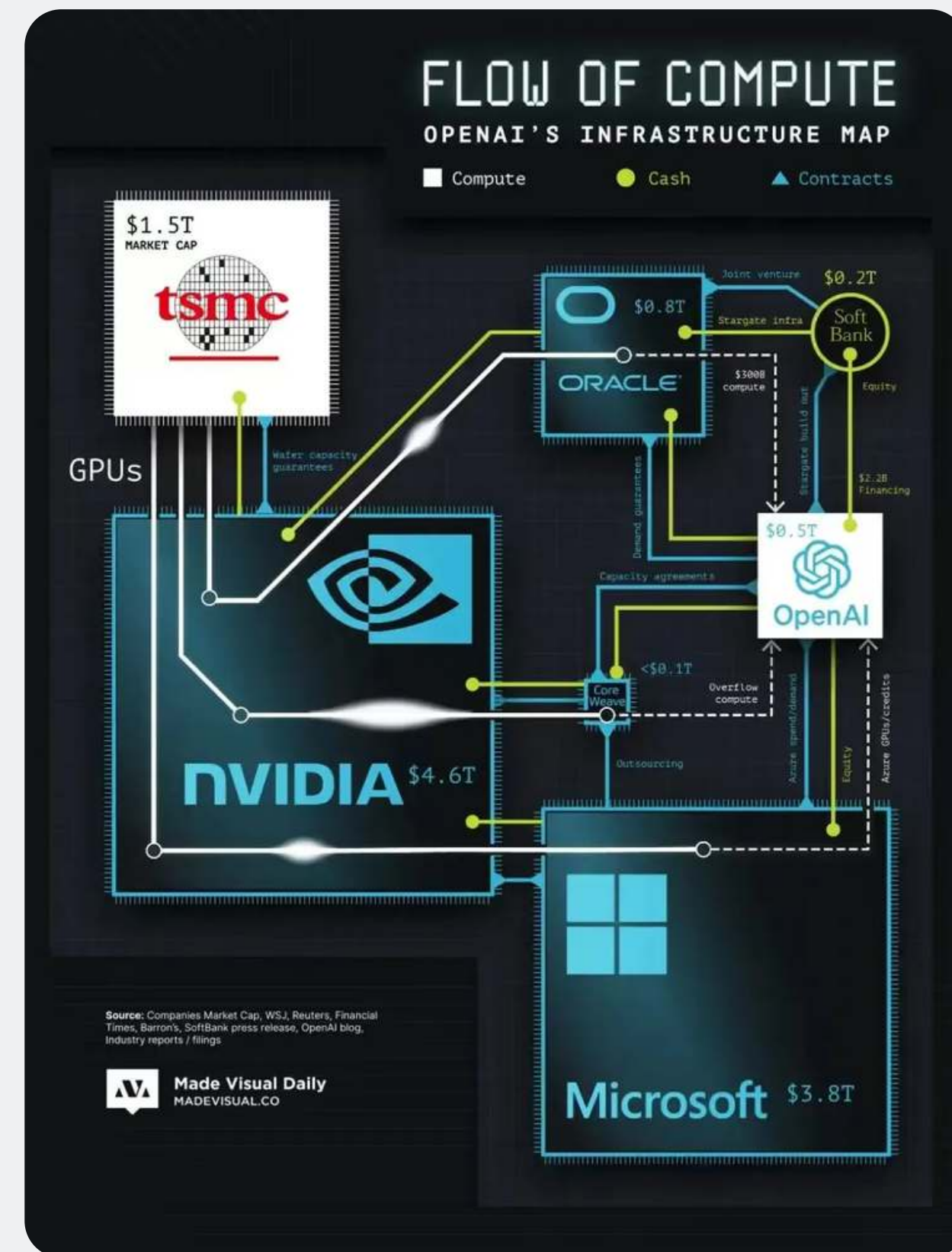
OpenAI and Broadcom announce strategic collaboration to deploy 10 gigawatts of OpenAI-designed AI accelerators

Multi-year partnership enables OpenAI and Broadcom to deliver accelerator and network systems for next-generation AI clusters.



Ask ChatGPT

← → 🔍 Q Nvidia老黄其实也“乐见其成”.....



← → 🔍 Q Nvidia现在是**什么江湖地位**?

NVIDIA MAKES HISTORY AS THE FIRST \$5 TRILLION COMPANY

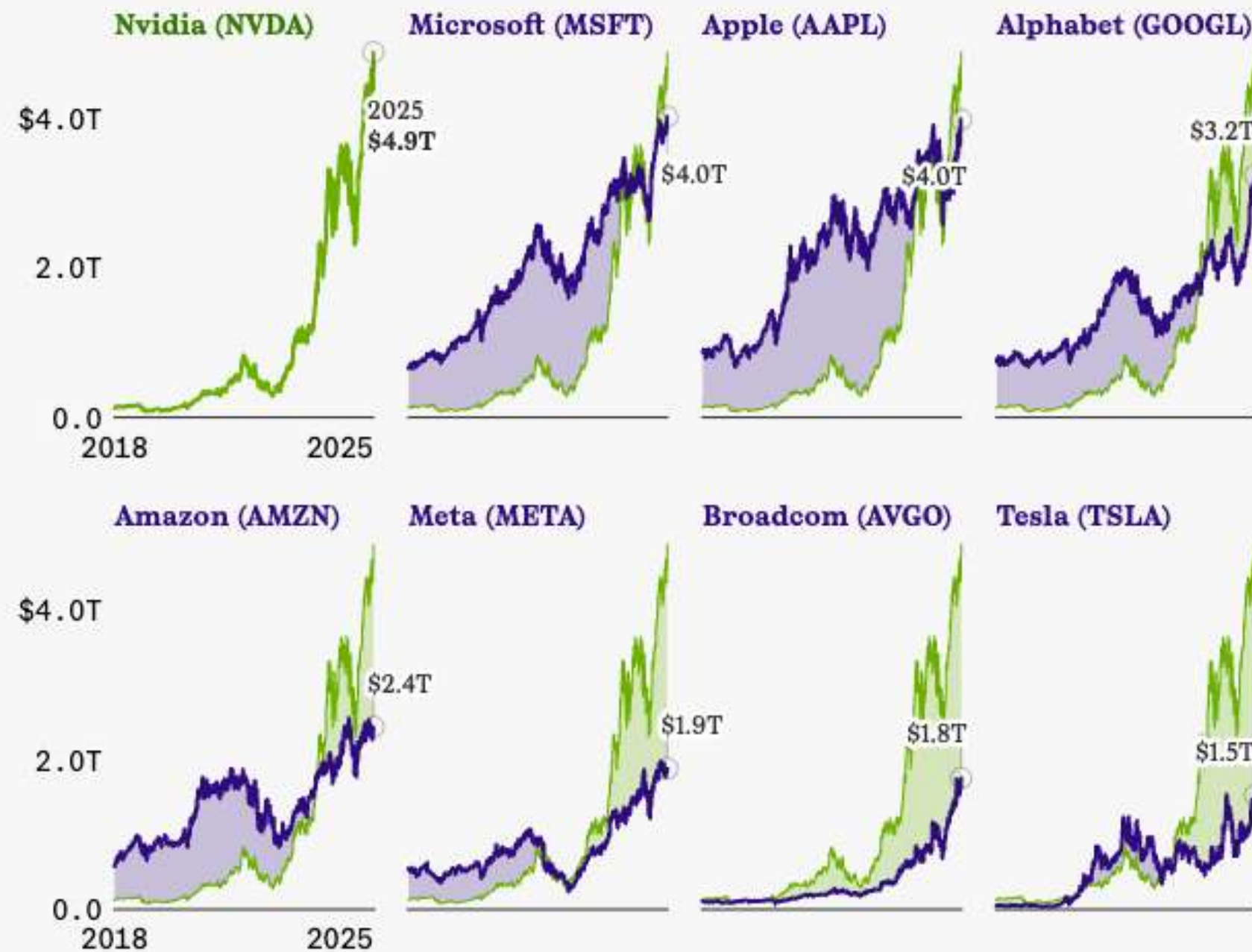


Sources: NASDAQ, Econovis

www.econovis.net @econovisuals EV

Nvidia Vs. Its BATMMAAN Peers

Market Caps, Relative to Nvidia [\$T, Daily]

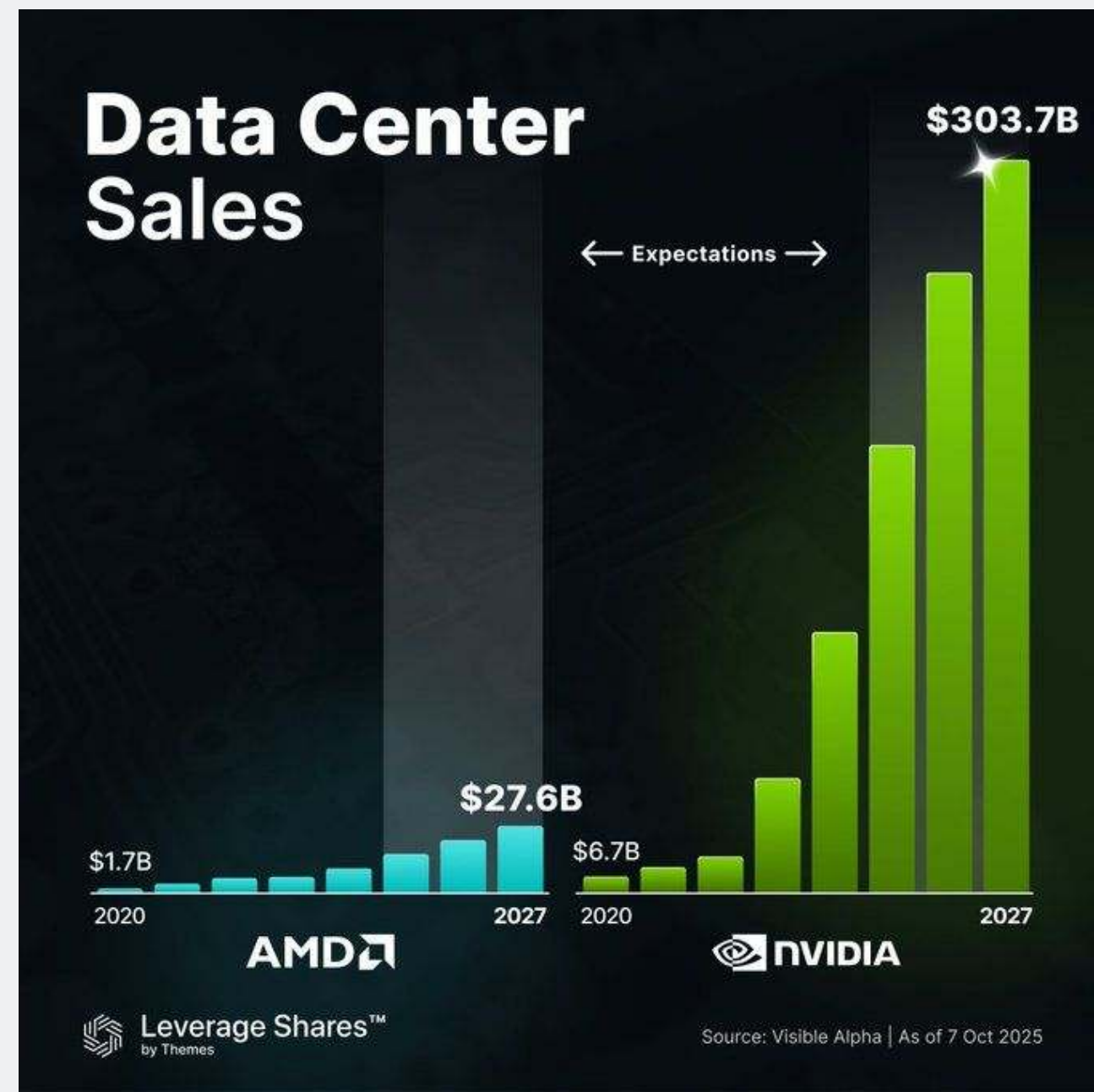
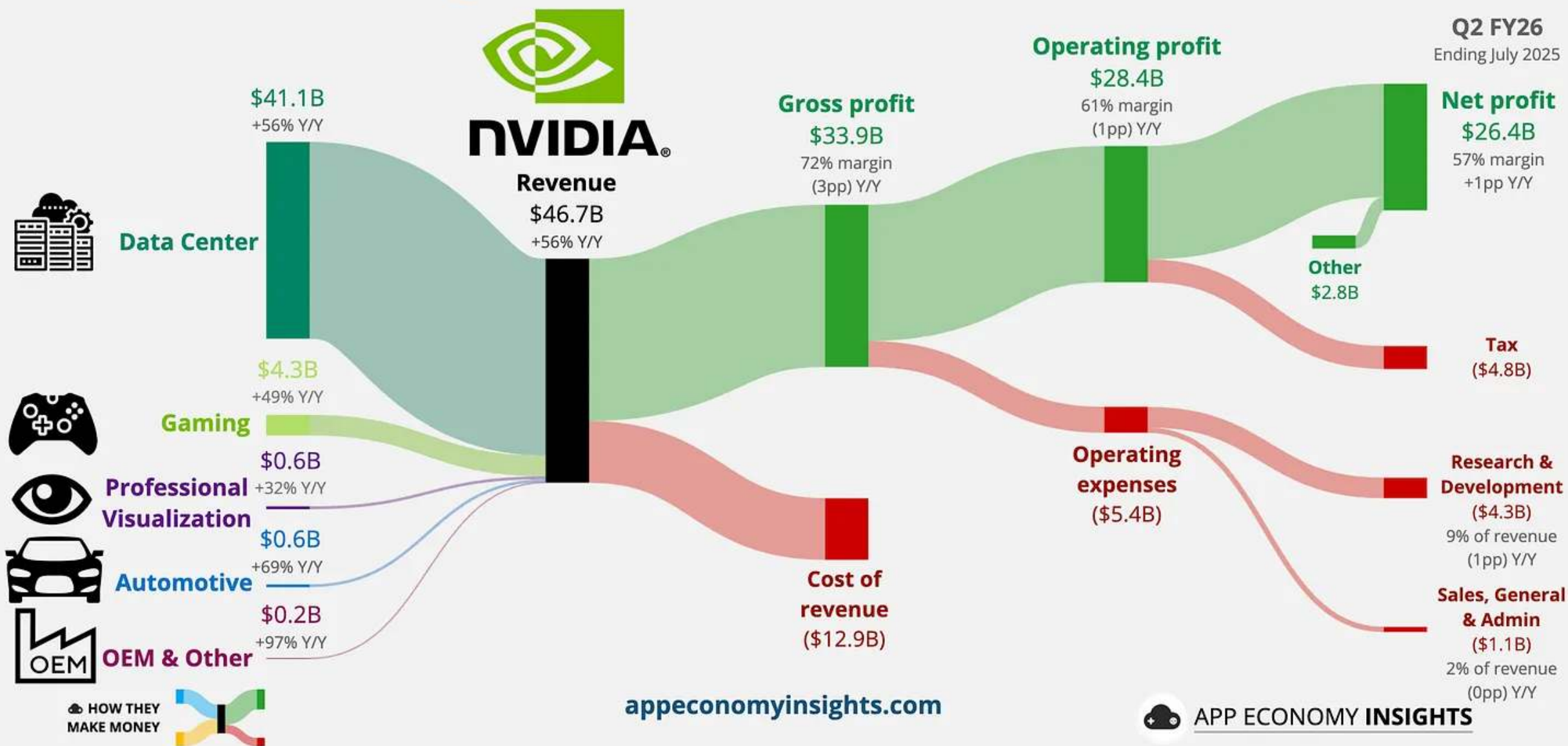


Source: YCharts

Sherwood

← → 🔍 Q Nvidia现在是什么地位？以及多能挣钱？

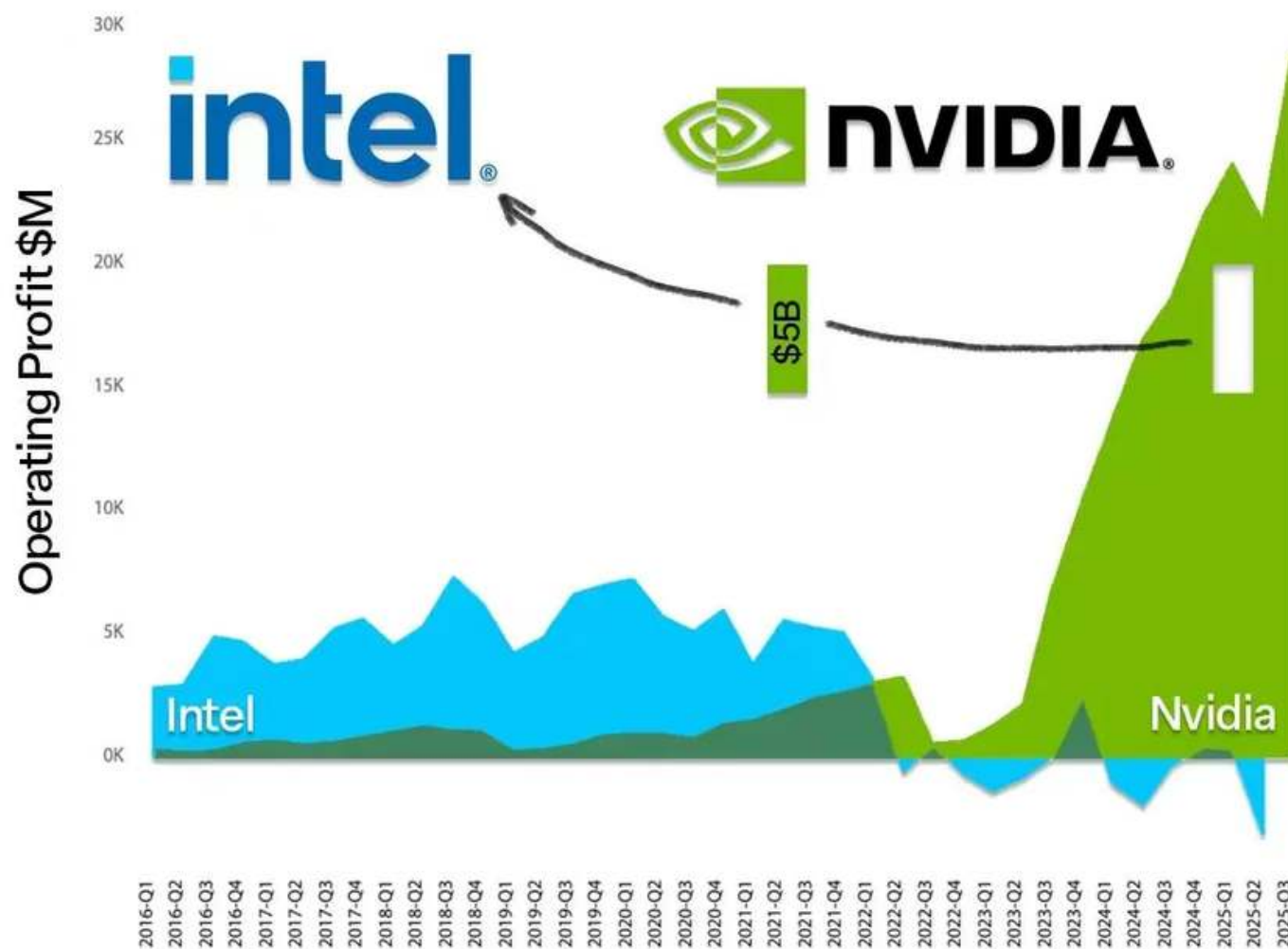
NVIDIA Q2 FY26 Income Statement



← → 🔍 钱在账上趴着干嘛？**燥起来啊！**

A Lifeline for Intel, a Footnote for Nvidia

Nvidia invests 5B\$ in Intel and Begin Collaboration



Announcing
Nokia to Build AI-Native 6G on New NVIDIA ARC
“Aerial RAN Computer”

