



The Strategic Playbook Series

Zhipu AI (2513.HK) China's Emerging Anthropic?

Equity Advisory, IPS, IWM

2026.7.27

Who is Zhipu AI (2513.HK)?

Tsinghua-affiliated, technology-driven large-model company built on a self-developed Open Source GLM architecture, focused on the enterprise (B2B)

Comparison dimensions	Anthropic	OpenAI	Zhipu AI (Zhipu)	MiniMax (Xiyu Technology)
Company profile	Currently the highest-valued AI company globally, focused on advanced productivity; renowned for top-tier logical reasoning, code generation, and AI safety. Financial position is strongest; management expects near-term profitability.	Global AI trend leader emphasizing full-modality capabilities and a platform-level AGI ecosystem. Strongest brand but faces the greatest financial pressure; returning to profitability is the core concern.	Leveraging Tsinghua academic roots, focused on domestically developed foundations and privatized government/enterprise deployments—positioned as a B2B enterprise service pillar.	Focused on native multimodality and globalization, leveraging voice/video generation and consumer 'Companion' apps for overseas expansion; a smaller player with relatively weaker risk resilience.
Primary business lines	Claude model series — safety-first and stability-oriented; primarily serves enterprise clients.	ChatGPT — largest global user base and strongest brand recognition among AI companies.	Chinese large-model company offering enterprise APIs and an AI coding tool, 'GLM Coding Plan.'	Chinese AI company driving overseas growth with the conversational companion app 'Talkie'; high proportion of revenue from overseas markets.
Listing status	Not listed; submitted a confidential filing to the U.S. SEC on 2026-06-01; earliest expected listing in Oct 2026.	Not listed; submitted a confidential filing to the U.S. SEC on 2026-06-08; expected to list between late 2026 and 2027.	Listed 2026.1.8 (2513.HK)	Listed 2026.1.9 (100.HK)
Current valuation/market cap	Approx. US\$965bn Priced after a US\$65bn financing round completed in May 2026, surpassing OpenAI to become the highest-valued private AI company globally.	Approx. US\$852bn Priced following a US\$122bn financing round completed in March 2026; ranks below Anthropic.	Approx. US\$52.6bn Peaked at US\$144bn on June 22; halved in mid-July after a near-28% one-day drop.	Approx. US\$8.6bn Spiked to nearly US\$50bn in March amid AI hype, then fell sharply.
Revenue scale	Annualized revenue rose from US\$1bn at end-2024 to US\$47bn by May 2026 within 17 months; revenue surpassed OpenAI from April.	Annualized revenue approx. US\$25bn — largest among the four, but growth noticeably slowed from Feb to Apr 2026.	Full-year 2025 revenue ~US\$0.1bn; latest annualized 2026 revenue ~US\$1bn.	Full-year 2025 revenue ~US\$0.079bn — smallest of the four.
Profitability	Relatively strongest financial position; management states a near-term return to profitability.	Largest losses — roughly US\$1.20 lost for every US\$1 earned; projected net loss US\$14–27bn in 2026.	2025 net loss approx. US\$0.47bn.	2025 net loss approx. US\$0.25bn.
Key strengths	AI coding tool 'Claude Code' drove rapid growth — related annual revenue reached US\$2.5bn within nine months, capturing over half of the AI coding market; significant enterprise budget shifts from OpenAI to Anthropic.	Clear brand and user-base advantages; enterprise customer revenue now accounts for ~40% of total.	Q1 API price increase of 83% accompanied by a 400% rise in usage, demonstrating material pricing power.	Over 70% of revenue from overseas; Open Source models offer strong cost-performance and enjoy a good reputation in the developer community.
Key risks	Not yet publicly listed; current valuation is largely supported by private investors — validation in the public markets remains uncertain.	Sustained long-term loss pressure; market forecasts indicate stable profitability may not occur until around 2030.	Small revenue base and high short-term stock volatility; valuation may be materially detached from actual earnings capacity.	Revenue highly concentrated in the single product 'Talkie'; if user engagement declines or major competitors re-enter, growth momentum will be at risk.

Open Source vs Closed Source

Open Source (like buying a washing machine): high autonomy and customizability;
Closed Source (like using a laundromat): low deployment barrier and rapid deployment

Fundamental difference: whether code and model weights are publicly released, and whether users have the right to

Comparison Dimensions	Open Source	Closed Source
Representative Companies/Models	Meta (Llama), DeepSeek, Zhipu AI (GLM Open Source)	OpenAI (GPT-4), Anthropic (Claude), Google (Gemini)
Suitable Customers (depends on the trade-off between data security and operational cost/complexity)	Large enterprises, financial/healthcare/government organizations with strict data-privacy compliance, or firms needing scenario-specific competitive advantages — self-hosted Open Source model deployment is the preferred option.	Startups, lightweight app developers, or non-sensitive businesses seeking rapid MVP validation — integrating Closed Source models via API is the fastest, most resource-efficient approach.
Advantages	Absolute data security: sensitive data (e.g., financial, medical, government secrets) stays on-premise, eliminating leakage risk. Deep customization: vertical fine-tuning for industry-specific models. No vendor lock-in: technical sovereignty retained.	Ready to use: a few lines of code to access world-class AI via API. High performance: backed by giants' multi-billion-dollar R&D and compute, with rapid iteration. Zero hardware maintenance: no need to buy GPUs or manage complex infrastructure.
Disadvantages	High hardware and talent barrier: requires expensive GPU servers and skilled ML engineers. Top-tier performance may lag: state-of-the-art general capabilities (e.g., inference limits) are often led by Closed Source giants.	Data privacy concerns: data must be uploaded to the cloud, unsuitable for highly sensitive operations. Uncontrolled costs: token fees scale linearly with user and call volume. Vendor dependency: if a provider changes API pricing, policies, or disables a model, the enterprise is vulnerable.
Transparency	High: model weights/source code open for inspection and audit	Low: black-box model accessible only via API
Data & Privacy	Very high: can be fully deployed on-premise	Lower: data must be transmitted to third-party cloud servers
Customization & Further Development	Highly flexible: supports fine-tuning and code modification	Limited: restricted to prompt engineering or vendor API parameter settings
Usage Cost	Requires own compute/servers; high initial hardware cost but no token fees	No need for own hardware; pay per-token/subscription (pay-as-you-go)
Operational Requirements	High: requires a professional MLOps/engineering team for maintenance	Low: turnkey solution; maintenance and upgrades handled by the model provider

China–US Large-Model Companies Comparison 1/2: Strategy, Positioning, Deployment, Service Scenarios

US: Closed Source, cloud-first, serving large enterprises on a variable-cost model

China: Open Source combined with privatized deployments, serving SMEs and government clients

