

Initiation of Coverage

Knowledge Atlas Technology

Aiming to become China's Anthropic; initiate at Buy

Strategically following a successful and proven path

We initiate coverage of Knowledge Atlas Technology (better known as Zhipu or Z.ai) with a Buy rating and a PT of HK\$1,160. We are encouraged by Zhipu's solid track record in advancing model intelligence, especially in coding capability, which could lead to a commercialization breakthrough as enterprise AI demand reaches an inflection as of early 2026. We think its strategy resembles Anthropic's, with a consistent focus on frontier AI coding capability that has led to rapid model iteration, strong user mindshare and faster-than-expected monetisation. Compared with the global leader, Zhipu has had an even steeper ARR ramp-up, albeit off a smaller base, surging over 6x to US\$250m over four months from Dec 2025 to Mar 2026 (vs. Anthropic's similar ramp over nine months in 2025). Despite its premium valuation (145x 2026 consensus P/S), we believe Zhipu is well positioned to benefit from the secular headwind of surging token demand, with early signs of pricing power, which could bring further growth upside.

Strong R&D capability to support further share gains in an expanding market

We are positive on Zhipu's significant market potential, as improving model intelligence could expand use cases from developer-centric coding to broader knowledge workers and enterprises with increasingly varied and complex tasks. We are also positive on Zhipu's potential to sustain leadership in frontier model development, supported by its early-mover advantage in coding, forward-looking model upgrade roadmap, and strong R&D team with a proven academic track record. Its latest release, GLM-5.1, is among the few Chinese models closing in on Anthropic in long-horizon tasks, a core capability to execute multi-step, real-world coding workflows with high value.

Robust revenue growth with a long runway

We forecast a robust revenue CAGR 231% in 2025-27E, driven by significant open platform scaling and continued rapid growth in on-premise deployment. We expect open platform to increase its revenue contribution from 26% in 2025 to 78% in 2027E, driven by rising token consumption and Zhipu's premium pricing. That said, its revenue growth may be non-linear, affected by model launch cadence and computing power availability, along with competition, especially improvement of the in-house model capability of internet leaders (several are Zhipu's current KA customers), which is a major downside risk. We also expect upward pricing and improved inference efficiency to support gross margin expansion and gradual net loss ratio narrowing over time.

Valuation: Initiate at Buy and a price target of HK\$1,160

Our PT assumes 145x 2026E P/S, on a 231% 2025-27E revenue CAGR and 0.6x 2026E PSG, cross-checked by SOTP analysis where we assume 100x P/ARR on open platform's ARR of US\$600m in Dec 2026E and 60x P/S on 2026E on-premise revenue.

Equities

China
Internet Services

12-month rating **Buy**

Prior : No Rating

12m price target **HK\$1,160.00**

Prior :

Price (17 Apr 2026) **HK\$891.00**

RIC: 2513.HK **BBG:** 2513 HK

Trading data and key metrics

52-wk range	HK\$957.50-131.50
Market cap.	HK\$359b/US\$45.8b
Shares o/s	403m (ORD)
Free float	20%
Avg. daily volume ('000)	2,842
Avg. daily value (m)	HK\$1,555.9
Common s/h equity (12/26E)	Rmb1.30b
P/BV (12/26E)	>100
Net debt to EBITDA (12/26E)	0.1x

EPS (UBS, diluted) (Rmb)

	From	To	% ch	Cons.
12/26E	-	(11.23)	-	(10.03)
12/27E	-	(10.29)	-	(9.95)
12/28E	-	(8.55)	-	(6.89)

Wei Xiong

Analyst

wei.xiong@ubs.com

+86-21-3866 8883

Kenneth Fong

Analyst

kenneth-kc.fong@ubs.com

+852-3712 3890

Charles Chen

Associate

S1460524020001

charles-za.chen@ubs.com

+86-21-3866 8907

Highlights (Rmbm)	12/23	12/24	12/25	12/26E	12/27E	12/28E	12/29E	12/30E
Revenues	125	312	724	3,208	7,941	13,574	21,787	33,943
EBIT (UBS)	(459)	(2,046)	(2,251)	(5,084)	(4,757)	(3,914)	349	4,196
Net earnings (UBS)	(621)	(2,466)	(3,182)	(5,157)	(4,747)	(3,963)	261	4,080
EPS (UBS, diluted) (Rmb)	(41.12)	(154.22)	(7.90)	(11.23)	(10.29)	(8.55)	0.56	8.71
DPS (net) (Rmb)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Net (debt) / cash	979	1,329	1,025	297	(5,250)	(10,508)	(12,561)	(12,038)
Profitability/valuation	12/23	12/24	12/25	12/26E	12/27E	12/28E	12/29E	12/30E
EBIT (UBS) margin %	(368.4)	<-500	(310.7)	(158.5)	(59.9)	(28.8)	1.6	12.4
ROIC (EBIT) %	37.3	53.4	29.9	114.7	(478.8)	(195.5)	9.3	63.1
EV/EBITDA (UBS core) x	-	-	-	NM	NM	NM	>100	43.8
P/E (UBS, diluted) x	-	-	-	(69.4)	(75.4)	(90.8)	NM	89.1
Equity FCF (UBS) yield %	-	-	-	(1.4)	(1.4)	(1.1)	0.3	1.6
Dividend yield (net) %	-	-	-	0.0	0.0	0.0	0.0	0.0