NOKIA

NVIDIA

**NVIDIA INVESTS \$1 BILLION
IN NOKIA: POWERING THE
AI-NATIVE 6G REVOLUTION**

Announcing Department of Energy
Partners with NVIDIA
to Build 7 New AI Supercomputers

Argonne National Laboratory

Solstice - 100K Blackwell GPUs

Equinox

Tara

Minerva

Janus

Los Alamos National Laboratory

Powered by Vera Rubin

Mission

Vision

NVIDIA

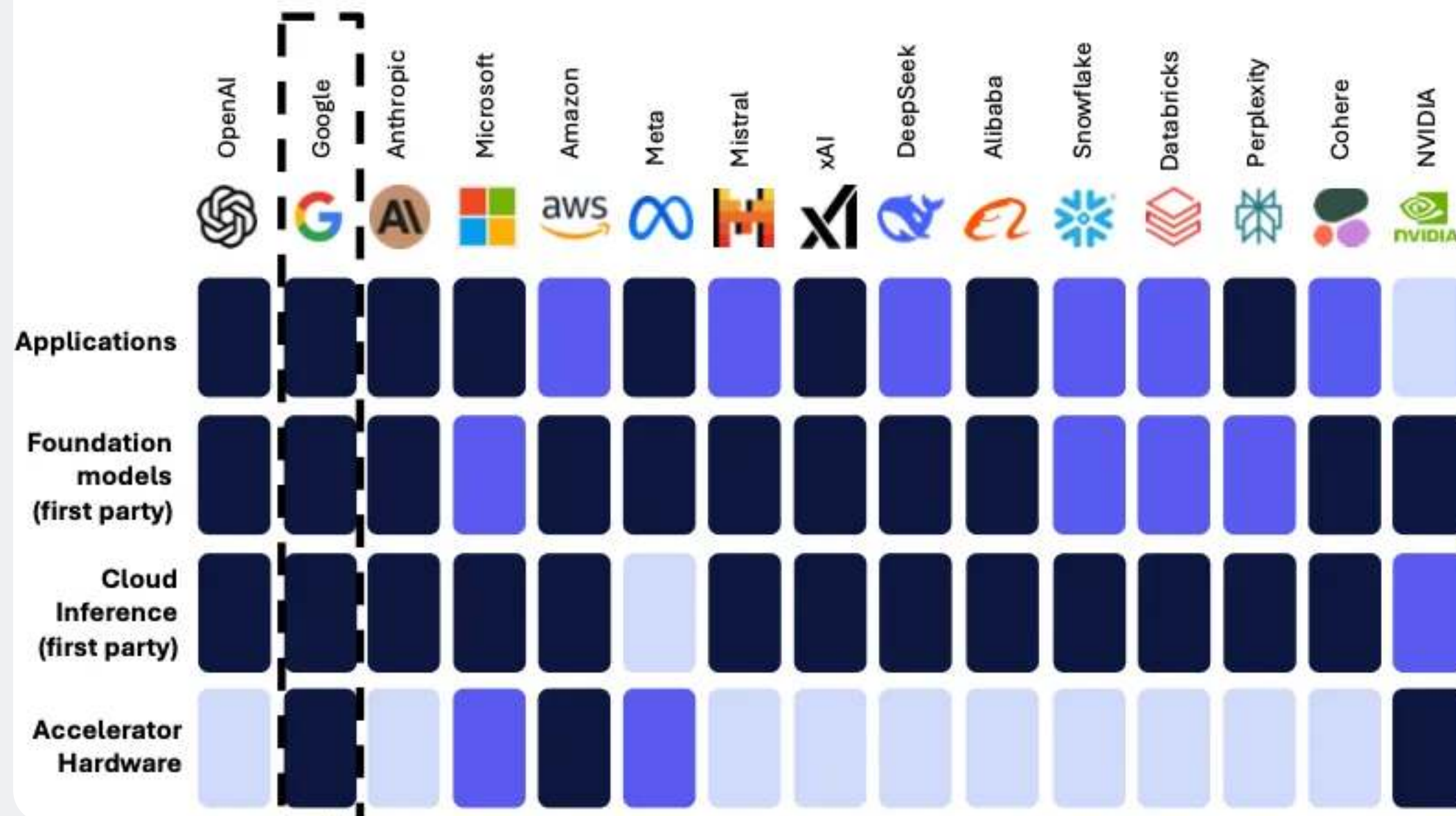
Los Alamos
NATIONAL LABORATORY

U.S. DEPARTMENT
of ENERGY

← → 🔍 🔍 这个局为什么Google敢不参与？---因为可以**再开一桌**

Key Players in the AI Value Chain

Classifications are indicative and determined based on a range of factors including market share and strength of offering



← → 🔍 我就是为了**这碗醋**包的这顿饺子.....



Sam和老黄到底在搞什么？



我们怎么一步步走到了这？



所谓“泡沫”到底是什么？



← → 🔍 我们在前面的100多页PPT里面见到了很多数字

“Feeds & speeds” and a lot of acronyms

Hundred of models, dozens of benchmarks and acronyms - MTOK replaces MHz



← → 🔍 🔍 数字单位---100 million



Cristiano Ronaldo
commanded €94M
to join Real Madrid in
2009.

Welcome to the new economy.



Jiahui Yu reportedly
received \$100M+ to
join Meta in 2025.



- 由一家科技公司/超级富豪持有
- 明星员工的年薪超过千万，甚至过亿美金
- 普通员工和明星员工薪资差距巨大

我说的是AI labs，不是NBA球队

← → 🔍

数字单位---100 million

EliseAI Joins the

“\$100M”
ARR Club

Company	ARR	Total Funding	Valuation
Elise ^{AI}	>\$100M	\$342M	\$2.2B
IIElevenLabs	>\$200M	\$281M	\$6.6B
glean	>\$103M	\$770M	\$7.2B
synthesia	>\$100M	\$335.6M	\$2.1B
Lovable	>\$100M	\$225.5M	\$1.8B
cohere	>\$100M	\$1.6B	\$7B

Rank ³	Company	Market Founded year	Annual revenue ¹ ▼	Revenue multiple ²	Revenue per employee
1	 CURSOR (Anysphere)	Software development 2022	\$500M	20x	\$3,205K
2	 replit	Software development 2016	\$150M	20x	\$758K
3	 glean	Enterprise workflow 2019	\$100M	72x	\$82K
3	Harvey	Legal 2022	\$100M	50x	\$148K
3	 Lovable	Software development 2023	\$100M	18x	\$2,222K
3	 MERCOR	HR 2023	\$100M	20x	\$1,333K
7	 crescendo	Customer service 2024	\$91M	6x	\$47K
NEW 8	 manus	Enterprise workflow 2022	\$90M ⁴	N/A	N/A
NEW 9	 Cognition	Software development 2023	\$73M	140x	\$702K
10	 StackBlitz	Software development 2017	\$40M	N/A	\$1,143K

垂类/初创明星AI公司ARR门槛

← → 🔍 数字单位---1 billion



AI unicorn formula: ½ the team, ¼ the time

Of the new unicorns in 2024...

	Non-AI unicorns	AI unicorns
Median employee headcount	414	203
Median time to unicorn status	9 years	2 years

Source: CB Insights. Data as of 12/9/24.

← → 🔍 🔍 数字单位---10 billion

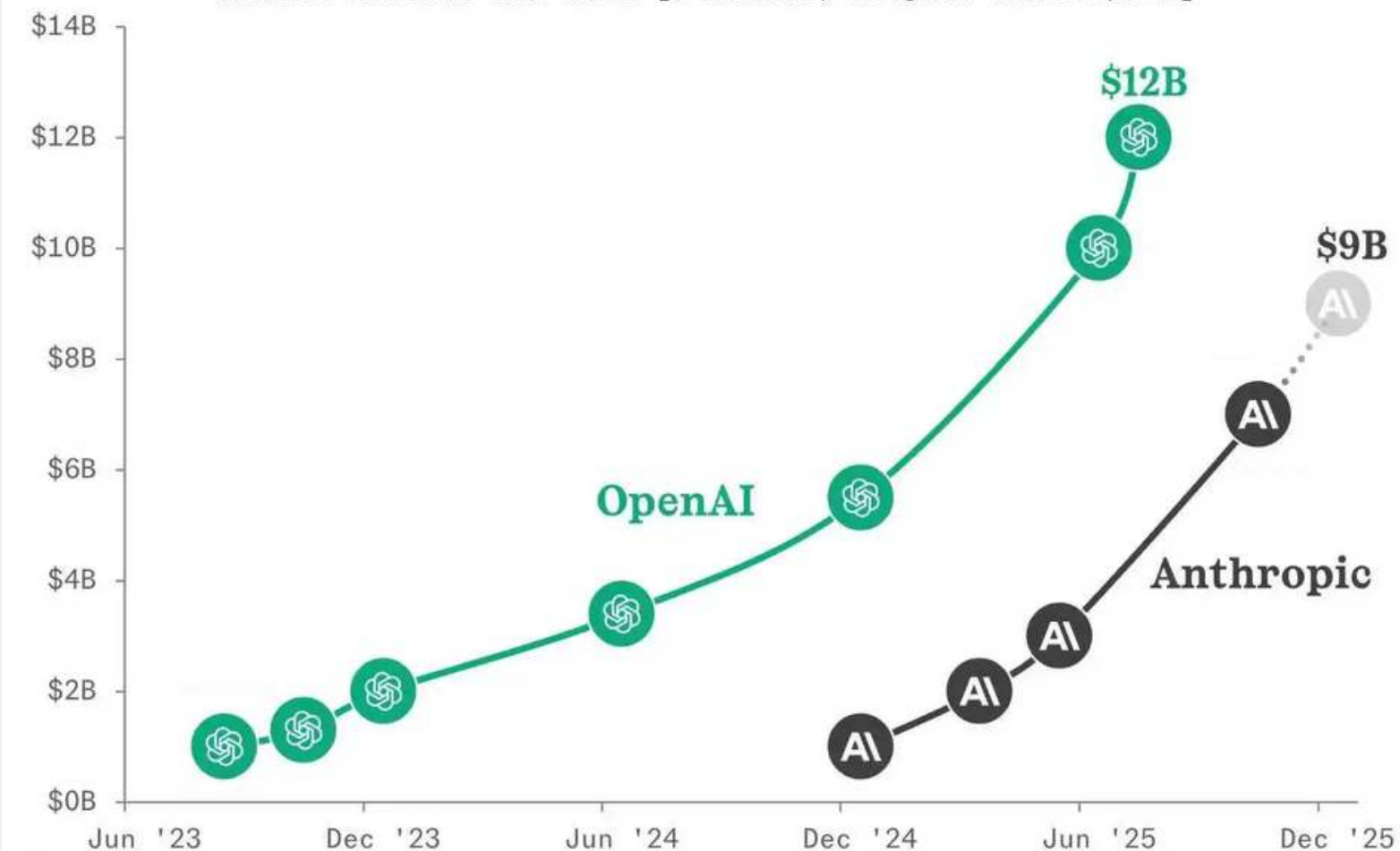
Most Valuable Private AI Startups							Oct 13, 2025
Minimum \$10B valuation							@ deedydas
#	Company	Latest Valuation (USD)	Founded	Annual Revenue	Year over Year growth	Revenue multiple	Comments
1	OpenAI	\$500B	2015	~\$12B	~100%	~41.7x	Growth based on mid-year 2025
2	xAI	\$200B	2023	~\$200M	—	~1000x	*Based on rumored latest round. Uses revenue projected from Q1
3	Anthropic	\$183B	2021	>\$5.5B	>500%	~33.2x	Growth based on mid-year 2025
4	Databricks	>\$100B	2013	~\$4B	>50%	≥25x	Not technically an AI company, but has AI products
5	Figure AI	\$39B	2022	\$0	—	∞	Robotics company, no revenue
6	Safe Superintelligence (SSI)	\$32B	2024	\$0	—	∞	Ilya's company, no revenue
7	Cursor (Anysphere)	\$30B*	2022	\$500M+	>500%	~60x	*Based on rumored latest round, lower margin revenue
8	Scale AI	\$29B	2016	~\$870M	~130%	~33x	Based on 2024 numbers
9	Perplexity	\$20B	2022	~\$148M	~135%	~135x	Last reported public metrics
10	Mistral AI	~\$14B	2023	> \$100M	—	~140x	Last reported public metrics
11	Thinking Machines Lab	\$12B	2025	\$0	—	∞	Mira's company, no revenue
12	Cognition	\$10.2B	2023	~\$150M	~7200%	~68x	Revenue combines Cognition and Windsurf revenue
13	Sierra	\$10B	2023	\$100M+	>400%	~100x	Fully enterprise revenue
14	Mercor	\$10B	2023	~\$450M	—	~22x	Lower margin revenue
15	Midjourney	\$10B*	2022	~\$500M	~66%	~20x	*Implied valuation based on analysts, hasn't raised outside capital.

敢被称为AI明星公司的估值门槛

← → 🔍 数字单位---10 billion

Could Anthropic Catch OpenAI In The Biggest Revenue Race In Artificial Intelligence?

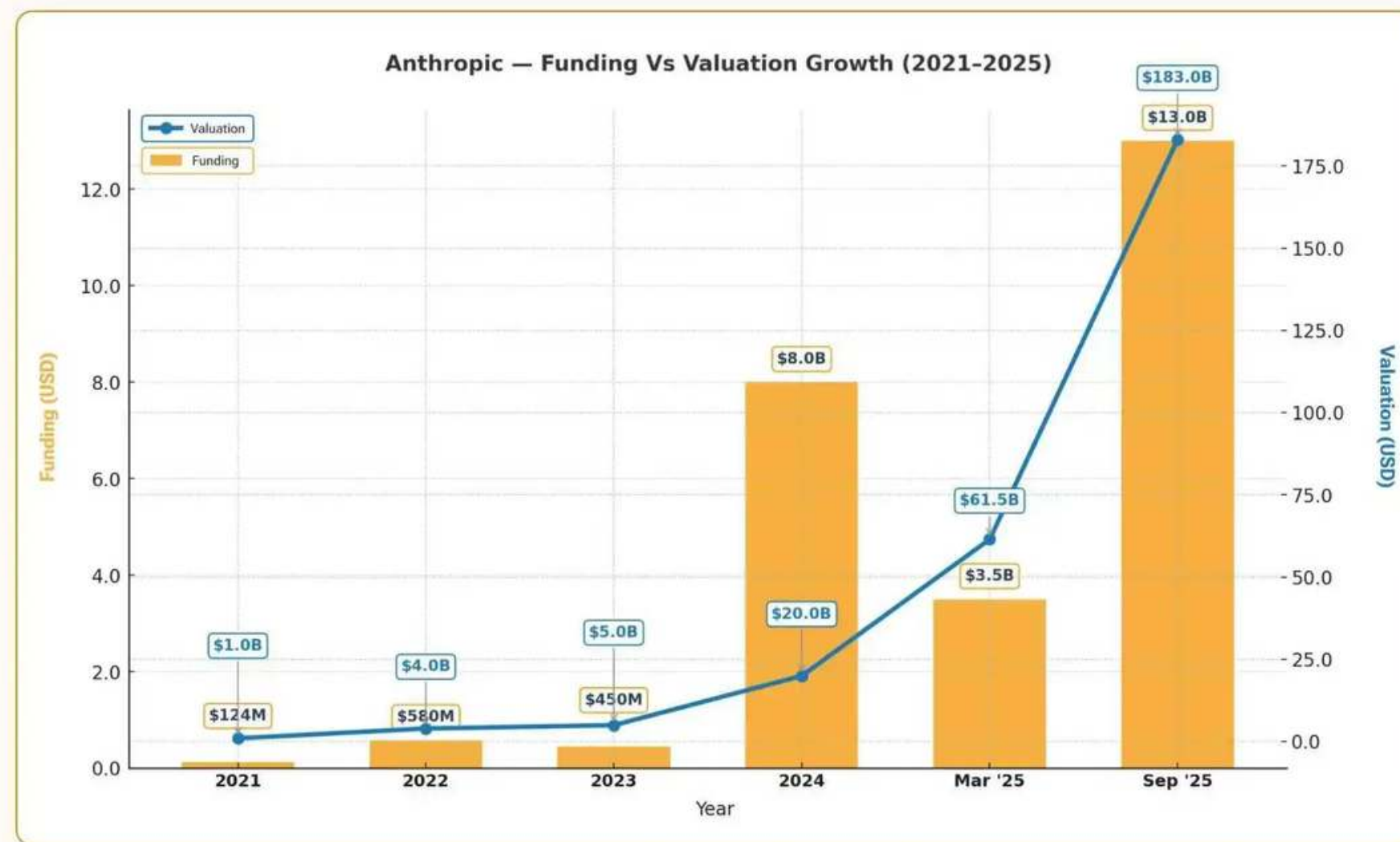
Annual Revenue Run Rate [\$ billions, irregular media reports]