Company/Organization	Core Strategy	Strategic Positioning	Deployment	Core Service Scenarios
Zhipu AI	Open Source + On-Premise balance full-stack in-house R&D with industry–academia collaboration; MaaS API and on-premise deployments equally emphasized; deep focus on government and enterprise markets	China leader in domestic AI substitution; Hong Kong-listed as the "first large-model" stock	SaaS APIs, privatized (On-Premise) local deployment, integrated hardware–software appliances	Government/enterprise knowledge bases; vertical large models for finance, healthcare, energy; AI Agent development platforms; academic and research tools
Alibaba Qwen	Open Source + cloud integration Qwen's Open Source ecosystem drives Alibaba Cloud MaaS consumption	Open Source ecosystem plus cloud-integrated strategy; Alibaba Cloud infrastructure provides core stickiness	Comprehensive coverage: public cloud APIs, private cloud, and Open Source models available for local deployment	E-commerce intelligence (customer service/recommendation/content generation), cloud enterprise APIs, developer Open Source ecosystem, DingTalk collaboration
ByteDance Doubao	Closed Source ultra-low-cost APIs + traffic ecosystem (90%+ cheaper than peers), capturing business (B) customers while building a consumer (C) super-entry	Traffic giant with highly cost-effective APIs; Volcengine as the cutting-edge weapon	Primarily public cloud APIs (Volcengine) and lightweight edge-side operation	Consumer entertainment/search, Douyin/TikTok content generation and marketing, high-frequency low-cost B2B APIs
Moonshot AI (Kimi)	Hybrid: APIs + partial Open Source tech-driven with viral consumer products, focused on long-context and agents; centered on Kimi subscription model	Consumer blockbusters and long-text/Agent-led unicorn (highest consumer user stickiness among native AI startups)	Primarily public cloud SaaS APIs and Kimi consumer App/Web services	Long-form analysis, office assistants, deep search, AI collaborative writing and coding
DeepSeek	Open Source algorithmic innovation + extreme compute efficiency, technologies shared Open Source	Pioneer in algorithmic efficiency innovation and Open Source disruption	Fully Open Source models (local/quantized deployment) + official low-cost APIs	Mathematics/logic reasoning, code generation (Coder), community-driven Open Source fine-tuning
Anthropic	Closed Source AI safety (Constitutional AI) + high-security enterprise-grade AI	Top independent giant second only to OpenAI; core to strategic alliances with Google and Amazon	Public cloud APIs, integrated with Amazon Bedrock and Google Cloud Vertex AI	Complex programming (Claude Code), enterprise compliance, long-form reasoning, data analysis
OpenAI	Closed Source pursuing AGI and a full-stack ecosystem, software–hardware integration, serving both consumer and enterprise markets	Global AGI leader and consumer super-portal; the highest-valued AI unicorn	ChatGPT platform (SaaS), OpenAI API, Azure OpenAI Service (private/hybrid cloud)	Consumer AI assistants, multimodal interaction (voice/video/image), B2B AI workflows, Agent automation

China–US Large-Model Companies Comparison 2/2: Customers, profitability & commercialization, shareholders and risks

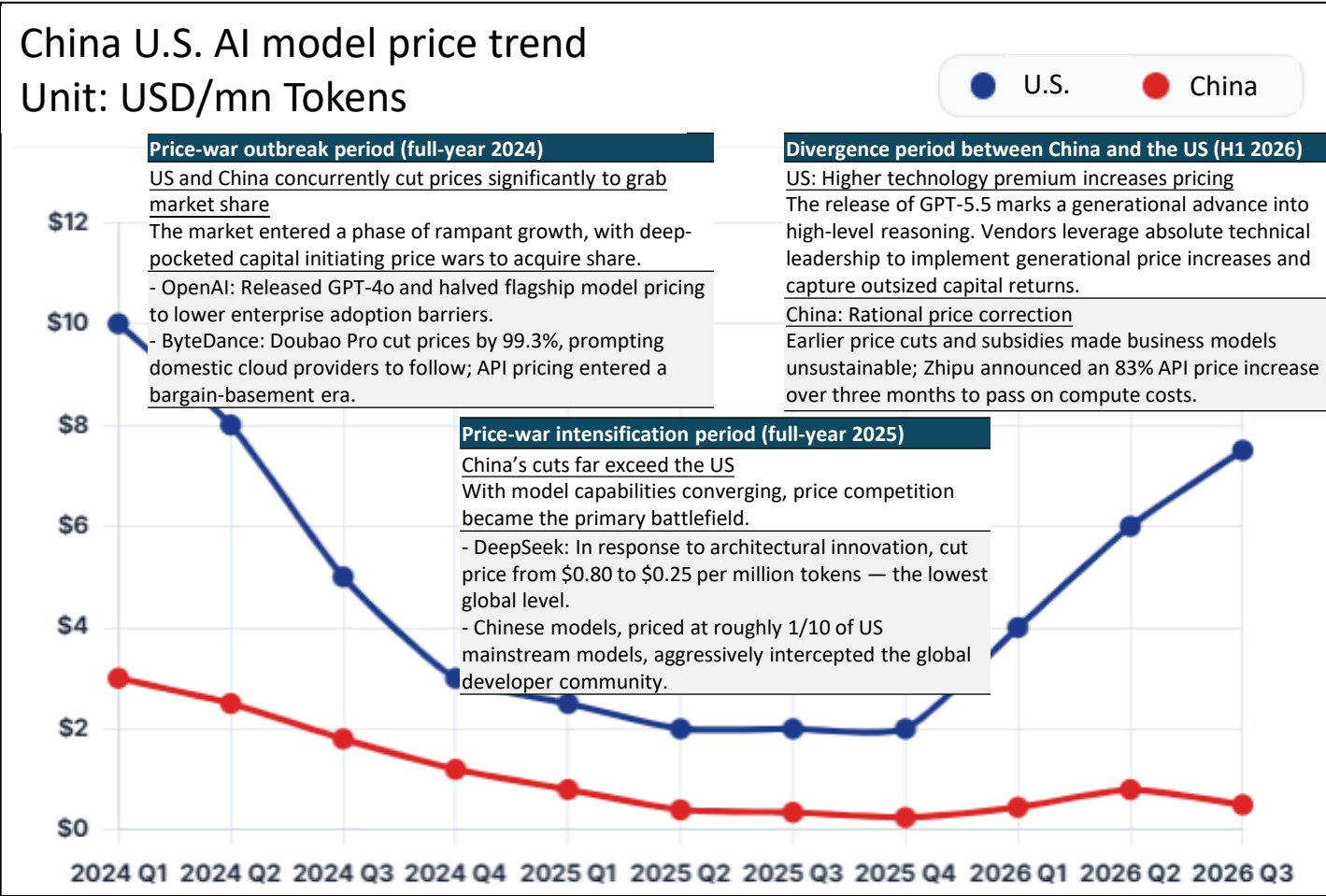
US: loss-making, relying on global enterprise customers and capital-market financing to drive commercialization;

China: loss-making, supported by internet giants and government resources, focused on aggressively capturing market share.

Company / Institution	Primary Customers	Profitability / Commercialization Status	Major Shareholders	Key Risks
Zhipu AI	Central/state-owned enterprises, government agencies, financial institutions, research institutes	Revenue growing rapidly; remains in strategic losses	Founding team: Tsinghua KEG Lab; partners/investors include Meituan, Alibaba, Tencent, state-owned funds	Relies on chip supply; private/on-prem deployment projects have long delivery cycles and high costs
Alibaba Qwen	Taobao/Tmall merchants, Alibaba Cloud enterprise customers, global Open Source community	Model-as-a-Service (MaaS) billing driving Alibaba Cloud revenue growth; commercialization is stable	Alibaba Group (100% owner)	Low monetization rate for Open Source APIs; limited availability of high-end GPUs constrains compute scaling
ByteDance Doubao	Consumer end-users, internet/game developers, brand marketers	Backed by ByteDance's cash flow; aggressive pricing has reduced margins and it is not yet independently profitable	ByteDance (100% owner)	Price wars compress margins; consumer retention and willingness to pay need improvement
Moonshot AI (Kimi)	Consumer knowledge users (students, researchers, white-collar workers) and developers requiring long-text/agent processing	Subscription and API revenues have surged, but losses persist due to high marketing and compute costs	Founders/early team from Tsinghua University, Alibaba, Tencent, Meituan Longzhu; investors include IDG Capital	Vulnerability to big-tech price wars and technology replication; high marketing and compute costs
DeepSeek	Global developers, research institutions, cost-efficient enterprise teams	Backed by parent High-Flyer Quant; API revenue emerging; not targeting short-term profitability	Founder Liang Wenfeng and entities affiliated with High-Flyer Quant	Lean commercialization team; may face geopolitical technology restrictions
Anthropic	Large multinational enterprises (financial / legal / healthcare), software engineering teams	B2B API and enterprise subscription revenue growing rapidly; remains deeply loss-making and in fundraising phase	Founding team (Dario Amodei, etc.); investors include Amazon, Google, Salesforce Ventures	Heavily reliant on AWS/GCP; squeezed by OpenAI and Open Source models
OpenAI	Hundreds of millions of consumer users globally, over one million developers, Fortune 500 companies	ARR in the billions of USD, but training next-generation frontier models incurs massive compute costs; continues to burn cash and post losses	Microsoft (largest shareholder), Thrive Capital, SoftBank, Khosla Ventures	Internal governance and restructuring risks; very high compute costs; competitive and regulatory pressures

Token Pricing Analysis 1/3: China–US Pricing Trend Shifts

US: Maintain high-end premium; China: Move toward tiered pricing



Current AI model pricing is primarily composed of three main factors: **compute cost** + **capability premium** + **competitive strategy**

Factors	Core concepts	Zhipu AI GLM-5.2 Price (\$4.40)	OpenAI GPT-5.5 Price (\$30.00)
Compute cost	The lower bound of model inference costs — including GPU time, memory usage, and energy consumption. This establishes the price floor; pricing below this implies operating at a loss.	25% (\$1.10)	15% (\$4.50)
Capability premium	Excess pricing attributable to performance differences, brand trust, and ecosystem lock-in. Mainstream providers therefore tier offerings into flagship (SOTA), mid-tier, and economical models for pricing.	35% (\$1.54)	50% (\$15.00)
Competitive strategy	Vendors price according to market-share objectives, competitive dynamics, and ecosystem strategy. During price wars margins can be compressed to near zero; when demand is strong there is scope for price increases.	40% (\$1.76)	35% (\$10.50)

Market	Trend direction	Drivers
China market	Shift from 'ubiquitous low pricing' to 'tiered pricing': mass-market offerings remain low-cost while premium lines gradually increase prices.	Rising compute costs, robust demand, and wider adoption of peak–valley (time-of-use) pricing. Goldman Sachs: adjusted blended average ~ \$0.35 per million tokens.
US market	High-end premium under pressure and forced to follow price cuts.	Chinese model performance catch-up (gap narrowed to 1–4%) is causing enterprise client churn. OpenAI CEO expects that by end-2027 the cost of achieving GPT-5.2-level intelligence will fall by at least 100x.

Token Pricing Analysis 2/3: Zhipu AI Pricing Strategy

Zhipu AI models: higher than domestic peers but price-advantaged versus US flagship models

Token: the unit of text processed by AI; analogous to 'water usage' billing — charged by consumption; 4 English letters = 1 Token; 1 Chinese character = 1.5–2 Tokens

Input: content fed to the AI (questions/commands)

Output: content generated by the AI in response; typically 3–6× more expensive than input (generation is more computationally intensive than reading)

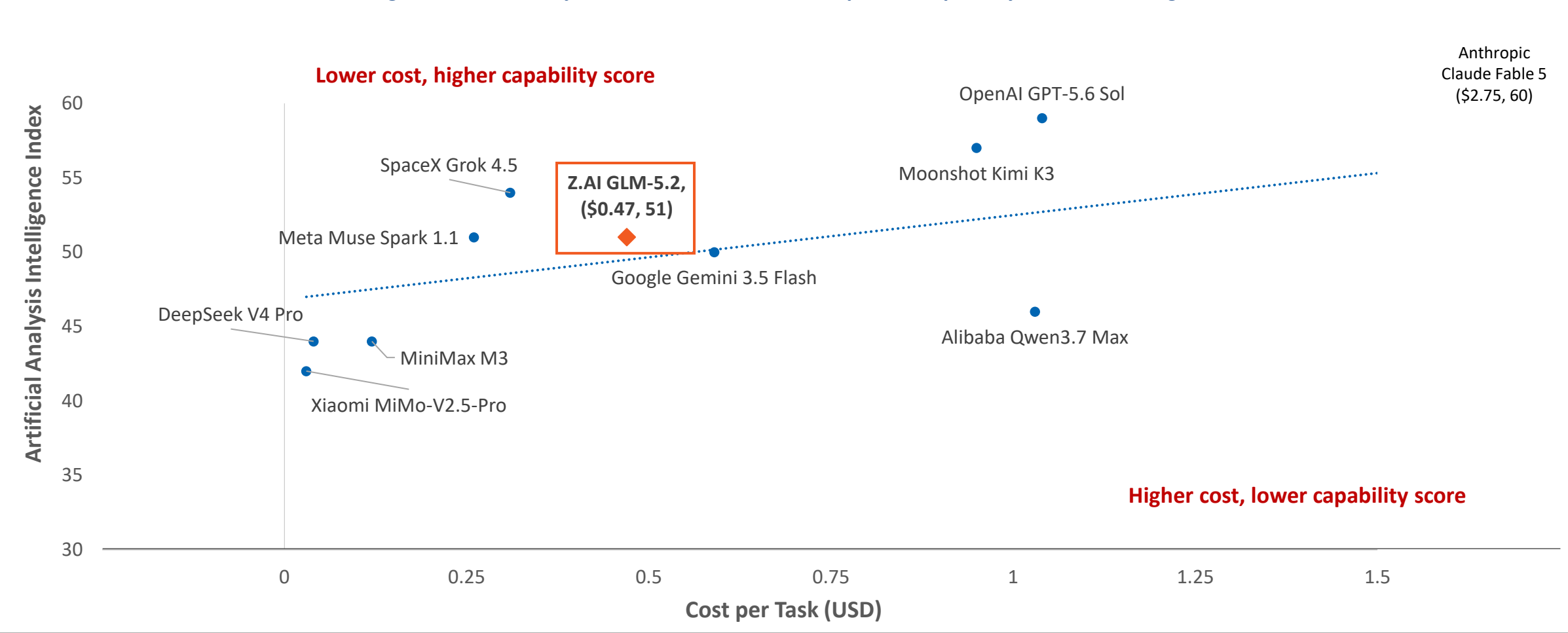
Company	Flagship Model	Input Price (USD per M tokens)	Output Price (USD per M tokens)	Positioning
MiniMax	M3	\$0.30	\$1.20	Cheapest on market
Alibaba Qwen	Qwen3-Max	\$0.37	\$1.47	Lowest-priced flagship in the Open Source cohort
Doubao (ByteDance)	Doubao-Seed-2.0-Pro	\$0.47	\$2.35	Primary mid-to-low price offering
DeepSeek	V4 Pro	\$1.74	\$3.48	Open Source high value flagship
Zhipu AI	GLM-5.2	\$1.40	\$4.40	Value-oriented offering
OpenAI	GPT-5.6 Terra	\$2.50	\$15.00	Mid-tier
Anthropic	Claude Sonnet 5	\$3.00	\$15.00	Key mid-tier player
Moon's Dark Side Kimi	K3	\$3.00	\$15.00	Open Source flagship, positioned against top closed-source models
Anthropic	Claude Opus 4.8 (flagship)	\$5.00	\$25.00	Top tier
OpenAI	GPT-5.6 Sol (flagship)	\$5.00	\$30.00	Top tier
Anthropic	Claude Fable 5 (most powerful)	\$10.00	\$50.00	Most expensive on the market

Note: RMB-priced models converted to USD using the July 2026 exchange rate of approximately 1 USD = CNY 6.8

Token Pricing Analysis 3/3: Zhipu AI pricing strategy

Zhipu AI models: capability score at 85% of the top but priced at only 17%; superior cost-performance is a competitive advantage

Leading models from major foundation model developers — capability scores and usage costs