CHARTR

Source: Media reports | Note: December 2025 is Anthropic's target ARR

ANTHROPIC funding & valuation



最头部模型厂商的2025年ARR以及单轮融资规模

← → 🔍 🔍 数字单位---10 billion

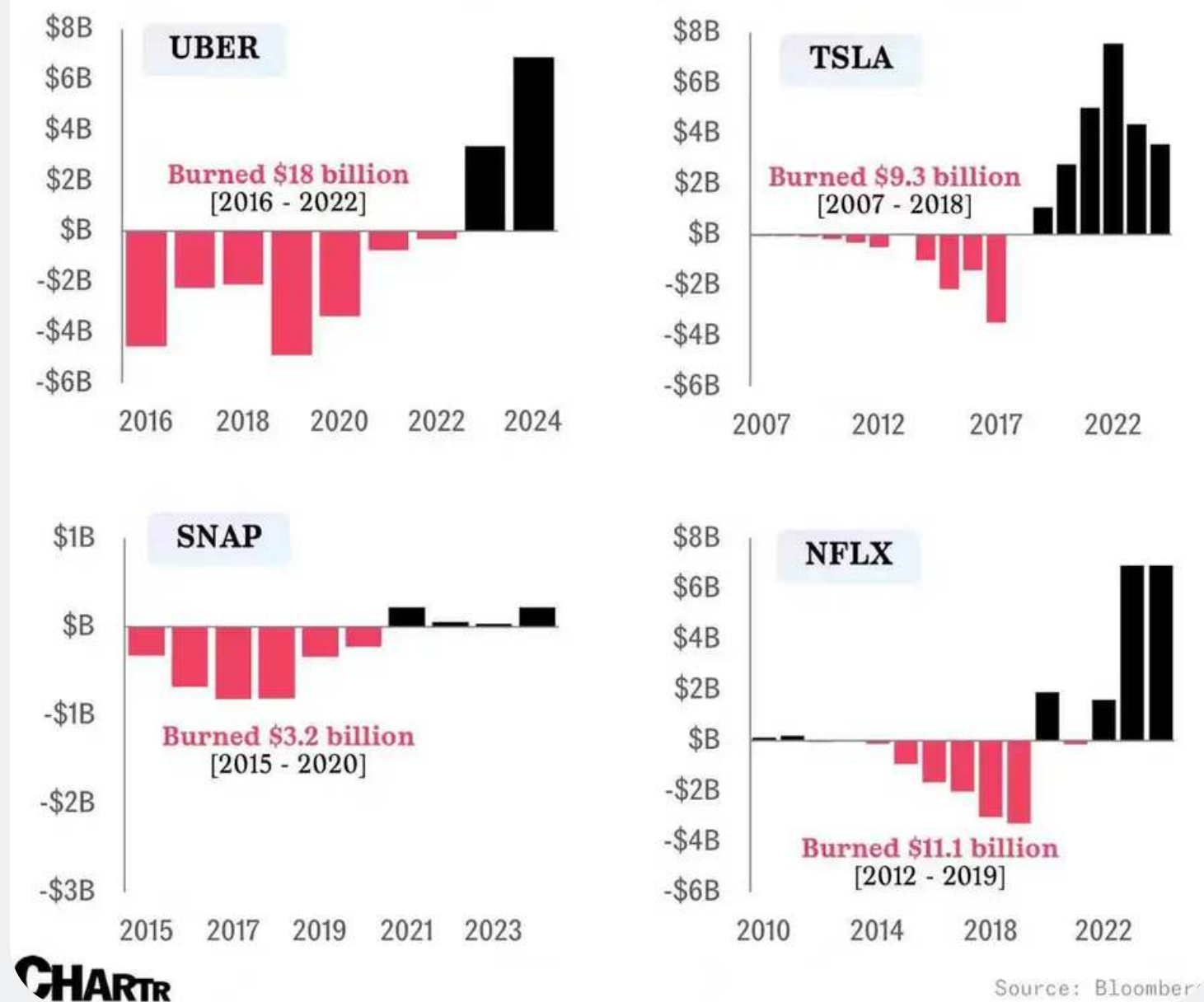


Neocloud厂商的“大订单”金额

← → 🔍 数字单位---10 billion

Cash Incinerators: Tech Companies Often Burn Billions Before They Earn Billions

Free Cash Flow, Selected Companies [Annual, \$ billions]



历史上最烧钱的公司所烧的钱

上市之前烧钱总额

amazon

≈

\$1B \$1B

9x

Uber

≈

\$1B \$1B \$1B \$1B \$1B \$1B
\$1B \$1B \$1B \$1B \$1B \$1B
\$1B \$1B \$1B \$1B \$1B \$1B

9x

OpenAI

≈

\$162b



← → 🔍 🔍 数字单位---小几十个billion



史上最大收购案之一



英伟达上个季度的收入



Baidu的市值

← → 🔍 🔍 OpenAI发明了一个新的单位---“GW”---约等于\$50b

A 1GW AI data center cheat sheet

1GW GPU cluster: \$50bn capex, \$8.6bn D&A p.a.

1GW NVL 72 GB200 capex/D&A	Capex	Useful life	D&A p.a.
Total per 1GW of capacity	\$50bn		\$8.6bn
Compute and storage	\$30bn		\$6.0bn
% of total	60%		70%
GPUs	\$21bn	5 years	\$4.2bn
% of total	42%		50%
ASP	\$35k		
Chips per GW	600k		
Average power/GPU	1.7kW		
CPUs	\$1bn	5 years	\$0.2bn
% of total	2%		2%
Other server and storage	\$8bn	5 years	\$1.6bn
% of total	16%		18%
Networking	\$6bn	5 years	\$1.2bn
% of total	12%		14%
Building, power & cooling	\$14bn	10 years	\$1.4bn
% of total	28%		16%

OpenAI | NVIDIA

10GW

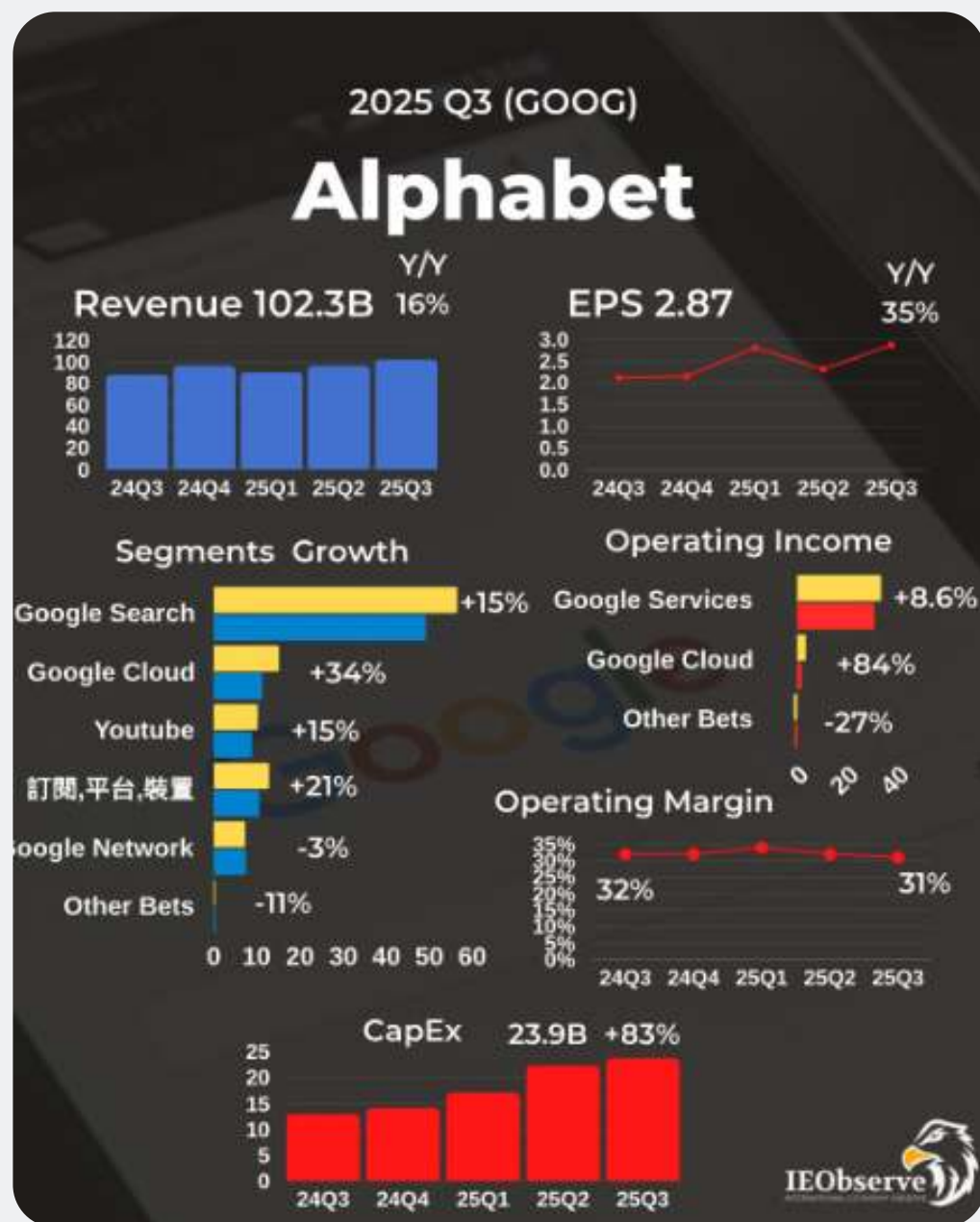
OpenAI | BROADCOM

10GW

OpenAI | AMD

6GW

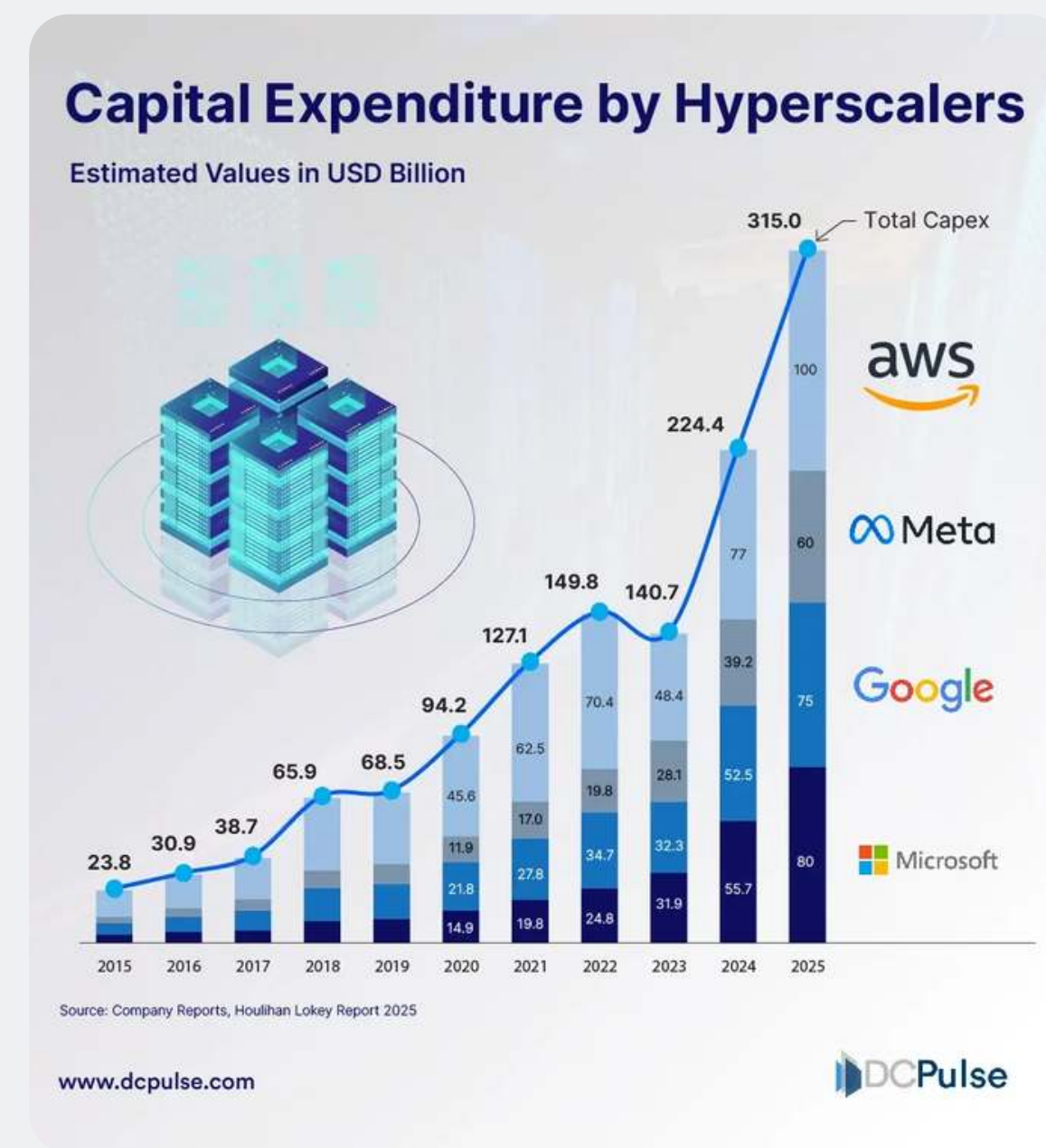
← → 🔍 数字单位---100 billion



Google最新一季收入



英伟达对OpenAI的投资额



巨头们的年度Capex

← → 🔍 🔍 数字单位---500 billion

OPENAI AT \$500B?

INSIDE THE \$6B EMPLOYEE STOCK SALE

TL:DR: It's a secondary, not an IPO

OpenAI当前的估值

Exceptionally Strong Demand for Grace Blackwell NVL72

Driven by Two Exponentials

Visibility into more than half a trillion dollars in
cumulative Blackwell and Rubin revenue
through CY2026 (so far)

More than 5X Hopper's lifetime revenue

\$100B
4M GPUs

2023-25
Hopper
Lifetime

B Rubin
(Far)

Shipped
6M GPUs

Blackwell and Rubin are 2 GPUs per chip.
Excluding China.

\$0.5T+
20M GPUs

英伟达芯片订单额

← → 🔍 数字单位---Trillion

AI's \$200B Question



GPU capacity is getting overbuilt. Long-term, this is good. Short-term, things could get messy.

BY DAVID CAHN
PUBLISHED SEPTEMBER 20, 2023

AI's \$600B Question



The AI bubble is reaching a tipping point. Navigating what comes next will be essential.