Financials 1/2 – Income Statement

Z.ai (2513.HK) in RMB mil	2022	2023	2024	2025	2026E	2027E	2028E	2029E	2030E	25-30 CAGR
Revenue	57 100%	125 100%	312 100%	724 100%	5,494 100%	13,531 100%	17,885 100%	22,975 100%	28,736 100%	109%
YoY %		117%	151%	132%	659%	146%	32.2%	28.5%	25.1%	
On-premise (本地部署)	55 95%	113 90%	264 84%	534 74%	1,007 18%	1,753 13%	2,620 15%	3,522 15%	4,576 16%	54%
YoY %		105%	134%	102%	89%	74.0%	49.5%	34.4%	29.9%	
Cloud-based (雲端服務)	3 5%	12 10%	48 16%	190 26%	4,487 82%	11,778 87%	15,264 85%	19,453 85%	24,160 84%	163%
YoY %		360%	307%	293%	2257%	162.5%	29.6%	27.4%	24.2%	
Cost of sales	(26) 100%	(44) 100%	(137) 100%	(428) 100%	(3,732) 100%	(8,122) 100%	(9,907) 100%	(11,528) 100%	(13,465) 100%	
Payroll	(16) 63%	(15) 35%	(74) 54%	(112) 26%	(201) 5%	(316) 4%	(393) 4%	(423) 4%	(458) 3%	
Computing (Inference)	- 0%	(12) 27%	(30) 22%	(149) 35%	(2,848) 76%	(6,125) 75%	(7,327) 74%	(8,365) 73%	(9,664) 72%	
as a % of cloud-based Revenue	0%	-101%	-62%	-78%	-63%	-52%	-48%	-43%	-40%	
Gross Profit	31	80	176	297	1,762	5,409	7,978	11,447	15,271	120%
Gross Profit Margin	54.6%	64.6%	56.3%	41.0%	32.1%	40.0%	44.6%	49.8%	53.1%	
YoY %		156.6%	118.5%	68.7%	494.0%	207.0%	47.5%	43.5%	33.4%	
Operating expense	(130) 100%	(706) 100%	(2,714) 100%	(4,083) 100%	(8,642) 100%	(12,908) 100%	(14,885) 100%	(16,365) 100%	(18,015) 100%	
R&D expenses	(84)	(529)	(2,195)	(3,180)	(6,846)	(9,930)	(11,501)	(12,579)	(13,854)	
as a % of revenue	-147%	-425%	-703%	-439%	-125%	-73%	-64%	-55%	-48%	
Computing (Training)	(15) 11%	(312) 44%	(1,553) 57%	(2,159) 53%	(3,898) 45%	(6,089) 47%	(7,154) 48%	(8,041) 49%	(9,196) 51%	
Payroll	(42) 32%	(137) 19%	(324) 12%	(479) 12%	(1,374) 16%	(1,624) 13%	(1,788) 12%	(1,838) 11%	(1,868) 10%	
Adjusted Operating profit	(52)	(459)	(2,046)	(2,251)	(5,431)	(4,251)	(2,794)	136	3,003	
Operating Profit Margin	-91.4%	-368.4%	-654.9%	-310.7%	-98.9%	-31.4%	-15.6%	0.6%	10.5%	
YoY %		/	/	/	/	/	/	/	2107%	
Adjusted Net Income	(97)	(621)	(2,466)	(3,182)	(5,504)	(4,246)	(2,857)	30	2,870	
Net Income Margin	-170%	-499%	-789%	-439%	-100%	-31.4%	-16.0%	0.1%	10.0%	
YoY %		/	/	/	/	/	/	/	9498%	

Key forecast risks:

Revenue

25–30E; projected CAGR +109% — assumes successful model upgrades, globally competitive performance and cost structure, and continued ARPU and user growth contributing to revenue.

Opex — R&D:

Compute for model training = 439% of revenue (2025) — as China's compute network develops and supply increases, the company's procurement scale expands and training/cloud compute costs decline, supporting margin improvement.

Corporate governance & regulation

- No material changes to the core R&D team
- Regulation is evolving alongside the industry; the AI total addressable market (TAM) continues to expand, with increasing consumer and enterprise use cases and rising adoption rates.

Financial Statements 2/2 – Balance Sheet, Cash Flow Statement

Z.ai (2513.HK)		2022	2023	2024	2025	2026E	2027E	2028E	2029E	2030E
Balance Sheet (Rmb millions)										
Key Metric	Net Debt	(219)	(1,249)	(2,269)	(2,175)	(29,015)	(25,947)	(23,226)	(23,588)	(26,776)
	Trade Receivable in days		651	633	344	60	48	52	40	29
	CAPEX	(32)	(507)	(126)	(11)	(824)	(1,353)	(1,073)	(1,149)	(1,437)
	Depreciation	(17)	(64)	(270)	(273)	(207)	(922)	(1,912)	(2,776)	(3,712)
Assets (Key Element)	Trade and other receivables	28	416	667	699	1,099	2,436	2,683	2,297	2,299
	Cash at bank and on hand	219	1,249	2,269	2,259	29,115	30,947	30,226	30,588	33,776
	Property, plant and equipment	41	788	866	656	2,921	6,059	8,797	11,765	15,237
	Total Asset	362	2,860	4,376	4,854	34,587	41,108	43,621	46,826	53,788
Liabilities (Key Element)	Trade and other payables	26	288	603	1,329	2,361	3,664	4,055	4,101	4,715
	Contract liabilities	35	74	75	149	1,127	2,706	2,683	2,297	2,299
	Bank loans				85	100	5,000	7,000	7,000	7,000
	Lease liabilities	0	204	458	294	1,556	3,468	5,761	8,489	11,686
	Total Liabilities	542	3,843	8,331	12,965	6,550	17,047	22,143	25,041	28,855
Equity (Key Element)	Share premium	0	0	0	0	41,246	41,246	41,246	41,246	41,246
	Reserves	(196)	(1,011)	(3,993)	(8,133)	(13,637)	(17,612)	(20,195)	(19,889)	(16,740)
	Non-controlling interests	1	0	1	(18)	(18)	(18)	(18)	(18)	(18)
	Total Equity	(180)	(983)	(3,955)	(8,111)	28,036	24,061	21,478	21,784	24,933
Cash Flow Statement (Rmb millions)		2022	2023	2024	2025	2026E	2027E	2028E	2029E	2030E
Operating CF (Key Element)	Profit/(loss) before income tax	(144)	(788)	(2,958)	(4,718)	(6,953)	(7,494)	(6,970)	(5,025)	(2,878)
	Depreciation of PP&E	17	64	270	270	207	922	1,912	2,776	3,712
	Equity-settled share-based payments	1	6	24	558	1,449	3,247	4,113	5,054	5,747
	Cash flows from operating activities	(68)	(648)	(2,245)	(2,246)	(3,324)	(1,126)	(501)	3,285	7,112
Investing CF (Key Element)	Purchase of PP&E	(32)	(507)	(126)	(11)	(824)	(1,353)	(1,073)	(1,149)	(1,437)
	Cash flows from investing activities	33	(785)	(49)	(376)	(915)	(1,393)	(1,117)	(1,197)	(1,489)
Financing CF (Key Element)	Issuance/replacement of common shares	0	98	0	0	41,652	0	0	0	0
	Issuance of convertible bonds and bank loans	0	0	267	1,252	15	4,900	2,000	0	0
	Issuance of financial instruments	203	2,406	3,011	1,579	(10,073)	0	0	0	0
	Cash flows from financing activities	191	2,463	3,312	2,614	35,367	4,080	623	(2,003)	(2,713)
Cash and cash equivalents at end of year		219	1,249	2,268	2,255	29,115	30,947	30,226	30,588	33,776

Zhipu AI asset-light strategy

- Zhipu AI management states: "Zhipu AI primarily procures services, supplemented by selective equipment leases to meet demand, thereby reducing capital expenditures."
- Globally, asset-light is a common strategy — e.g., NVIDIA does not build fabs, Apple does not operate assembly plants; Zhipu AI's strategy is not to own large data centers
- Compute networks are one of the six national networks in Mainland China and a construction priority, aimed at meeting demand for compute resources and lowering enterprises' compute costs and barriers

Zhipu AI completed two fundraising rounds in H1

	New shares (mn)	% of post-fundraising total shares	Issue price	Proceeds (RMB bn)
2026.1.8 IPO	43.0	9.65%	HKD 116.2	4.5
2026.7.9 Placement	19.8	4.25%	HKD 1588.0	27.1
Total proceeds				31.6