BY DAVID CAHN
PUBLISHED JUNE 20, 2024



THE VENTURE ALLOCATOR

The \$1 Trillion Question

Without The AI Buildout, The U.S. Economy Might Already Be In Recess

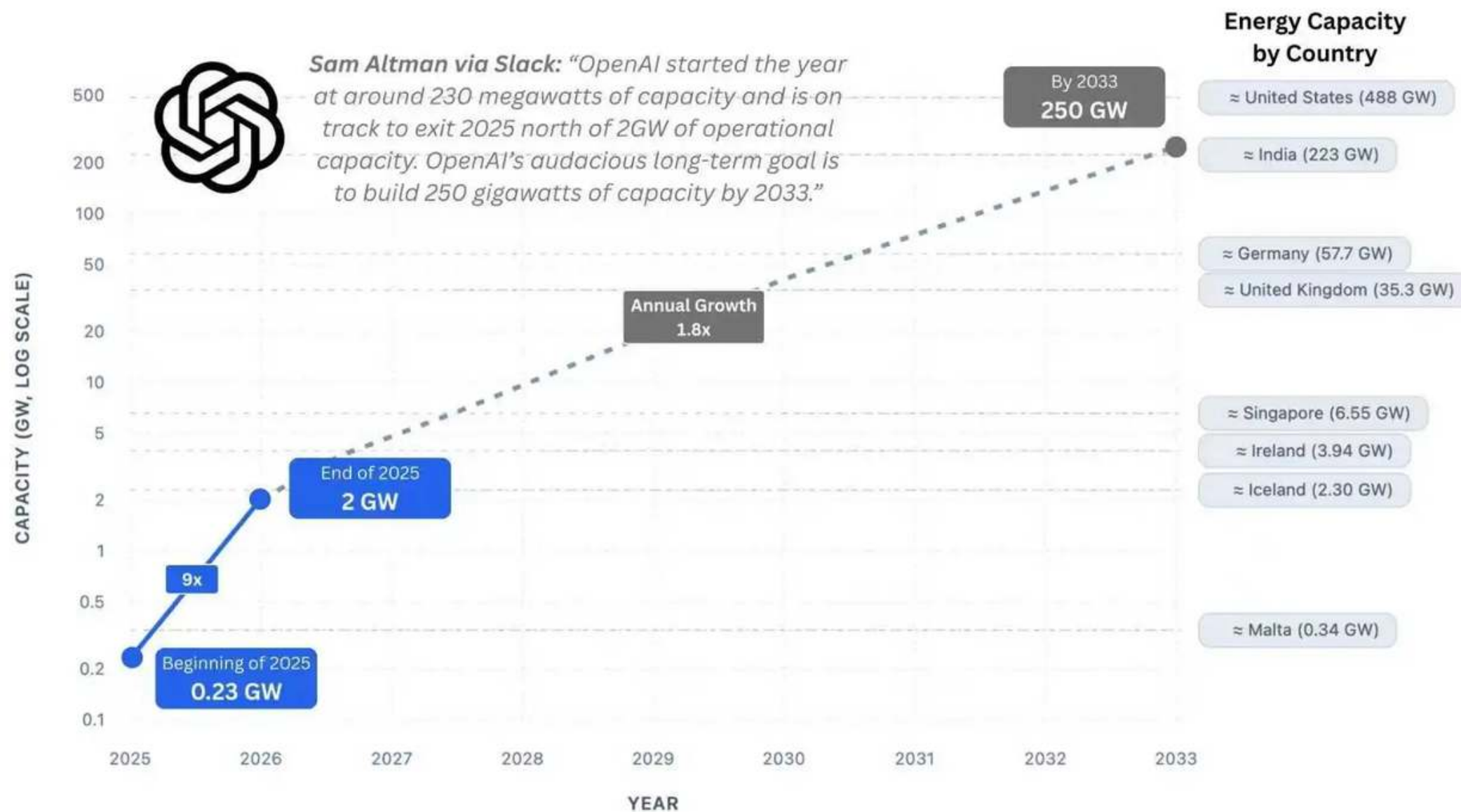


JAMES HEATH
OCT 08, 2025

← → 🔍 数字单位---Trillion

OpenAI planning to 125x energy capacity in 8 years

This would mean using more than India's energy capacity today



OPEN AI PREPARES
FOR

IPO
\$1 TRILLION
VALUATION

 **OpenAI**

← → ↺ 🔍 数字单位---Trillion

MICROSOFT AND APPLE JOIN THE \$4 TRILLION CLUB

4.5 Market Capitalization (Trillion US\$)



Sources: NASDAQ, Econovis

www.econovis.net @econovisuals EV

ALPHABET SURGES PAST AMAZON TO BECOME THE WORLD'S FOURTH \$3 TRILLION COMPANY

Market Capitalization (Billion US\$) 3,500



Sources: NASDAQ, Econovis

www.econovis.net @econovisuals EV

← → 🔍

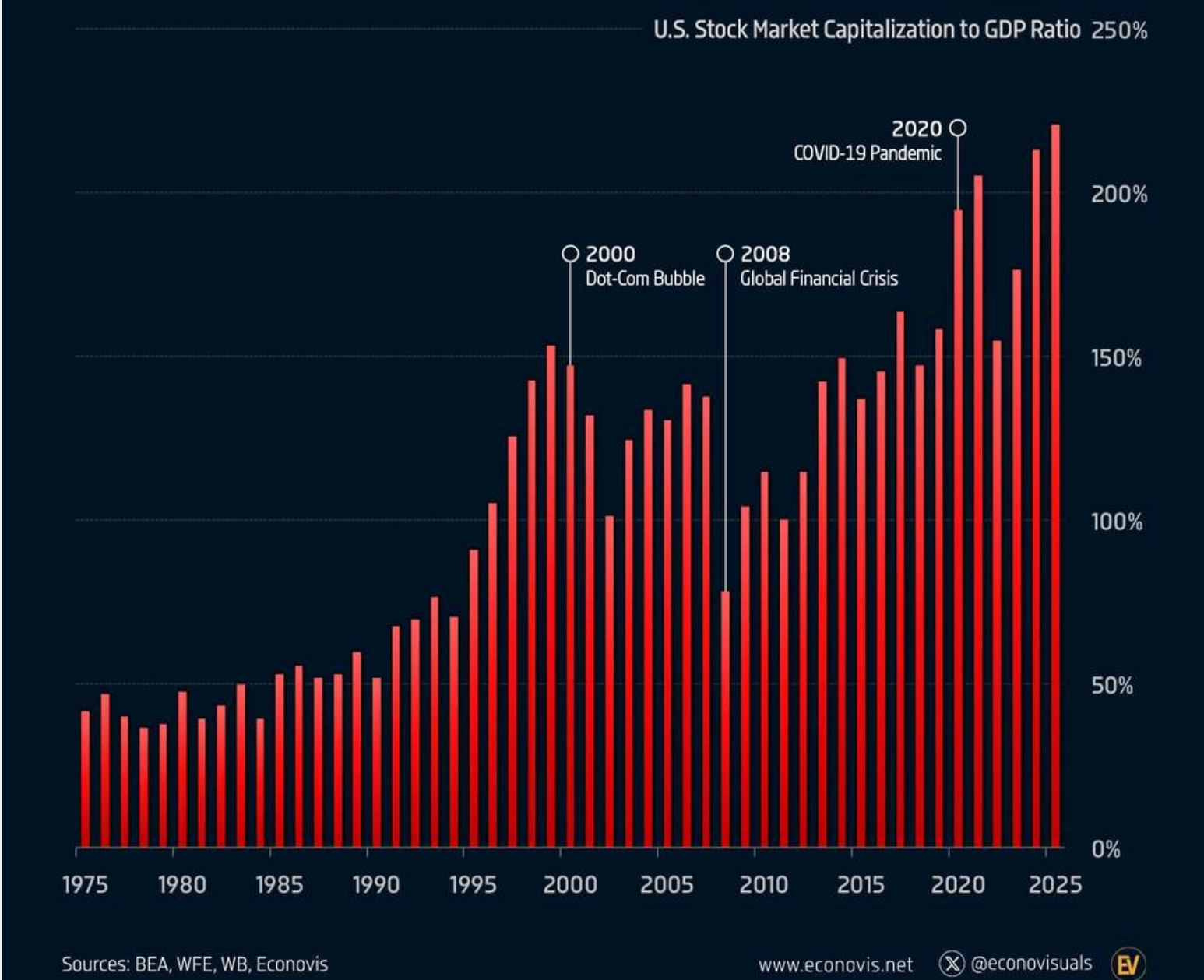
数字单位---Trillion

The BATMMAAN Stocks Tale Of The Tape

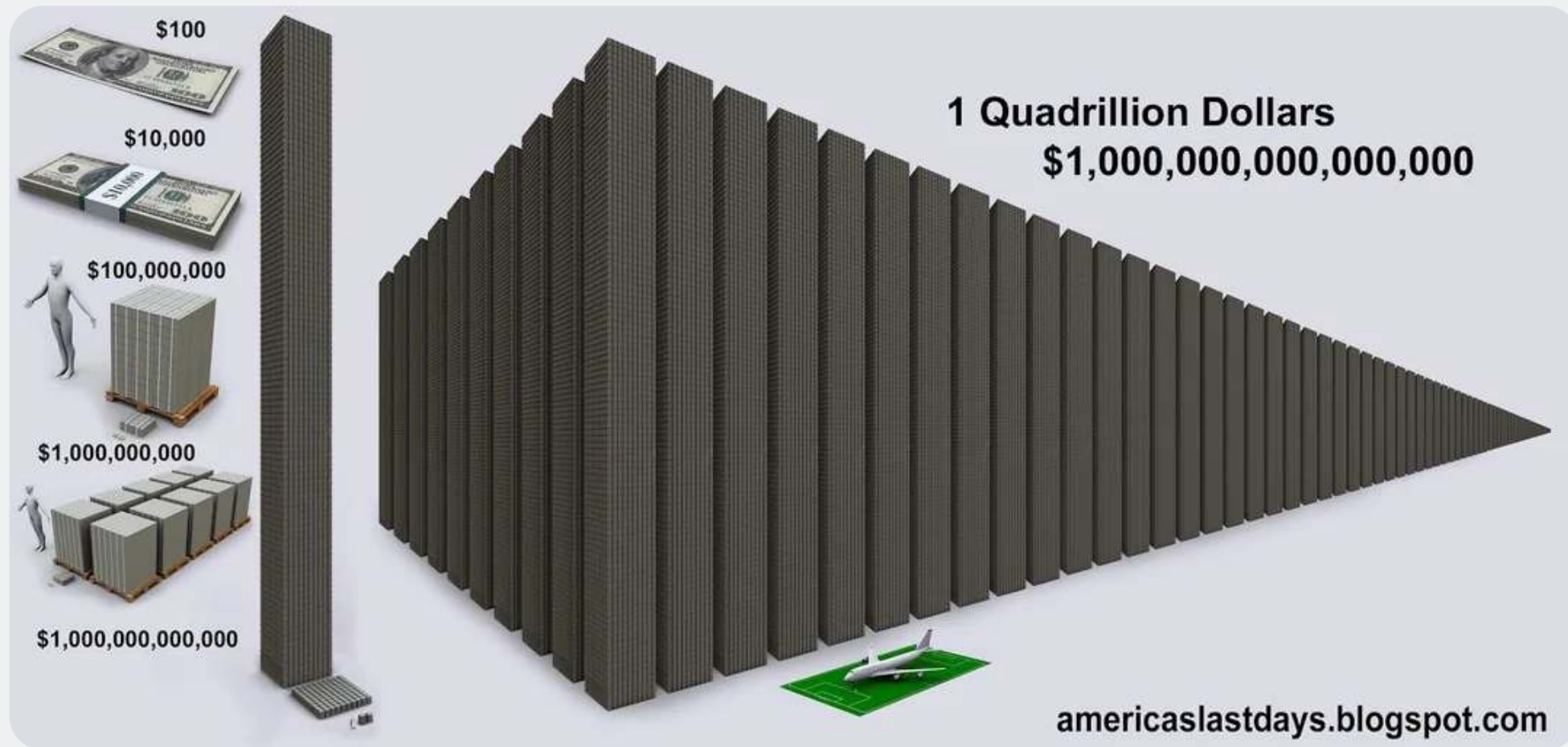
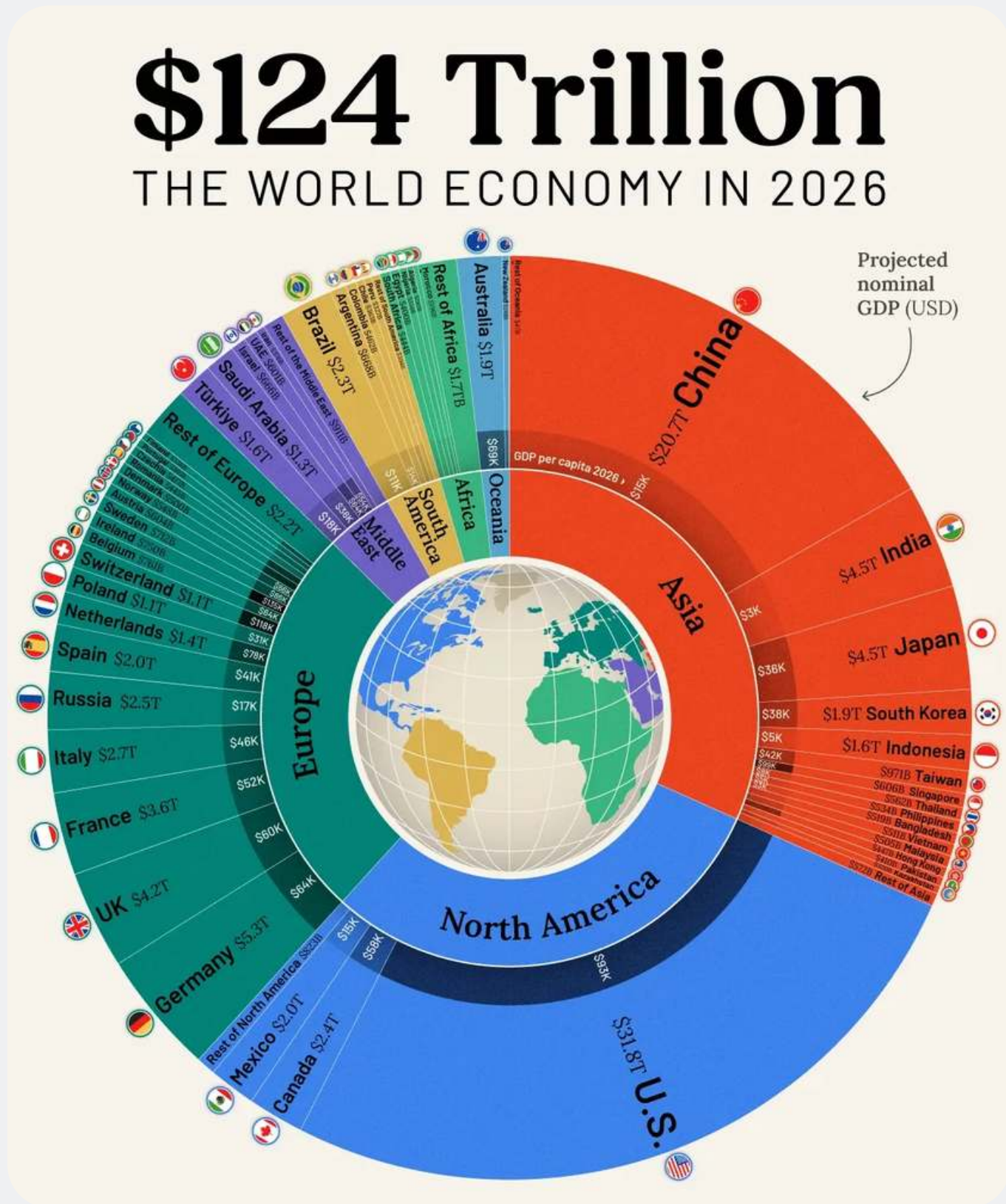
[As of October 29, 10:50am ET]

Letter	Name	Ticker	Mkt Cap [\$T]	Industry	% YTD Return
B A T M M A A N	Broadcom	AVGO US	1.8	Semiconductors	64%
	Apple	AAPL US	4.0	Telecoms Equipment	7%
	Tesla	TSLA US	1.5	Motor Vehicles	13%
	Microsoft	MSFT US	4.0	Packaged Software	28%
	Meta	META US	1.9	Internet Software/Services	28%
	Amazon	AMZN US	2.5	Internet Retail	5%
	Alphabet	GOOGL US	3.3	Internet Software/Services	43%
	Nvidia	NVDA US	5.1	Semiconductors	55%

U.S. STOCK MARKET CAP-TO-GDP RATIO CLIMBS TO RECORD 221% IN 2025



← → 🔍 你们知道**Trillion**一般什么时候用么？以及Trillion之后是什么吗？



1. 主要使用场景（非常有限）

quadrillion 因代表的数值极大 (10^{15})，日常场景几乎用不到，仅在以下领域偶尔出现：

- **宏观经济与金融**：描述全球级别的债务、GDP 总和或跨国公司的市值。例如，全球政府债务总额曾接近 100 万亿美元 (10^{14})，尚未普遍达到 quadrillion 级别。
- **科学与工程**：计算天体质量、宇宙距离或微观粒子数量。例如，地球的质量约为 5.97×10^{24} 千克（接近长尺度的 quadrillion），但科学界更习惯用科学计数法而非文字单位。
- **数据存储与运算**：描述未来超大规模的数据量，如全球数据总量预计在多年后可能达到 quadrillion 字节（即 PB 级以上），但当前主流仍以 TB、PB 为主。

← → ↺ 🔍 还是英国人敢说啊.....



← → 🔍 都聊到这了，你觉得讨论是否是泡沫还有意义吗？



Sam和老黄到底在搞什么？



我们怎么一步步走到了这？



所谓“泡沫”到底是什么？



← → 🔍 过去的一个月，我的Substack收藏夹几乎全都是这个话题



SHILPA REDDY

10月27日

The AI Bubble: Risk or Opportunity?

A case for cautious optimism

SHILPA REDDY • 13 MIN READ

Valuation Metrics of AI-Centric Companies (Oct 2025)			
Company	Market Cap	Forward P/E	F
Advanced Micro Devices (AMD)	\$380B	28.6x	
Alphabet (GOOG/GOOGL)	\$3.1T	23.3x	
Amazon (AMZN)	\$2.3T	28.6x	
Airbus (AVGO)	\$1.6T	36.8x	
Meta (META)	\$1.8T	24.6x	
Microsoft (MSFT)	\$3.9T	33.3x	
Nvidia (NVDA)	\$4.4T	29.9x	
Oracle (ORCL)	\$802B	42.2x	



THE AI OPPORTUNITY

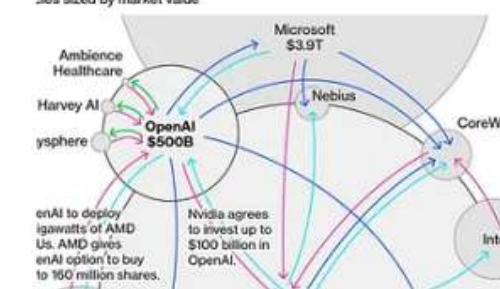
10月15日

Sh*t, are we in a bubble? - 10 Things You Need to Understand About the OpenAI Bubble

Had to read about this and it's not looking good. OpenAI circular financing is looking like a bubble about to burst

PAID • GUILLERMO FLOR • 8 MIN READ

How Nvidia and OpenAI Fuel the AI Money Machine
Hardware or Software / Investment / Services / Venture Capital
Firms sized by market value



ARTIFICIAL INTELLIGENCE MADE SIMPLE

10月17日

The "AI is a Bubble" Narrative is Stupid, Wrong, and Dangerous

How the memes and surface level analysis are distracting from Nvidia's play to control the ecosystem.

DEVANSH • 24 MIN READ



INTERESTING ENGINEERING++

10月13日

ALL IN: AI's House of Cards?