Zhipu AI year-end cash balance approximately Rmb 29.1bn
Current cash level can support roughly six years of CAPEX

Cash flow assumptions (RMB mn)	
Operating cash flow	-3,324
Capital expenditures	-1,487
Total	-4,813

*In the model, the higher figures were assumed

Valuation methodology

P/S and PSG as benchmark multiples

	Fair value	Valuation methodology			Fair value implied P/S multiple			
	HK\$/share	Valuation method	Benchmark multiple (2026E)	Rationale	2027E	2028E	2029E	2030E

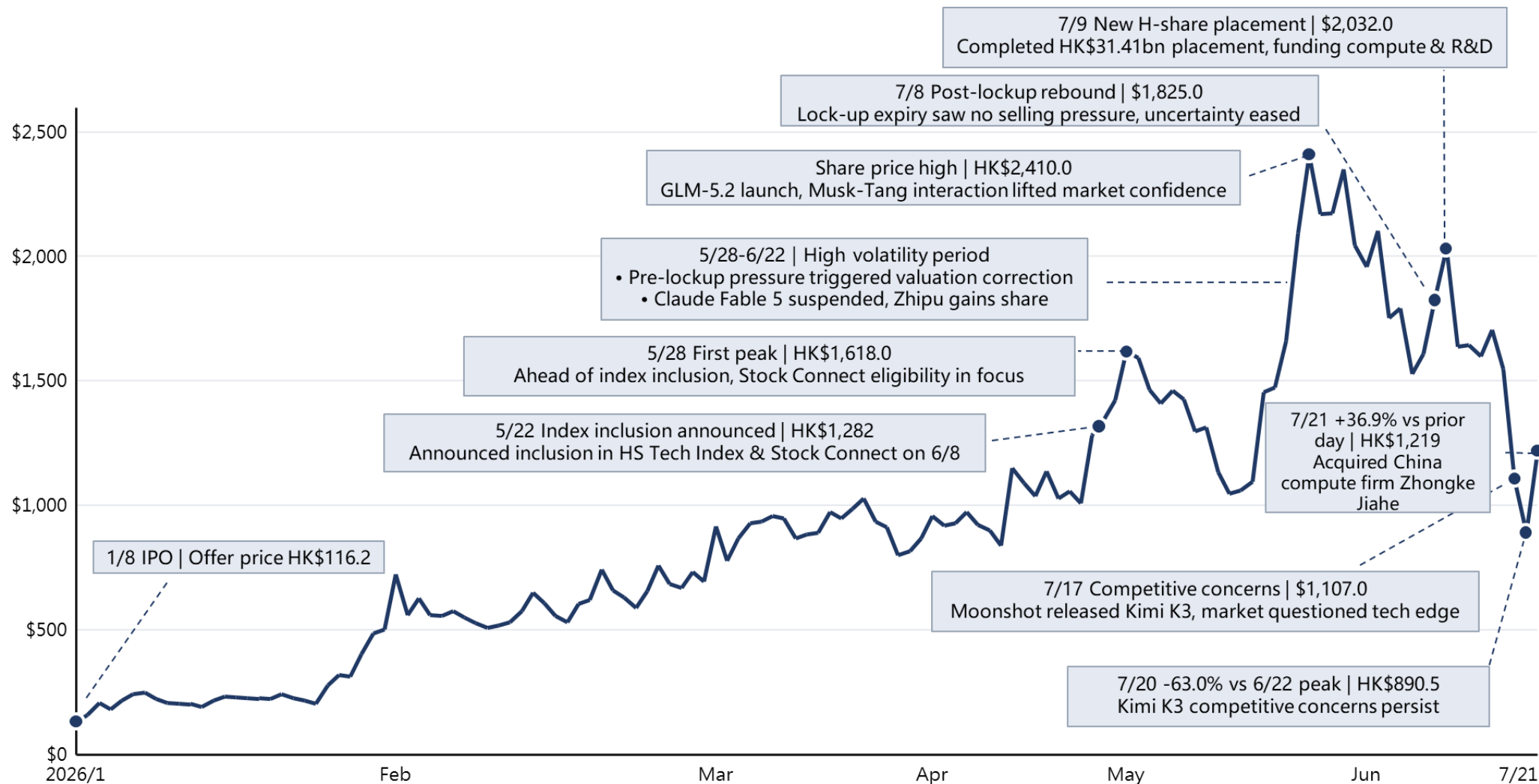
Analyst average	1,784	P/S	128x	1. Domestic AI-related companies (Minimax, Moore Threads, Cambricon, Metax, Biren Technology) currently trade on average at ~55x 2026 P/S or ~0.5x PSG.	52.2	39.5	30.7	24.6
	vs			2. Overseas AI model peers currently valued at: OpenAI 34.1x, Anthropic 20.5x 2026 P/ARR.				
	(IPO \$116.2)			3. Considering that AI model updates are a key risk to financial forecasts and share-price volatility, we recommend a conservative pricing level of ~ 0.4x PSG . (2025–2027 CAGR: 333%)				
	(Placement \$1,588)							

Sell-side institutions (Post GLM 5.2 announcement)		Target price	Scenario analysis	2026E		2027E	
				On-premise deployment	Cloud deployment	On-premise deployment	Cloud deployment
				(RMB mn)	(RMB mn)	(US\$ 100m)	(RMB 100m)
Haitong International Research Ltd		2,300	Bull Case (HK\$2,300/shr)	1,175	6,350	2,045	15,422
UBS		2,200					
Macquarie		2,157					
CLSA		2,061	Base Case (HK\$1,784/shr)	1,007	4,500	1,753	11,796
SDIC Securities		2,043					
CCB International Securities Ltd		2,000					
HSBC		1,900	Bear Case (HK\$1,250/shr)	801	3,100	1,394	8,099
Goldman Sachs		1,880					
Average		1,784		534	190		
JP Morgan		1,600	Implied CAGR in revenue (2025–2027)				
86Research Inc.		1,474	On-premise deployment		Cloud deployment		
CICC		1,250	Bull Case (HK\$2,300/shr)	96%	801%		
Aletheia Capital Limited (Sell)		545					
Buy							
Hold			Base Case (HK\$1,784/shr)	81%	688%		
Sell							
			Bear Case (HK\$1,250/shr)	62%	553%		

* By late 2025 through Q1 2026, AI coding and agentic engineering surged, with paid token consumption increasing exponentially

Factors Driving Share Price Volatility 1/3

Since IPO (Jan 2026): low/high HK\$116.1/HK\$2,980 (26x); largest one-day gain +43% / loss -28%.



Drivers of Share Price Volatility 2/3

Kimi K3 model launch triggered Zhipu AI shares to swing 20%–37% in a single day

2026	Market News	Share Price Impact
July 17	Moonshot AI launched its next-generation flagship model, Kimi K3, in the early hours. It topped the Code Arena coding test with a 76% win rate, and its overall performance approaches US frontier models such as GPT-5.6.	Amid intensified competition in the premium coding segment and concerns over competitive moats, Zhipu AI shares plunged 28.49% that day, erasing over HK\$200 billion in market capitalization. Sector peer MINIMAX-W also fell 15.63%.
July 20	(July 19) Moonshot AI circulated an IPO proposal to investors, expected to list on the HKEX within six months at the earliest, targeting a valuation up to US\$30 billion.	Under dual pressure from Moonshot AI’s IPO plans and K3’s technical competition, Zhipu AI shares continued downward, falling another 19.56%. Sector peer MINIMAX-W declined a further 10.6% that day.
July 21	Zhipu AI announced completion of the acquisition of AI compute-software firm Zhongke Jiahe to improve compilation efficiency, and the deployment of a 1GW hyperscale data center using exclusively domestic AI chips, effectively resolving market concerns about compute constraints.	Shares rallied strongly after the open, closing up 36.89%. Market cap rose by over HK\$150 billion in a single day, restoring total market capitalization to approximately HK\$567.6 billion, nearly recovering losses from the prior two panic-selling sessions.