A Speculative Analysis of the \$500 Billion ~ \$1 Trillion Circular Investing/Financing Ecosystem

INTERESTING ENGINEERING ++ • 1 HR 8 MIN READ



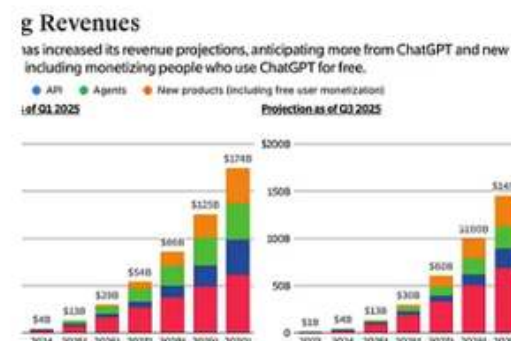
ENTERPRISE AI TRENDS

10月20日

AI bubbles are not evenly distributed

Thoughts on the most hated rally of all time

PAID • JOHN HWANG • 10 MIN READ



N. BARATTE - TECHSTOCK01

10月13日

What are the arguments in favor and against an AI bubble?

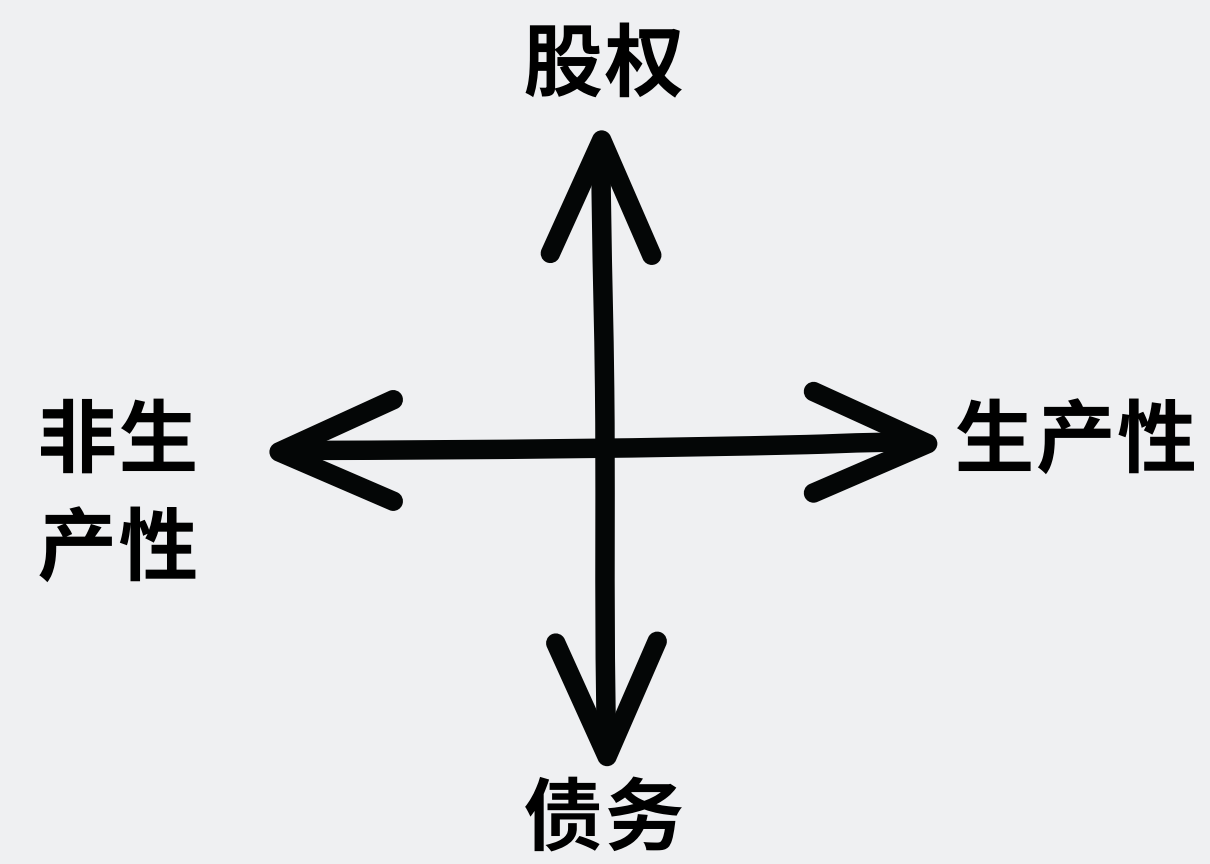
AI today is like Cloud Computing in 2014

NICOLAS BARATTE - TECHSTOCK01 • 8 MIN READ

US Hyperscalers Cloud revenues US\$ bn



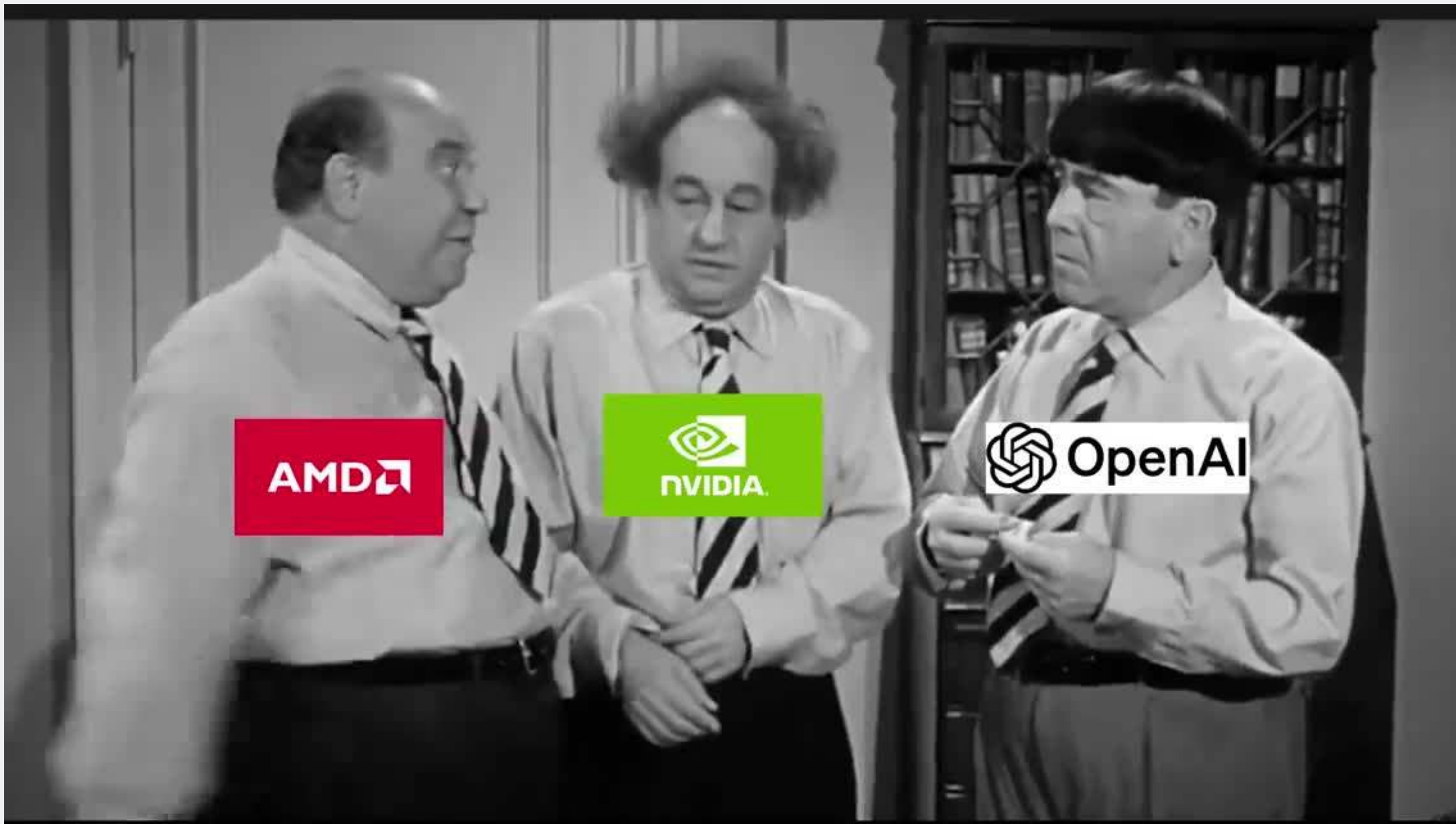
← → 🔍 已经不是泡沫的确认问题，大家开始讨论是**什么类型的泡沫**



COATUE
identified **6** distinct bubble types...

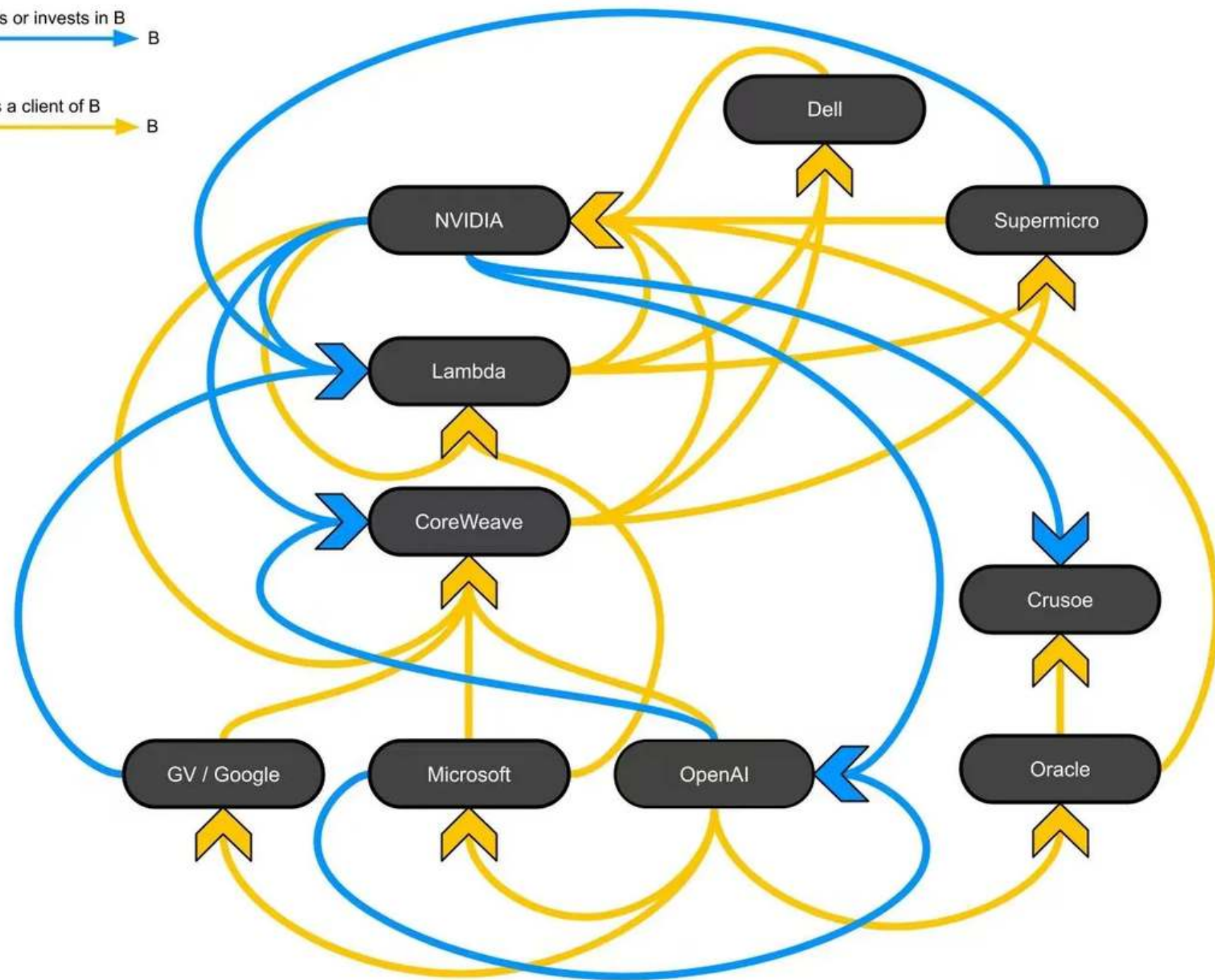
- Stock Market
- Infrastructure
- Real Estate
- Credit / Debt
- Commodity
- Collectible / Alternative Asset

← → 🔍 🔍 引发这一轮讨论的最直接因素.....循环融资

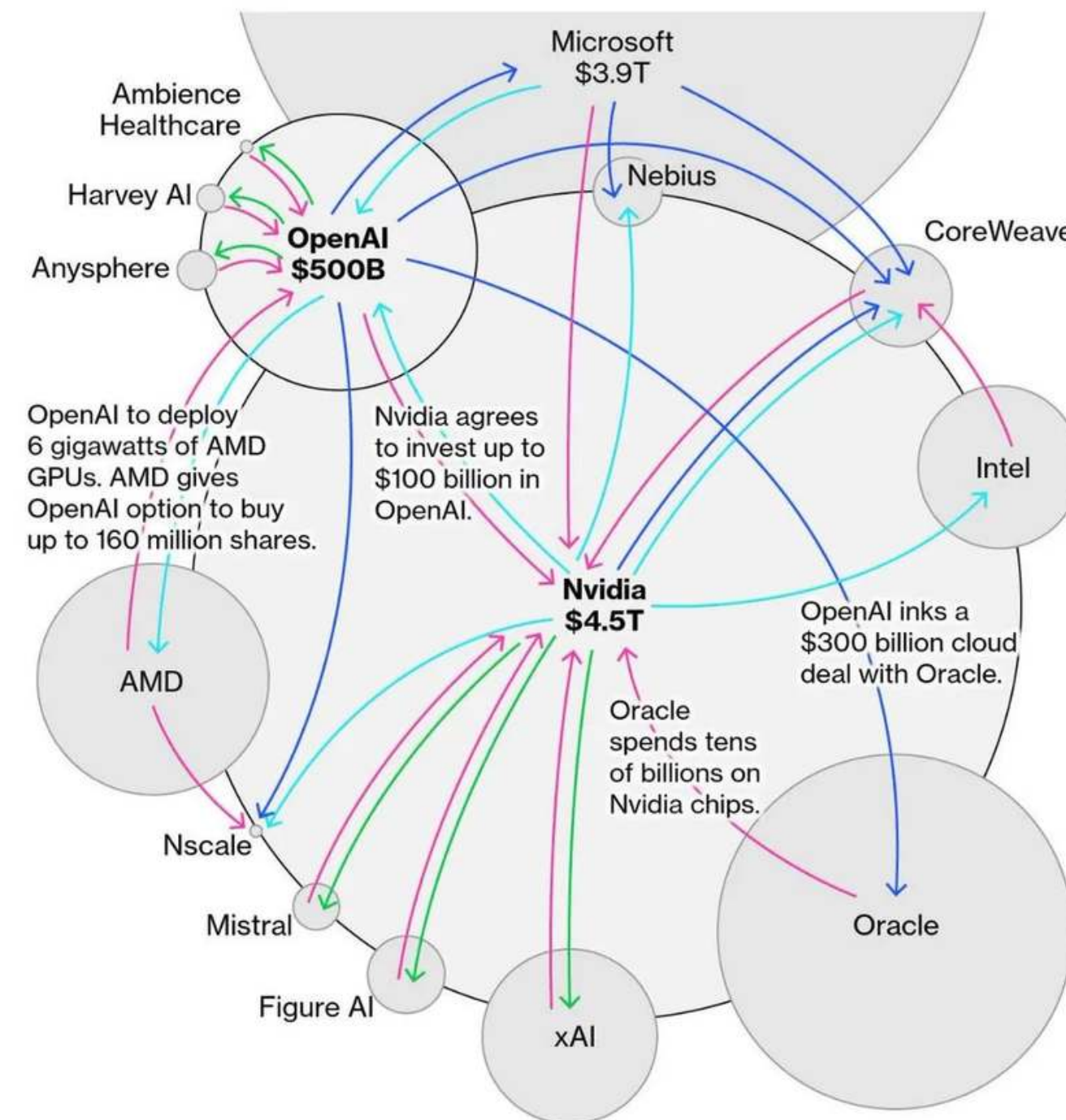


← → 🔍 这样的图你们最近一段时间应该看过很多很多了

A funds or invests in B
A is a client of B

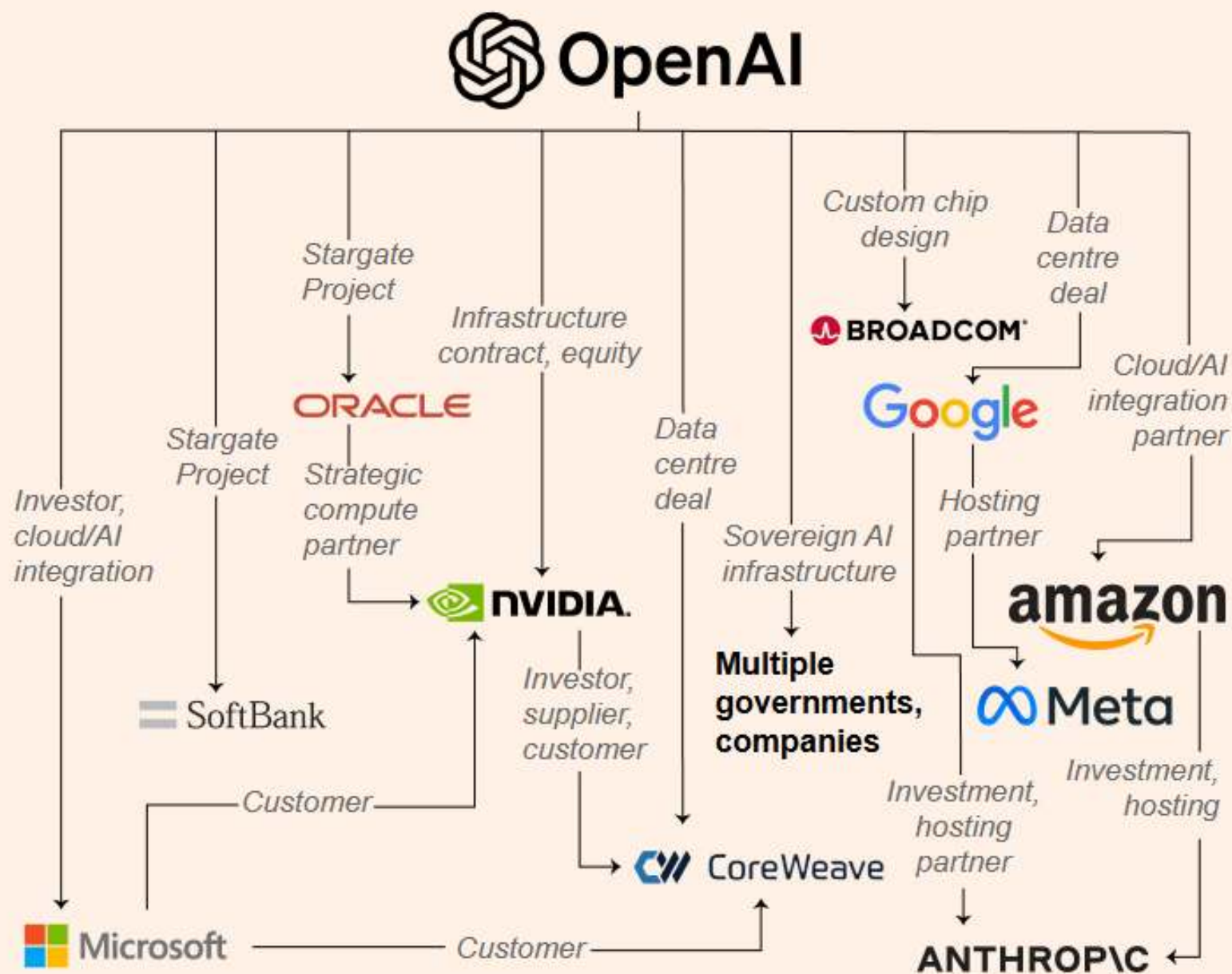


Hardware or Software Investment Services Venture Capital
Circles sized by market value



← → 🔍 这样.....

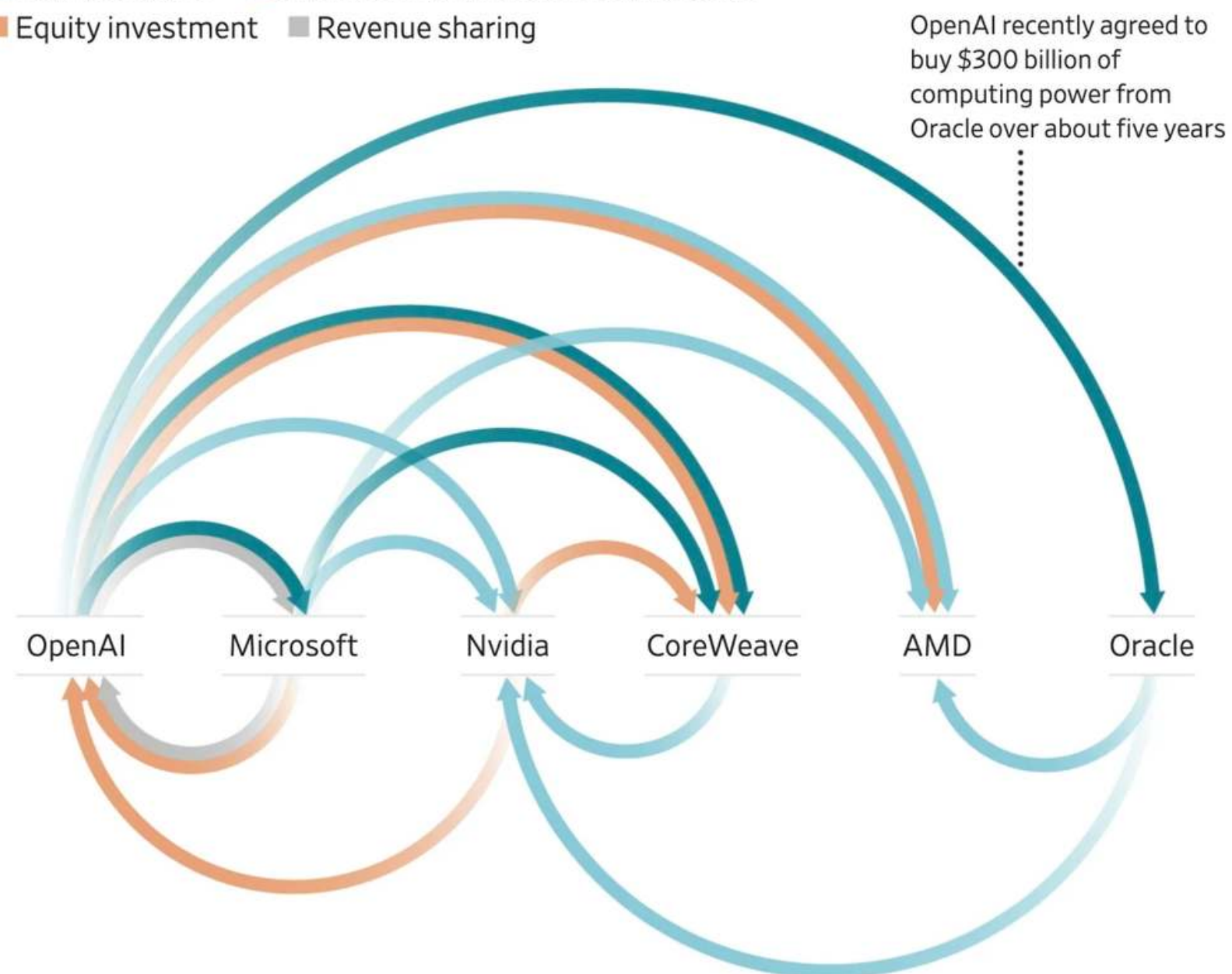
Large tech companies are tying their fortunes to OpenAI

Source: Citi Research, FT research
FT graphic: Eva Xiao

FINANCIAL TIMES

Select capital flows among six AI-industry companies

■ Chip purchases ■ Infrastructure purchases and rentals
■ Equity investment ■ Revenue sharing



Note: Some chip purchases are through intermediaries. Some investments and other arrangements subject to conditions.

Sources: staff reports; Morgan Stanley
Nate Rattner/WSJ

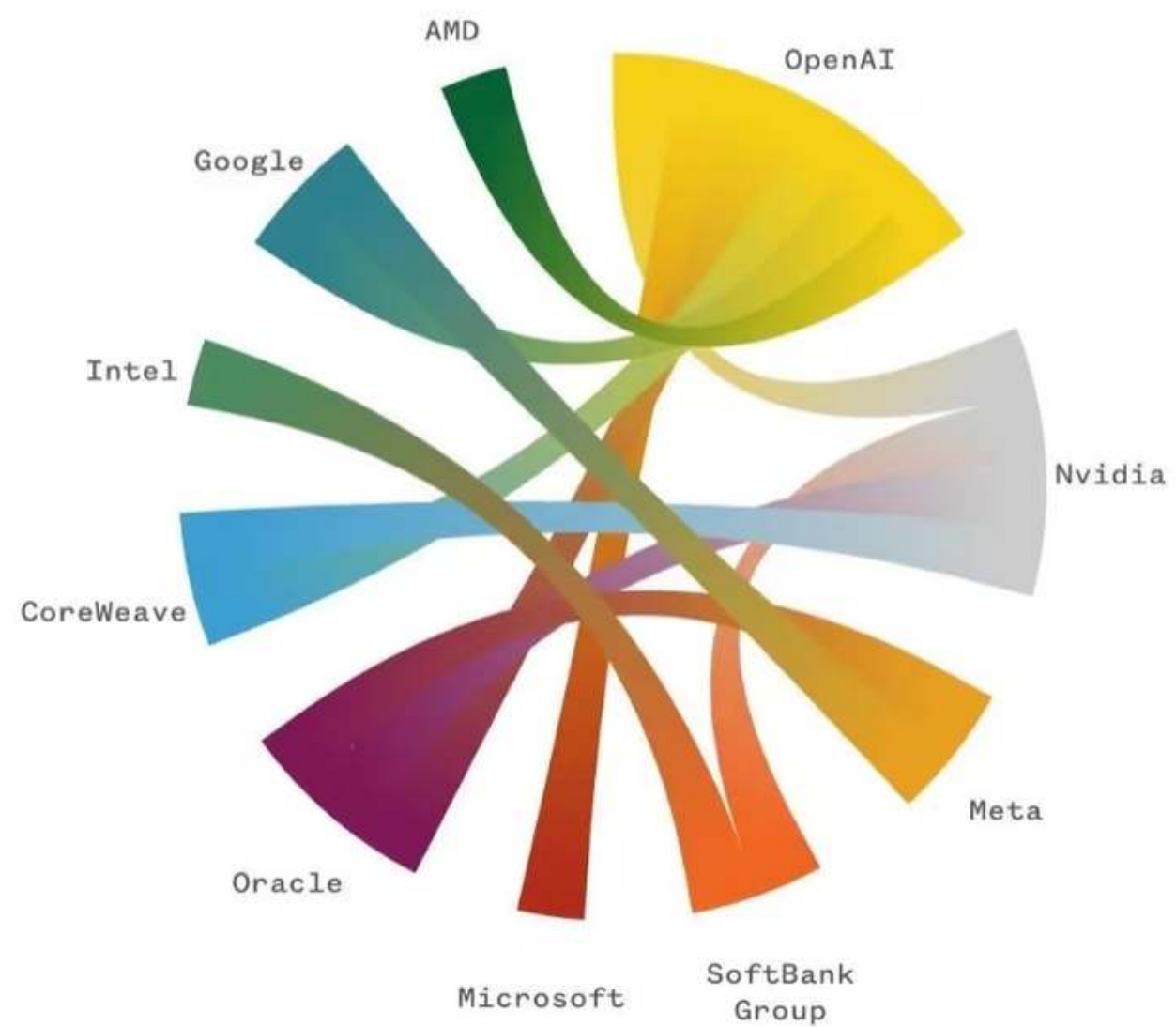
00175

小红书 6721048205

← → 🔍 甚至逐渐Meme化...

AI funding's circular web

The Nvidia-OpenAI partnership is the latest in a series of investments as tech companies pour billions into data centers and compute power.



← → 🔍 🔍 所有人不可避免的和.com泡沫做对比

The Rise and Fall

The Nasdaq Composite Index, daily close



S&P 500: Then and Now



Source: SoFi, Bloomberg

← → 🔍

我本想着用AI来做这次的ppt，后来只觉得这一页能用

历史先例与未来展望

🕒 技术资本支出超级周期

- 🚆

铁路
1840s – 1870s
- 📻

无线电
1920s
- 📶

光纤
1997 – 2001
- ☀️

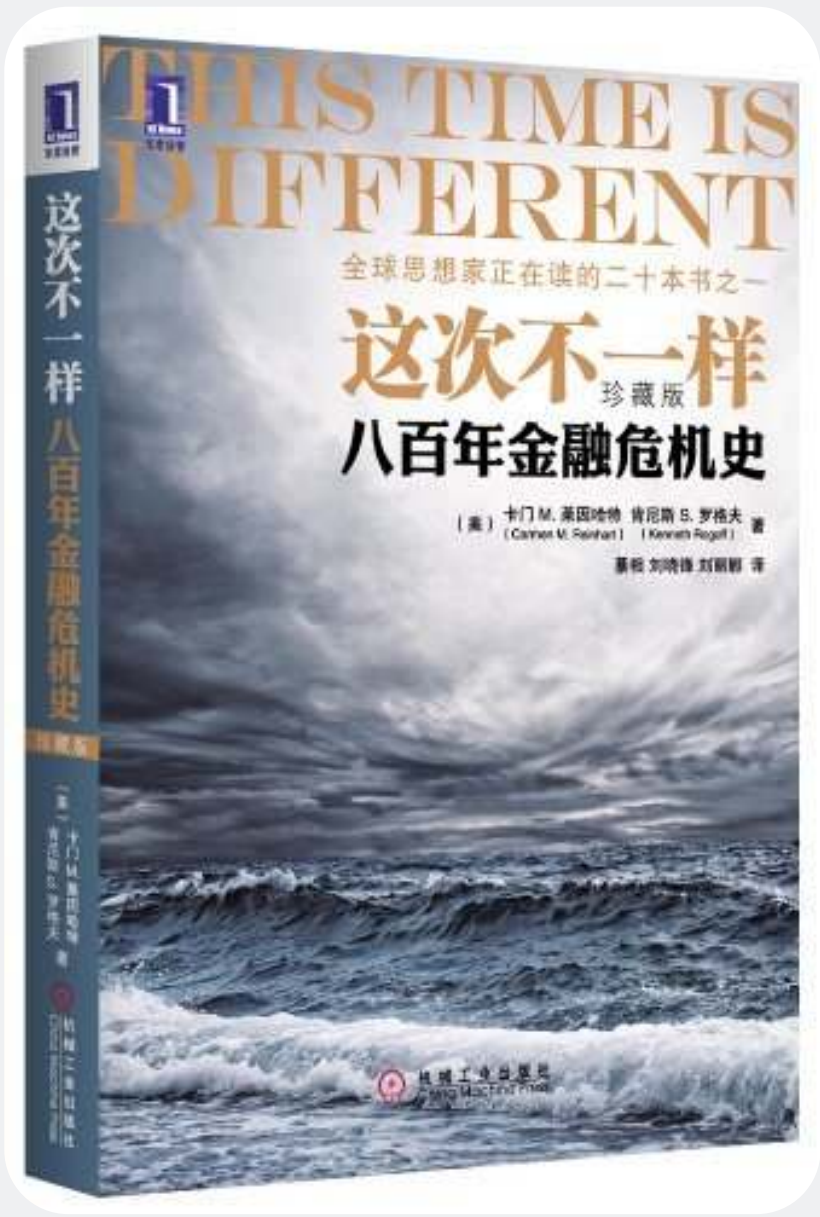
太阳能
2005 – 2012
- 🤖

人工智能
现在

🔊 AI投资与历史案例共同模式

特征	AI投资
革命性技术	✅ 是
大规模资本需求	✅ 是
政府支持	✅ 是
"这次不同"叙事	✅ 是
供应商融资	✅ 是
过度建设	✅ 是
崩溃与整合	🟡 待观察
最终价值创造	🟡 待观察

历史表明，技术资本支出超级周期总是以糟糕的结果结束，但最终创造长期价值



← → 🔍 🔍 所以“这次不一样”.....吗?

Exhibit 11: The 2000s saw a decoupling of prices from earnings

Cisco share price and earnings expectations

Index level, rebased to 100 in January 1998



Source: IBES, LSEG Datastream, J.P. Morgan Asset Management. Past performance is not a reliable indicator of current and future results. Data as of 12 November 2024.

Exhibit 12: Today, rapidly rising earnings have supported tech valuations

Nvidia share price and earnings expectations

Index level, rebased to 100 in January 2020

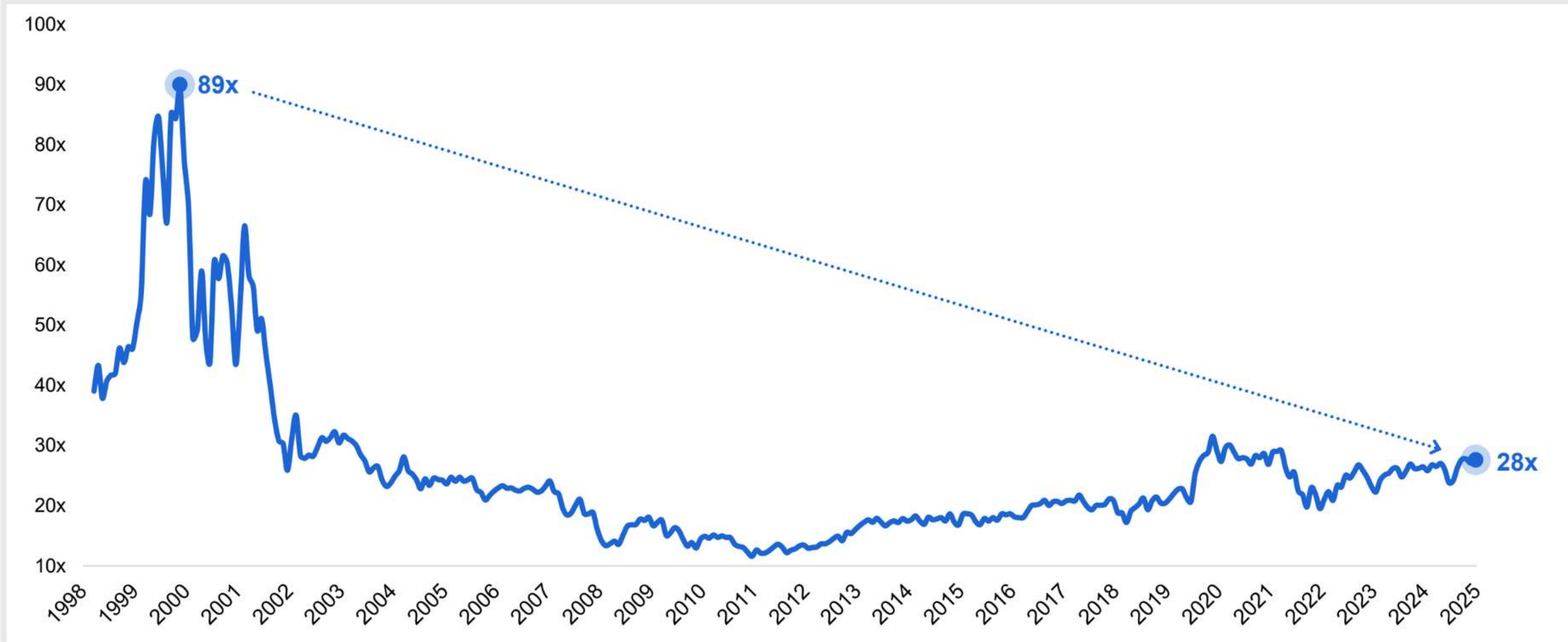


Source: IBES, LSEG Datastream, J.P. Morgan Asset Management. Past performance is not a reliable indicator of current and future results. Data as of 12 November 2024.

by Lin @speculator_io

← → 🔍 所以“这次不一样”.....吗?