DOW JONES
NEWSWIRES

Z.AI Shares Slump on Competitor's Model Release, Broader Tech Decline

2026新上市

当前价

涨跌幅

成交量

总市值

年初至今

智谱

02513 延时

1107.00

-441.00 (-28.49%)

1459.03万股

5154.45亿

+852.67%

MINIMAX-W

00100 延时

216.0

-40.0 (-15.63%)

842.06万股

754.35亿

+30.91%

Zhipu AI model release history – industry rankings steadily rising

Release Date	Zhipu AI model releases	OpenRouter ranking data	GLM API Pricing (per 1M input tokens)	Pricing comparison of close competitors (per 1M input tokens)	Relative Advantages	Relative Disadvantages
16 Jan 2024 29 Aug 2024	GLM-4 GLM 4-Plus	Chinese/Bilingual ranking: Top 10 Early general conversation ranking: #24	\$0.50	Qwen 2: \$0.30 GPT-4o: \$2.50 Claude 3.5: \$3.00	Versus GPT-4o: delivers comparable Chinese local semantic understanding while reducing API costs by roughly 80%.	On complex mathematical/logical reasoning (STEM) and advanced code generation, peak capability still lags GPT-4o.
28 Jul 2025 15 Sep 2025	GLM-4.5 GLM-4.6	Global overall ranking: #14 Agentic/Code category ranking: #8	\$0.60	DeepSeek V3: \$0.14 Qwen 2.5: \$0.30 Claude 3.5 Haiku: \$0.80	Versus Claude 3.5 Haiku: offers more stable logical computation on long-chain reasoning tasks and a lower per-output cost.	The 'Thinking' mode increases time-to-first-token (TTFT), causing additional output token consumption.
12 Mar 2026 18 Apr 2026	GLM-5 GLM-5V	Multimodal Agent category: #6 Web/UI Code ranking: #9	\$1.20	DeepSeek V4 Flash: \$0.09 Gemini 3 Flash: \$0.50 GPT-5.4: \$2.50	Versus DeepSeek V4 Flash: context capacity is 8× larger, significantly improving whole-project codebase and long-document retrieval quality.	Multimodal visual token processing exhibits higher latency; dynamic video-stream understanding is slightly inferior to Gemini.
9 Jun 2026	GLM-5.2	Open-weights (Open Source) model overall ranking: #1 Global model overall ranking: #8 Code/Agent specialty ranking: #2	\$1.40	DeepSeek V4 Pro: \$0.21 Kimi K3: \$3.00 GPT-5.5: \$5.00	Project-level software refactoring (FrontierSWE score 74.4) nears top closed-source models ; adopts Open-weights (Open Source) under MIT license and supports enterprise on-premise private deployments. Versus Claude Opus 4.8: comparable agentic coding capability to leading US models while reducing API costs by ~80%.	The peak code score (74.4) still slightly trails Claude Opus 4.8 (75.1) and the newly released Kimi K3. Versus DeepSeek V4 Pro: API price is roughly 6–7× higher ; the 744B large MoE architecture also imposes very high GPU memory requirements for enterprise on-premise deployment.

Drivers of Share Price Volatility 3/3

Recommend participating in the growth of China's large AI model sector via structured products

What we like:

- **Zhipu AI demonstrates strong, sustained R&D capability:**
Zhipu AI models are steadily closing the gap with top U.S. models — GLM5.2 scores at 85% of leading models while costing only 17% of their cost
- **'On-premises + Cloud' dual-engine growth model:**
On-premises deployments serve institutions that can build data centers and prioritize data security; the cloud is highly scalable and is the company's largest growth driver
- **Service procurement first, equipment leasing as a supplement:**
The company reduces capital expenditure accordingly while aligning with national policy to leverage China's compute network
- **Benefiting from increased compute supply:** As procurement scales up, training and cloud compute costs should continue to decline, improving margins

What we dislike:

- **Concentrated revenue stream:** Both on-premises models and cloud services are heavily dependent on the large model's capabilities and pricing (single business line)
- **Intense competition and weak user stickiness:** The cloud services / Model-as-a-Service (MaaS) sector is highly competitive and users can switch to alternative models
- **Open weights increase revenue-leakage risk:** Releasing model weights expands the potential user base, but various compute platforms can run Zhipu AI models without relying on Zhipu AI's cloud services
- **Geopolitical risk:** The U.S. government could restrict American firms from using non-U.S. models; the company's shares could be subject to sanctions lists
- **Loss of scarcity premium as the only listed large-model company:** Potential IPOs by other major model developers would increase comparable listed peers and could compress valuation premiums

High Level of Volatility:

Ticker	Name	Mktcap (US\$ mn)	Historical Volatility (%)	
			30-day	90-day
2513 HK	Zhipu AI CO LTD	73,631	240.72	187.62
100 HK	MINIMAX GROUP	8,596	172.12	142.10
000660 KS	SK HYNIX INC	867,725	125.64	107.80
SPCX US	SPACE X	1,555,944	114.72	N/A
MU US	MICRON TECH	1,118,336	111.73	99.09
TSM US	TAIWAN SEMI	2,155,395	55.98	48.48
META US	META PLATFORMS	1,538,538	52.13	46.39
AVGO US	BROADCOM INC	1,867,208	48.58	52.64
GOOGL US	ALPHABET INC	3,889,213	40.36	38.04
NVDA US	NVIDIA CORP	5,051,992	37.26	39.82
MSFT US	MICROSOFT CORP	2,834,542	35.63	33.30
AMZN US	AMAZON.COM INC	2,513,506	34.80	32.05
AAPL US	APPLE INC	4,724,335	33.43	26.95

Conclusion:

Recommend participating in the growth of China's large AI model industry via structured products

Appendix: Zhipu AI company development history

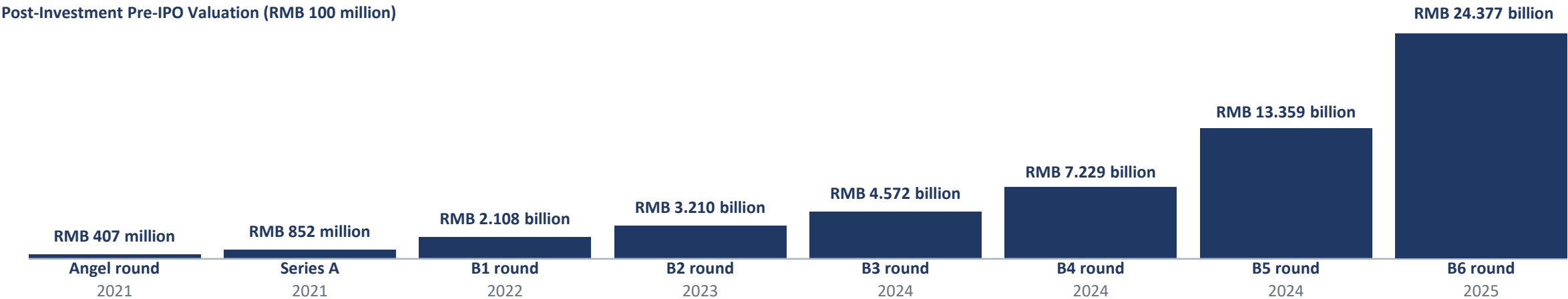
Seven-year dual-track path of proprietary foundation models and commercialization