NDX NTM P/E multiple over time

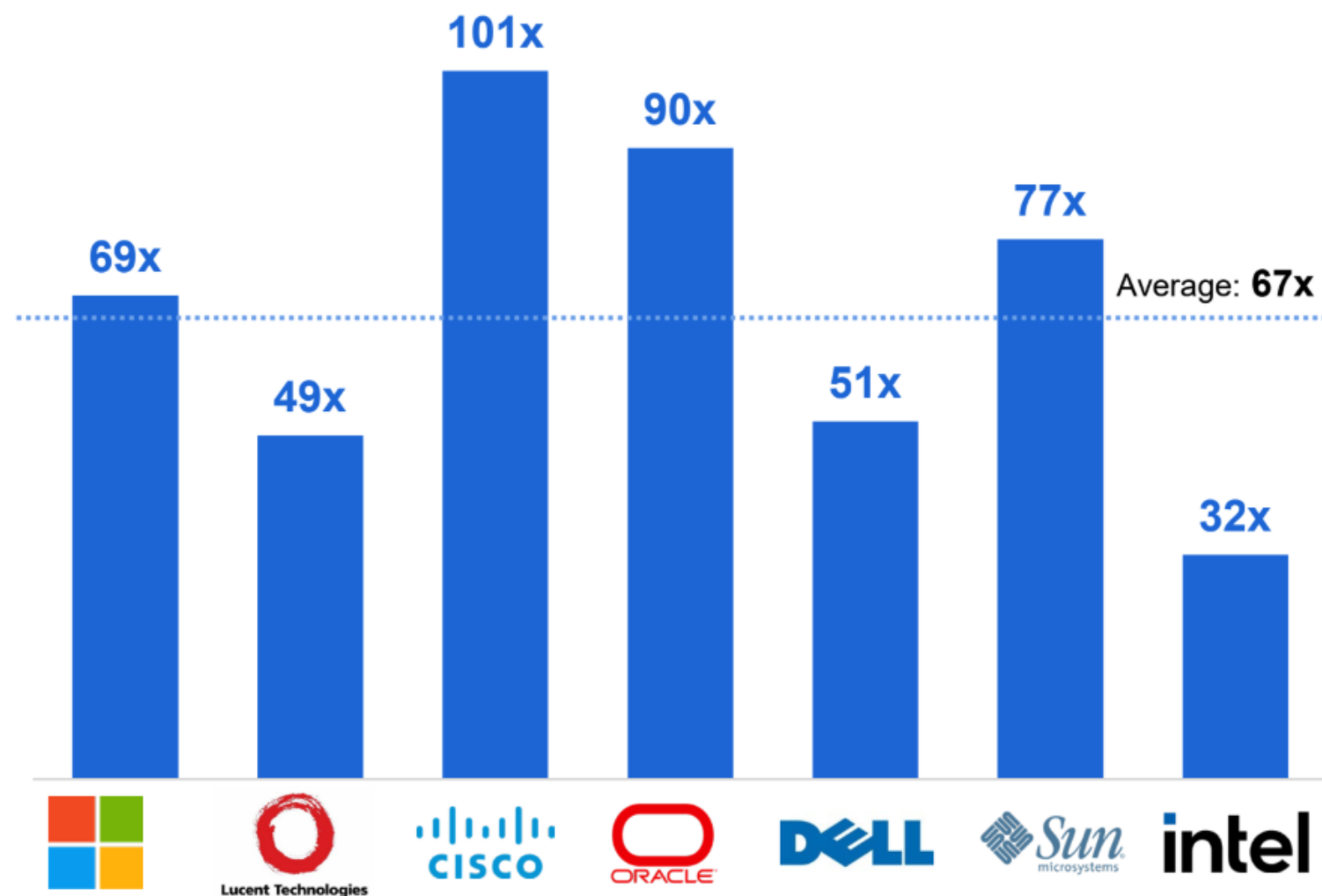


← → 🔍 所以“这次不一样”.....吗?

Valuations of the largest tech companies are lower vs. 1999

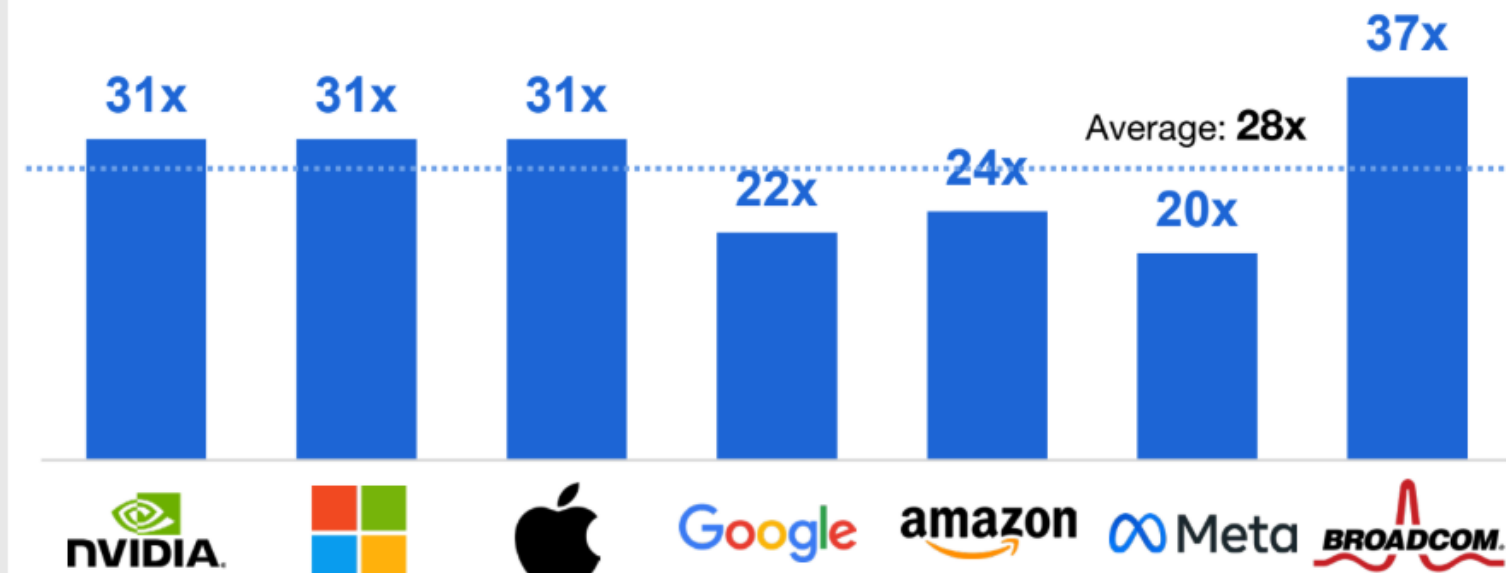
NTM P/E (1999)

Top 7 Tech of the 90s



NTM P/E (today)

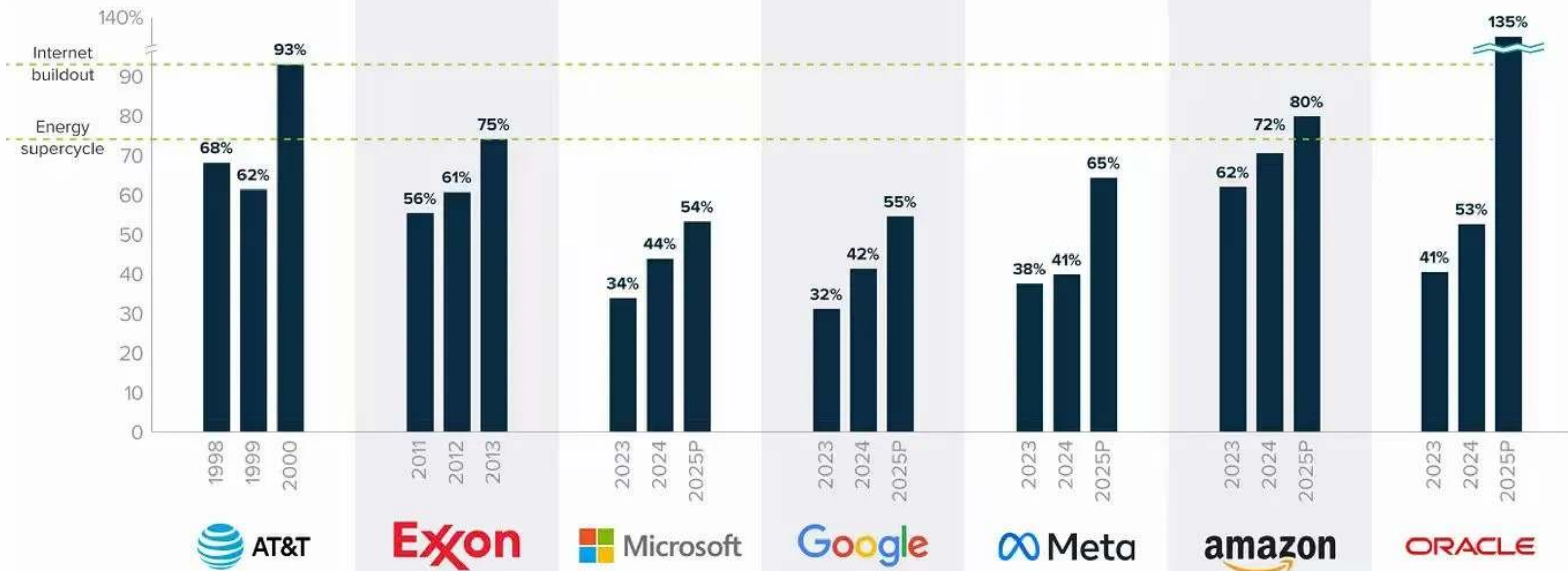
Top 7 Tech of Today



← → 🔍 所以“这次不一样”.....吗?

CapEx as % of operating cash flows

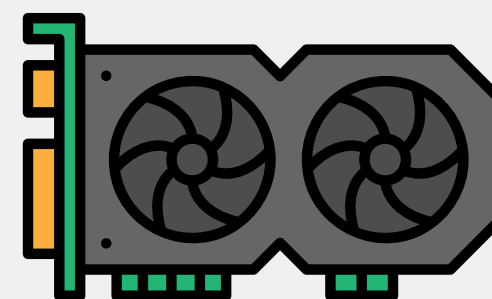
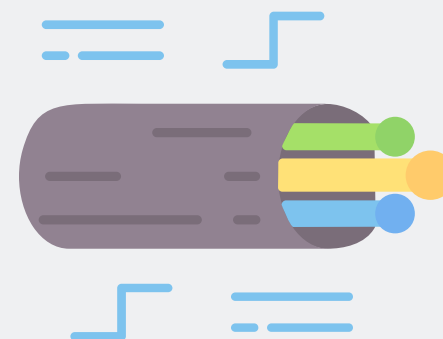
al6z Growth



← → 🔍 真的不一样，但好像更糟了.....



降息周期， 停止缩表



折旧期

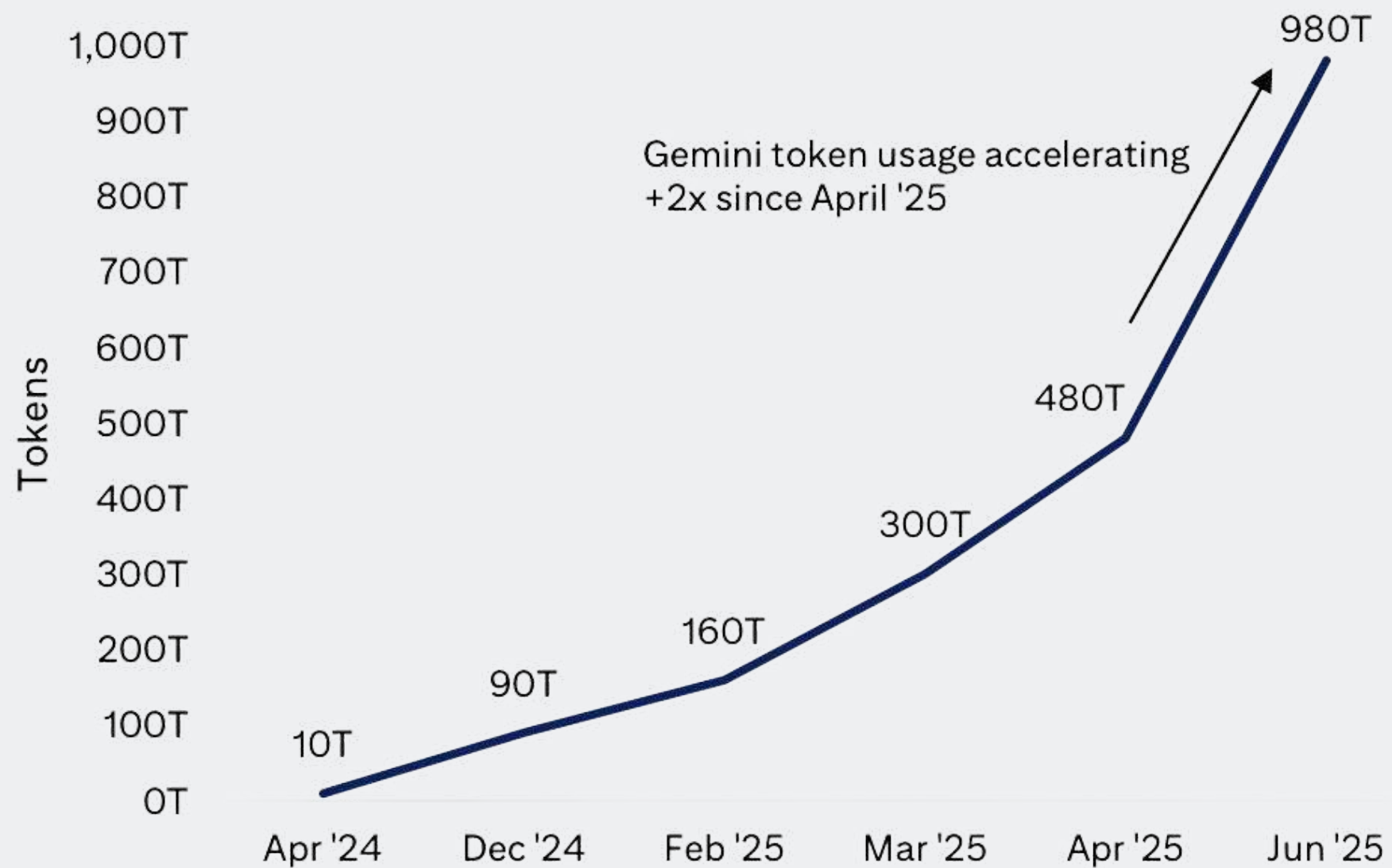
25-30年

25-30年

6年？3年？
甚至2年？

← → 🔍 我们该如何看待现阶段巨头们的Token的爆炸增长?