Year	Key Events	Strategic Significance
2019	Technology spin-out from Tsinghua KEG Laboratory; received RMB 40 million angel investment from China Science & Technology VC	Academic-industry benchmark: Tsinghua affiliation established technical talent base; angel-round valuation RMB 375 million
2020–21	Key decision: independently develop the GLM architecture, not based on GPT/BERT paths	Technical moat: self-developed core architecture avoids international patent and licensing risks, establishing an independent technical route
2021	Launched GLM-10B and Model-as-a-Service (MaaS) platform	Commercial validation: among China’s earliest to trial a MaaS business model, initiating revenue generation
2022	Open Source GLM-130B (only Asian model selected in Stanford evaluation); released CodeGeeX	Technical validation: first to Open Source a core model, building a developer ecosystem and entering the code-generation track
2023	Open Sourcing ChatGLM and launching Zhipu Qingyan; raised over RMB 2.5 billion	Consumer penetration: initial products completed regulatory filings, creating compliance barriers; strategic investments from Alibaba, Tencent, Xiaomi, etc.
2024	GLM-4 series and AutoGLM; CodeArena ranked No.1 globally	Agent strategy: autonomous AI agents targeting the next-generation direction led by OpenAI
2025	International brand Z.ai; GLM-4.6; compatible with 40+ domestic (China-made) chips	Domestic substitution: chip self-sufficiency reduces supply-chain risk while supporting overseas expansion
2026.1	GLM-5 released and listed on the HKEX Main Board (02513.HK)	First listed large-model company: raised ~HKD 4.3 billion, initiating public market financing

Appendix: Zhipu financing history

8 Pre-IPO rounds totaling RMB 8.364 billion; Jan–Jul 2026 IPO and placements net proceeds ~HK\$36.27 billion

Post-Investment Pre-IPO Valuation (RMB 100 million)



Round / Event	Date	Amount / Valuation	Key investors
Angel round	2021.4	20.37 million RMB (IP valuation)	Beijing Innovation Zhiyuan Technology
Series A	2021.7	152 million RMB	Dachen Chuanghong; Beijing Huakong; Rongpin Investment; Jiangmen Venture Capital; Lingyun Optics; etc.
B1 round	2022.2	208 million RMB	Junlian Xiangdao; Qiming Rongqian; Qiming Rongkai
B2 round	2023.6–8	410 million RMB	Tianjin Sankua (Meituan) exclusively led the first tranche; Junlian Xiangdao; Xinglian Zhaoji; etc. participated
B3 round	2023.11 – 2024.1	609 million RMB	Quande Meijia (Capital Today); Shanghai Yunya (Ant Group); Junlian Xiangdao; etc.
B4 round	2024.8	1.229 billion RMB	Tencent Investment; Social Security Zhongguancun Innovation Fund; Beijing Shunying (Lei Jun-affiliated); etc.
B5 round	2024.11–12	1.359 billion RMB	Saudi Prosperity7 (Aramco-affiliated); Beijing AI Industry Investment Fund; etc.
B6 round	2024.12–2025.5	4.377 billion RMB	Zhongguancun Science City; Daxing Industry Fund; Hangzhou Urban Investment Industrial Fund; Zhuhai Huafa; Lenovo Ventures; etc.
IPO listing	2026.1.8	Net proceeds HK\$4.896 billion (issue price HK\$116.2)	Public offering + cornerstone investors (global offering)
Placement of new H shares	2026.7.9	Net proceeds ~HK\$31.375 billion (placement price HK\$1,588)	International long-term funds; strategic/industry investors; professional institutions (China International Capital Corporation — sole placement agent)

Appendix: Zhipu AI core founding team

Tsinghua KEG Lab alumni: professors lead technology, engineers lead operations, PhDs lead strategy

Tang Jie

Founder, Chief Scientist

Age 49 | Shareholding: pre-IPO 6.66% → post-IPO 6.12%

Academic background

Professor, Department of Computer Science, Tsinghua University; Fellow of IEEE, ACM and AAAI. PhD, Tsinghua University (2002); ~24 years in AI, research in Semantic Web and data mining.

Key experience

Academic lead of Tsinghua KEG Lab; led development of AMiner (globally recognized academic intelligence platform); led development of China's first trillion-parameter model, Wudao 2.0.

Role at Zhipu AI

Chief technical architect — leads GLM framework development and the next-generation AGI roadmap. In July 2026 initiated the “Touch High” program to accelerate AGI efforts.

Zhang Peng

Co-founder, Chief Executive Officer

Age 47 | Shareholding: pre-IPO 0.10% → post-IPO 0.09%

Academic background

PhD in Innovation Leadership Engineering, Tsinghua University; core member of KEG Lab, fully involved in AMiner development.

Key experience

Responsible long-term for KEG Lab technology engineering and productization; led AMiner's transition from a lab project to an internationally recognized research platform.

Role at Zhipu AI

Oversees daily operations, commercialization and team execution. Has stated “we had revenue from day one”; serves as a pragmatic bridge between technology and market.

Liu Debing

Co-founder, Chairman

Age 50 | Shareholding: pre-IPO personal 0.23% + Beijing Lianpai 8.45% → post-IPO 7.98%

Academic background

PhD, Institute of Computing Technology, Chinese Academy of Sciences; mentored by Academician Gao Wen; ~18 years in computing technology.

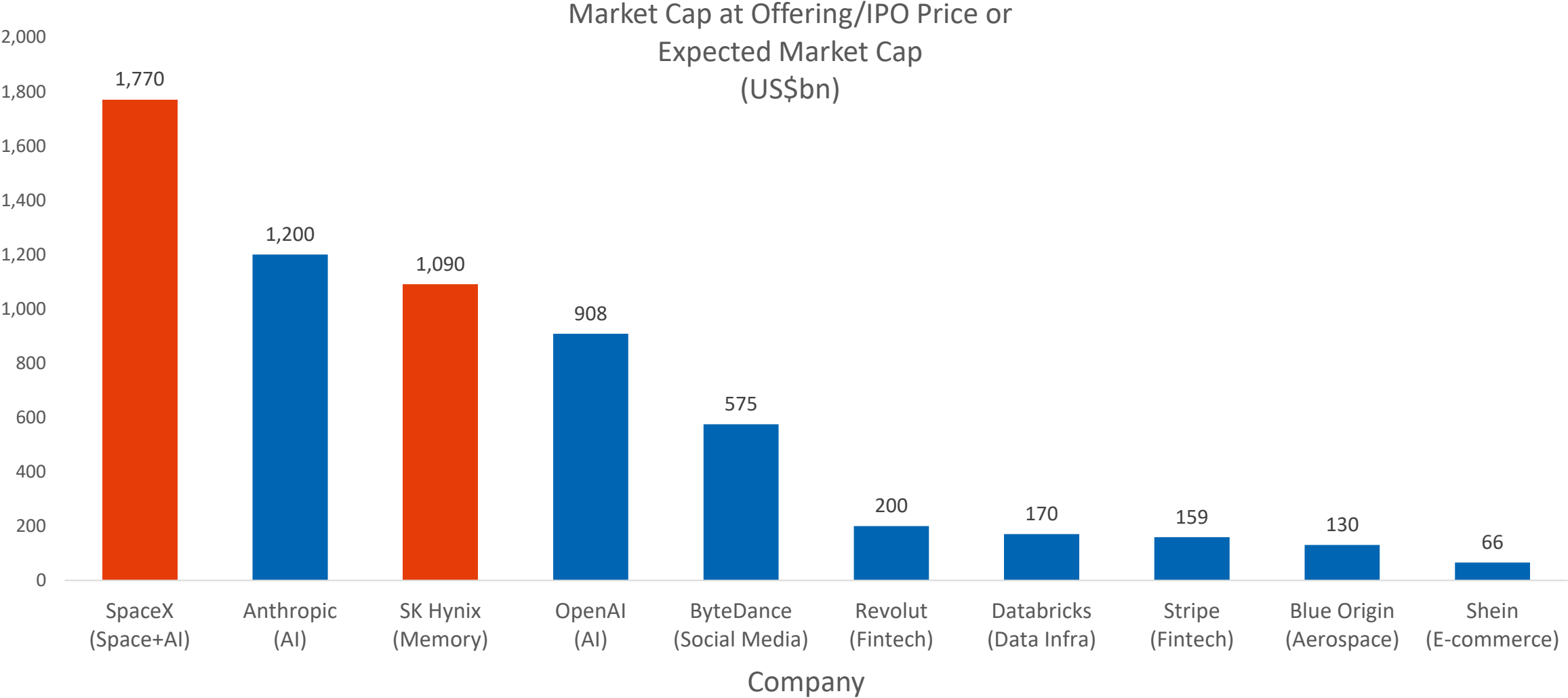
Key experience

Former research engineer at Teyi (China) Technology; later senior engineer at Tsinghua University with deep collaboration with KEG Lab; focused on computer vision and multimedia analysis.

Role at Zhipu AI

Responsible for top-level strategy, board governance and major capital operations; led multiple financing rounds and listing strategy from founding to IPO.

Appendix:
Upcoming Mega IPOs



Sources: Bloomberg, Company News, KGI

Q & A

投資產品與服務部



尹賀
Kevin Yin
Equity Advisory
kevin.h.yin@kgi.com
+852 2878 6798
SFC CE: AJM689



莫偉民
Raymond Mok, CFA
Equity Advisory
raymond.mok@kgi.com
+852 2878 4663
SFC CE: BHJ465



葉晉宜
Derek Yip, CFA
Equity Advisory
derek.yip@kgi.com
+852 2878 4263
SFC CE: BSQ196



梁啟棠
Cusson Leung, CFA
CIO
cusson.kt.leung@kgi.com
+852 2878 6998
SFC CE: ADU276

Disclaimer and Important Notice

Hong Kong: This document has been prepared by KGI Securities (Asia) Limited ("KGI"). The information contained herein is confidential and is not intended for public distribution or for use by any person or entity located in or resident in any jurisdiction that restricts KGI or its affiliates from providing such information. This document and its contents must not be forwarded to any other person. The information does not constitute investment advice to any person or entity in any jurisdiction, nor an offer, invitation or solicitation to subscribe for or purchase any securities, insurance or other investment products or services. In particular, the information in this document must not be distributed in the United States, to U.S. persons (i.e., U.S. residents or entities organized under the laws of the United States or any state, territory or possession thereof), or for the benefit of U.S. persons, nor used for the purpose of offering or soliciting any securities in the United States. All information is provided for general information and reference only and does not take into account any investor's individual objectives, financial situation or needs. No part of this document may be reproduced, circulated or published without KGI's prior written consent. The information is not intended to constitute legal, financial, tax or other professional advice and should not be relied upon as such. All investments involve risk; securities prices may be highly volatile and can rise or fall and may even become valueless. There is no assurance that trading in securities will be profitable; investors may incur losses. You should exercise caution, conduct your own independent review and obtain independent professional advice before making any investment decision. Investors should carefully consider whether an investment is suitable in light of their own risk tolerance, financial circumstances, investment experience, objectives, time horizon and knowledge. KGI makes no representations or warranties, express or implied, as to the accuracy, adequacy or completeness of the information provided. Any person relying on or using the information herein should independently verify its accuracy, completeness, reliability and suitability. Simulated, historical or forecast performance is not necessarily indicative of future results. Information provided in this document (including any data) may not have been independently verified and should not be relied upon as the sole basis for any investment decision. Neither KGI, its affiliates nor any of their directors, officers, employees or representatives accept any liability for any loss or damage (direct, indirect or consequential) arising from any omission, error, inaccuracy, incompleteness or from reliance on the information provided. KGI, its affiliates and their respective directors, officers, employees and representatives accept no responsibility for information provided by or attributed to third parties.

Risks of Trading Outside the Hong Kong Market: Trading on markets or exchanges in other jurisdictions may expose you to additional risks. Regulatory protections in such markets may differ from or be weaker than those provided in Hong Kong. Before trading on such markets or exchanges you should ask about and understand the rules applicable to your specific transactions and be aware that you may be subject to tax implications, currency risk and additional transaction costs. Investor assets received or held outside Hong Kong are governed by the laws and regulations of the relevant foreign jurisdiction, which may differ from the Securities and Futures Ordinance (Cap. 571) and its rules in Hong Kong; accordingly, such assets may not enjoy the same protections as assets received or held in Hong Kong. You should only consider trading outside the Hong Kong market when you fully understand the nature of the foreign market and the risks involved. Consider carefully whether such trading is suitable for you in light of your experience, risk profile and other relevant circumstances, and seek independent professional advice if you are in any doubt. Ensure you are familiar with and understand the trading rules of overseas markets, including any restrictions on holdings and disclosure obligations. Members of the KGI Group or their affiliates may provide services to any companies referred to in this document and their affiliates. KGI Group members, their affiliates and their directors, officers and employees may from time to time have interests in the securities referred to above.

Complex Products: "Complex products" are investment products whose terms, features and risks are sufficiently complex that they cannot reasonably be expected to be understood by retail investors. Investors should exercise caution when considering complex products. Investors may lose their entire investment or incur losses greater than their initial investment (where applicable). For complex products offered on the basis of offering documents or materials that have not been reviewed by the Hong Kong Securities and Futures Commission (the "SFC"), investors should exercise particular caution regarding such offerings. Where a product is described as having SFC recognition, such recognition does not constitute an official endorsement; SFC recognition is not equivalent to a recommendation or approval of the product, nor does it guarantee the commercial merits or performance of the product. Past performance, if provided, is not indicative of future results. Some complex products are available only to professional investors. Before making any investment decision, investors should read the offering documents and other relevant materials to understand the nature, features and risks of the complex product, obtain independent professional advice, and ensure they have sufficient net assets to bear the risks and potential losses of trading such products.

Bonds are not equivalent to time deposits and are not protected by the Hong Kong Deposit Protection Scheme. Bondholders are exposed to various risks, including but not limited to: (i) credit risk: interest and principal payments on a bond are payable by the issuer. If the issuer defaults, bondholders may not recover interest and/or principal; bondholders bear the issuer's credit risk. Credit ratings assigned by rating agencies do not guarantee the creditworthiness of the issuer; (ii) liquidity risk: secondary markets for certain bonds may be illiquid, and it may take time to sell bonds prior to maturity; (iii) interest rate risk: when interest rates rise, the prices of fixed-rate bonds typically fall. If you seek to sell a bond before maturity, the proceeds received may be less than the purchase price. You should not invest in bonds unless you fully understand and are willing to assume the associated risks.

Investments involve risk; prices of securities and fund units may rise or fall, and past performance is not indicative of future results. Before making any investment decision, investors should read the relevant fund offering documents in detail (including the full description of risk factors, particularly for funds investing in emerging markets). Funds are subject to normal market volatility and other risks inherent in the underlying assets. Changes in political, financial, economic, social and/or legal conditions—factors outside the control of the fund manager—may prevent the fund manager from achieving the fund's investment objectives despite their best efforts. Therefore, you may not recover your original investment and could lose part or all of your initial capital. Past performance of a fund should not be taken as a guide to future performance. There is no guarantee of capital return or any return. Some funds are intended as medium- to long-term investments. Specific investments made by a fund may not be able to be liquidated or hedged promptly or at reasonable prices due to market depth, lack of demand or market disruption. Investments held by a fund may become illiquid, particularly during periods of market turmoil or economic uncertainty, which may affect the fund's ability to meet your redemption requests. A fund's investments may be concentrated in particular industries, instruments or countries, making the fund's value more volatile; such concentration entails the specific risks of those industries or countries. Equity funds are subject to general equity market risk; equity prices may rise or fall. Factors affecting equity prices include, but are not limited to, changes in local and global investor sentiment, political environment, economic conditions, business and social developments, and issuer-specific factors. For passively managed funds (i.e., index funds), the fund manager generally does not have discretion to adjust holdings in response to market movements; accordingly, index funds may be affected by declines in market segments related to their reference index. Index funds are subject to tracking error risk, meaning their performance may not precisely follow the index; there is no assurance that an index will be replicated accurately or completely at any time.