Figure 1. Gemini Token Usage Accelerating through 2Q

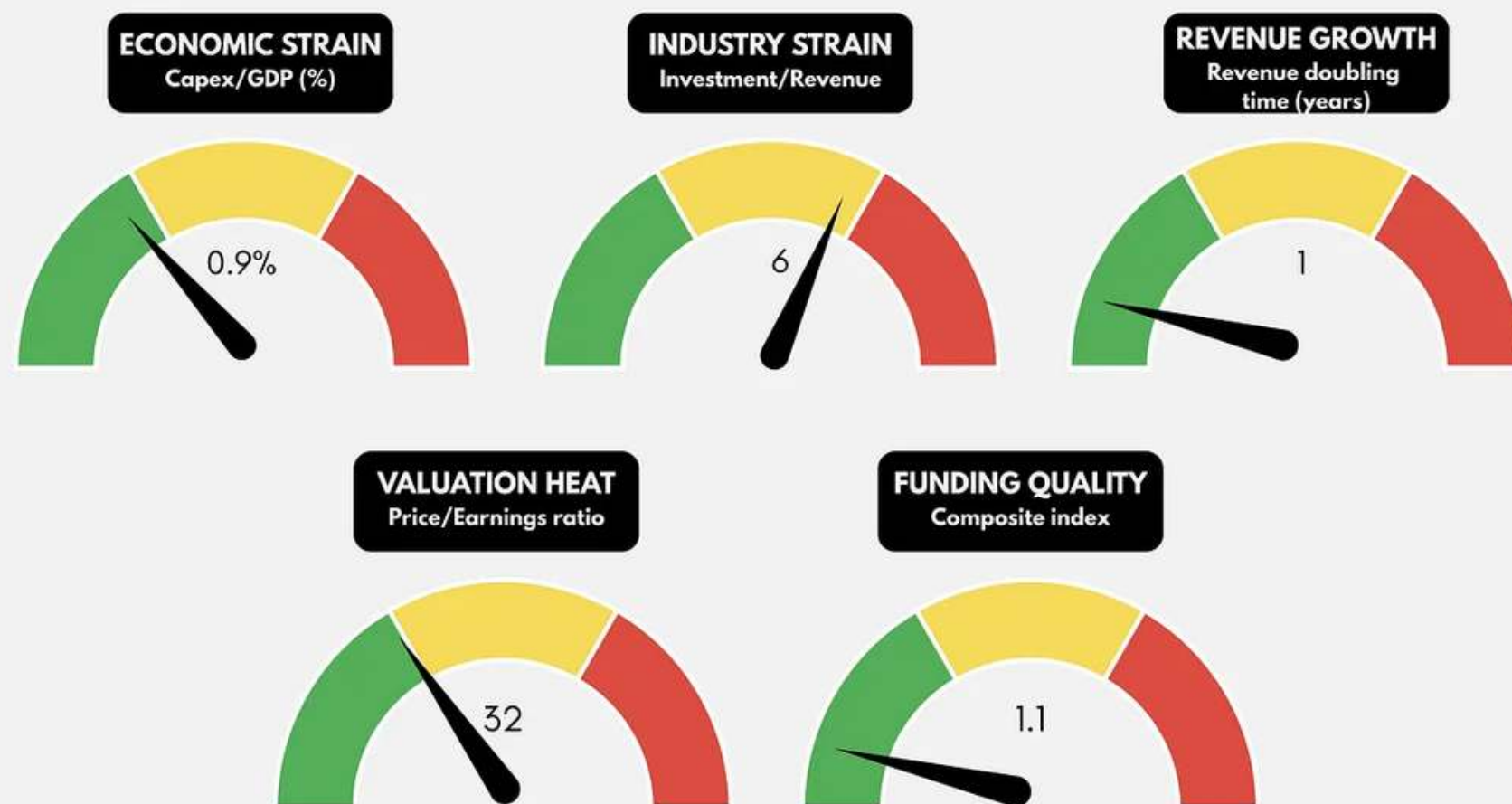


我们终于见到了Quadrillion

← → 🔍 我们能否用点更“科学”的评判方式？

Two reds = trouble

AI boom or bubble panel

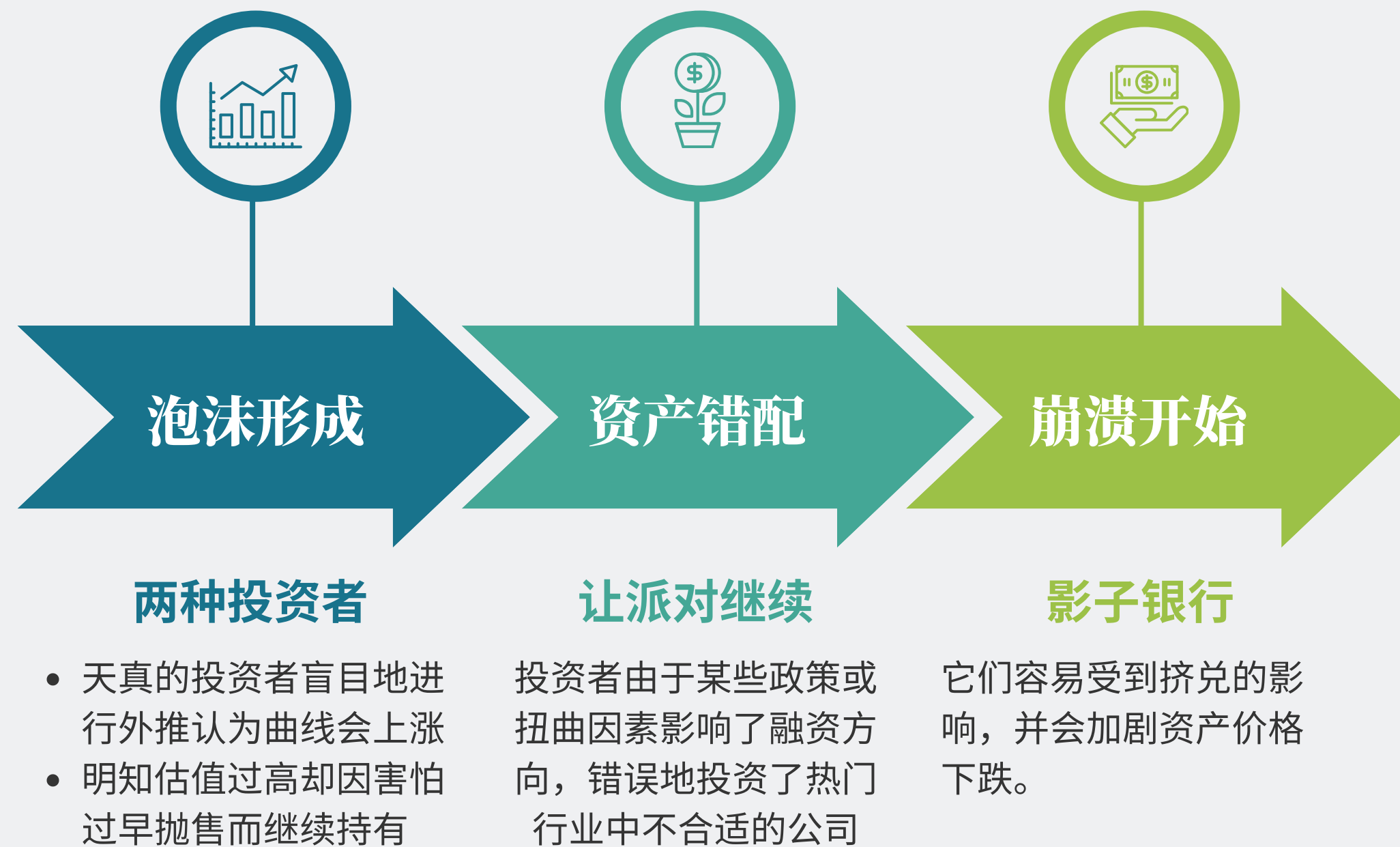
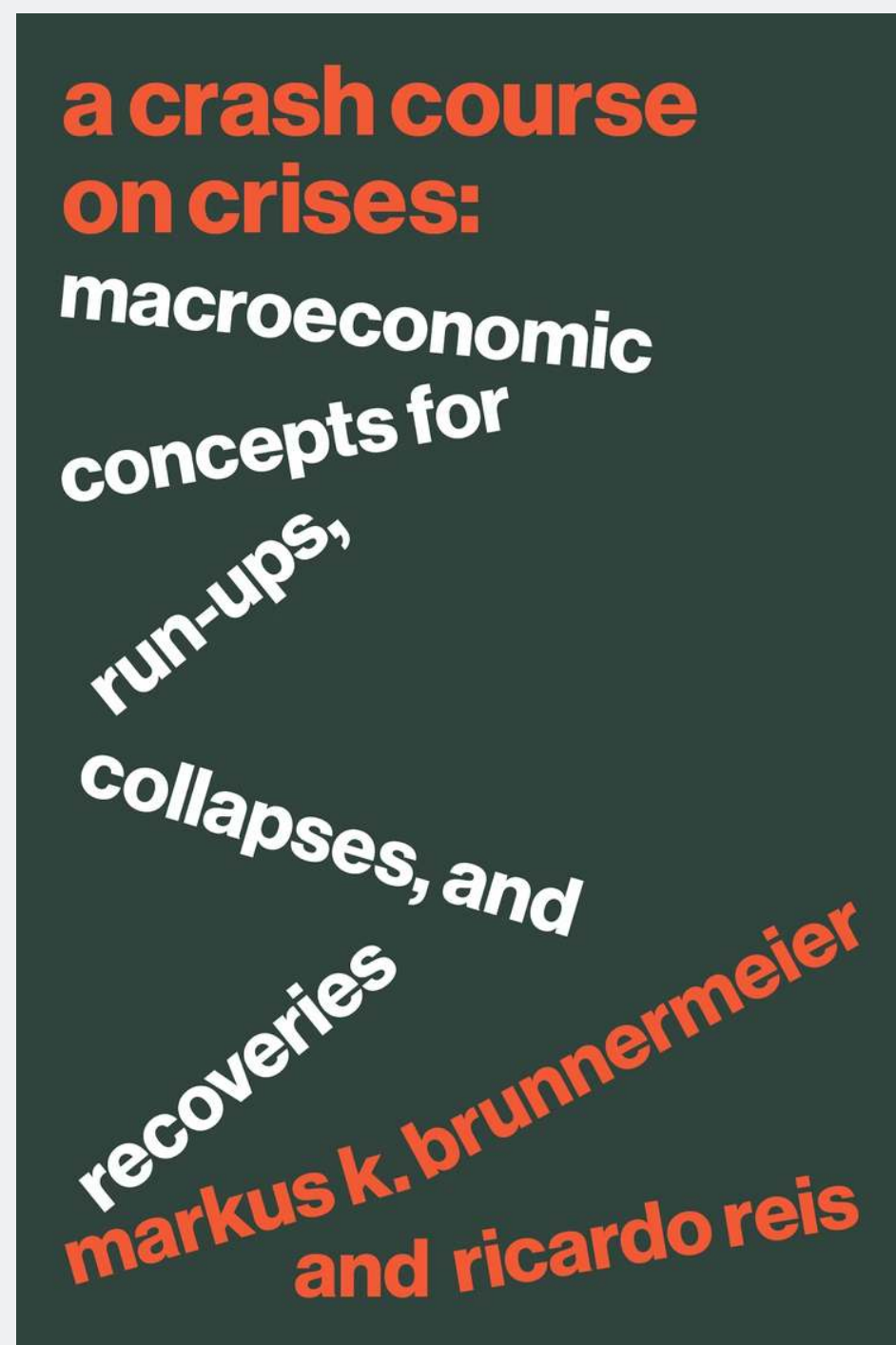


Exponential View 2025

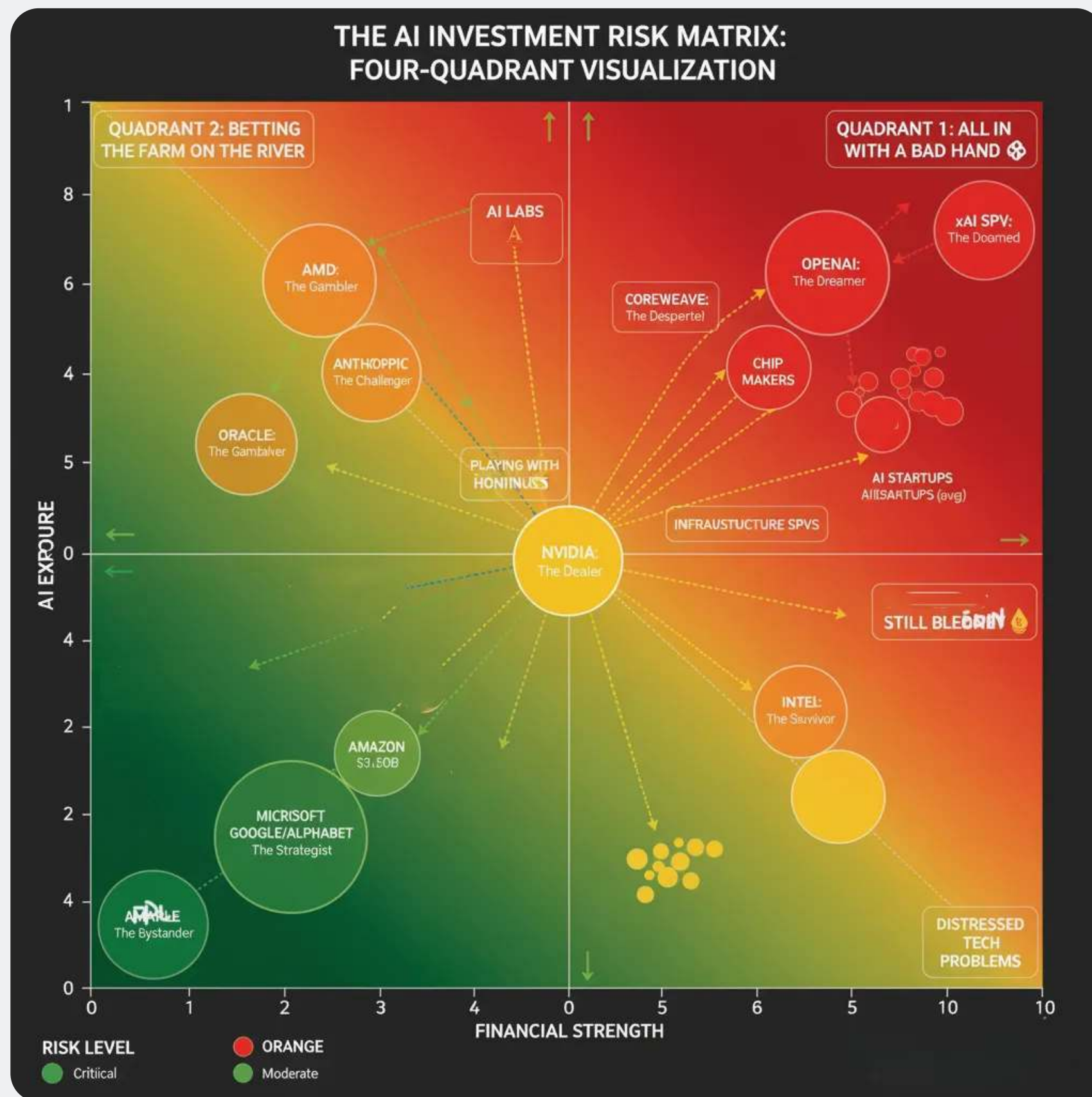
- 指标1 – 经济压力：当前的投资规模大到足以使整体经济“变形”吗？
- 指标2 – 行业压力：行业营收与投入的资本支出相称吗？
- 指标3 – 收入增长：收入增长得足够快、范围足够广，能赶上投入吗？
- 指标4 – 估值热度：估值有多火热？对比历史，股票是否被严重高估？
- 指标5 – 资金质量：资金来源来自稳健的资产负债表，还是脆弱的逐利资本？



《危机速成课：宏观经济概念在经济上升、崩溃和复苏中的应用》



← → 🔍 我们用前面所有这些方法论看看围观的公司



全力以赴：人工智能的纸牌屋？

对价值 5000 亿美元至 1 万亿美元的循环投资/融资生态系统的推测性分析



有趣的工程++

2025年10月13日



分享

...

“在扑克中，你玩的不是牌，而是对手。”

在人工智能中，每个人的筹码都是从桌子上借来的。”

- 高风险集中在（象限1）：CoreWeave和xAI SPV
- 微软和亚马逊等稳居“堡垒”（风险极低）象限
- 英伟达是“央行”
- AMD这样的公司正在进行“生存赌注”（象限2）

← → 🔍 可如果我们站在公司的角度上看“**杠杆**”



- 绝对意义上金融背景核心管理团队
- 在之前的核心业务是区块链矿场

NEBIUS

- 原Yandex, 因俄乌制裁被迫拆分
- 1000员工背井离乡在荷兰另起炉灶



**然后你和我说
我们杠杆高?
风险系数高?
COME ON...**

← → 🔍 正说着反着说都是你，就你“**大聪明**”是吧？

• 不要做**早下车**的大聪明

最大的常识最后往往会形成最大的共识，由此带来的价格走势往往都是非常极致的。

聪明人的诅咒往往是因为太早看到了常识，从而看不起常识；

在行情远没有演绎完的时候，就去刻意寻找反共识或非共识的观点。

@ 某天使基金GP

• 很多人**急于套现**，这样他们就可以说“我赢了”

投资者为市场调整做准备或试图预测市场调整而损失的钱，

远远多于在市场调整本身中损失的钱。

@ 彼得·林奇

• 对很多人来说，泡沫的定义其实只是一个他们**没有参与的牛市**

所以当我们讲“担心泡沫”的时候，究竟是指担心泡沫马上要爆，

还是担心出现泡沫而自己不在场？

@ 投资聚义厅-宝哥



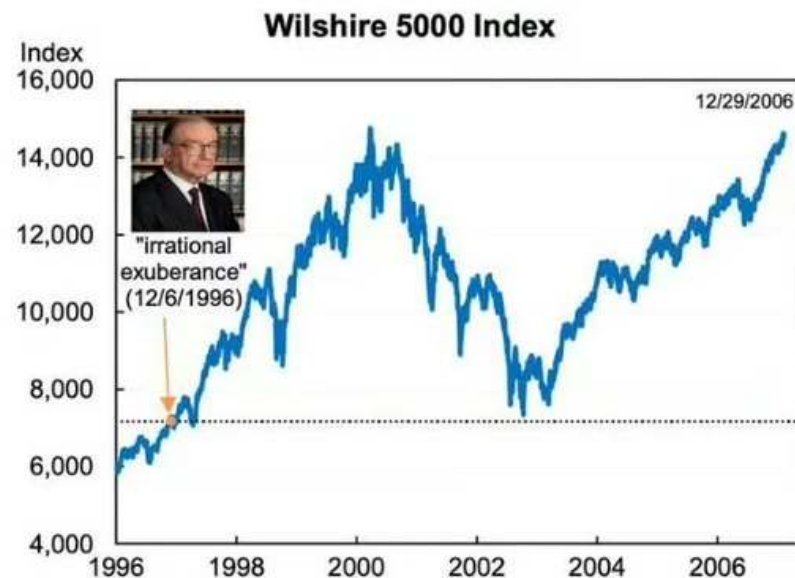
杰森·弗曼 @jasonfurman · 1小时

如果你过早地宣布泡沫，你就没有先见之明，你就错了。

如果你在艾伦·格林斯潘宣布“非理性繁荣”的那天买入股市，无论你最终在哪一天卖出，你都会赚钱。

(不要@我关于通货膨胀调整。)

If you call a bubble early you're not prescient, you're wrong



14

30

171

18K

18



开场

回顾

技术演进

产品叙事

资本趋势

泡沫大讨论

收尾



← → 🔍 🔍 我今天要讲的内容差不多就是这些了

听完看完什么感受？

是不是感觉：
好像什么都讲到了，
但又什么都没有讲…

那就对了！

狂喜播客节
POD KUANJIAL



屠龙之术

 庄明浩

你好，我是庄明浩。
这里是我的个人播客节目《屠龙之术》。

 小宇宙

长按扫码订阅节目



$$12/101 = 11.9\%$$



$$8/128 = 6.4\%$$

什么他妈AI不AI的
请享受生活
享受狂喜吧