Source: Company accounts, LSEG Eikon, UBS estimates. Metrics marked as (UBS) have had analyst adjustments applied. Valuations: based on an average share price that year, (E): based on a share price of HK\$ 891.00 on 17-Apr-2026

UBS Research THESIS MAP a guide to our thinking and what’s where in this report

PIVOTAL QUESTIONS

Q: Can Zhipu stay at the forefront globally in terms of coding and agentic capability?
Likely. We think model intelligence is the key competitive moat for model providers, which serves as the core differentiating factor and could expand the boundary of commercialization. We highlight the resemblance of Zhipu’s model development and monetization strategy to that of global leader Anthropic, which is highly focused on coding and agentic capability and has had accelerating demand from enterprises and developers with proven monetization potential. While industry competition remains intense, we are constructive on Zhipu’s early-mover advantage, strong R&D team and experience, and continued technological innovation to keep developing frontier models with improving capability and rapid iteration.

Q: What is Zhipu’s commercialization potential in open platform and on-premise deployment?
We forecast a 231% revenue CAGR in 2025-27, with the open platform portion rising from 26% in 2025 to 78% in 2027. We see significant monetization potential for China’s LLM API services, driven by robust enterprise AI demand, while improving model capability supports user growth and token consumption, providing meaningful upside for Zhipu’s open platform business. We also expect superior model intelligence to drive demand and pricing power, supporting volume and price growth in API services, although tight computing power availability could constrain upside. Meanwhile, we expect on-premise deployment to continue growing, supported by penetration among government and enterprises, though likely a lower priority.

UBSVIEW

We initiate on Zhipu at Buy. We are encouraged by its solid track record in advancing model intelligence, particularly in coding capability, which could drive strong revenue ramp-up with a long runway as enterprise AI demand accelerates. We note its strategy resembles Anthropic’s, and Zhipu has had an even steeper ARR ramp-up, albeit from a smaller base. We believe Zhipu is well positioned to benefit from secular growth in token demand with early signs of pricing power.

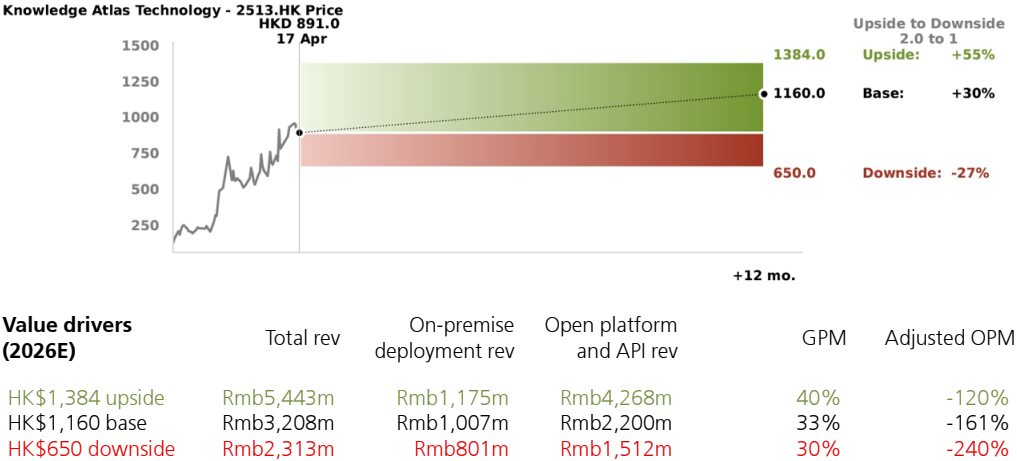
EVIDENCE

1) As per Artificial Analysis, GLM-5.1 ranks seventh globally in overall model intelligence (as of 17 April 2026). 2) GLM-5.1 achieved about an eight-hour task completion time horizon (human equivalent), ranking first among global open-source models on METR. 3) Zhipu’s API pricing has increased by 83% YTD (as of Mar 2026), including a 30% direct price hike on its Coding Plan. 4) Following that price adjustment, Zhipu’s total token usage surged ~3x MoM in March 2026, according to OpenRouter data. 5) Zhipu’s ARR has surged 6.4x to US\$250m in March 2026 from December 2025, higher than Anthropic’s growth rate (3.3x from December 2025 to April 2026).

WHAT’S PRICED IN?

Given the rapidly evolving technological and competitive landscape in the global foundation model industry, we believe investors lack conviction of Zhipu’s ability to sustain model leadership or its revenue growth trajectory. Also, we believe the market has yet to fully price in the monetisation potential of LLM API services and the broadening use cases enabled by improving model capability, which could bring meaningful revenue upside for Zhipu and help digest its premium valuation.

UPSIDE/DOWNSIDE SPECTRUM



Source: UBS estimates

COMPANY DESCRIPTION

Zhipu is a China-based developer of general-purpose large models pursuing customised on-premise solutions and cloud-based services.



PIVOTAL QUESTIONS

Q: Can Zhipu stay at the forefront globally in terms of coding and agentic capability?

UBSVIEW

Likely. We think model intelligence is the key competitive moat for model providers, which serves as the core differentiating factor and could expand the boundary of commercialization. We highlight the resemblance of Zhipu's model development and monetization strategy to that of global leader Anthropic, which is highly focused on coding and agentic capability and has had accelerating demand from enterprises and developers with proven monetization potential. While industry competition remains intense, we are constructive on Zhipu's early-mover advantage, strong R&D team and experience, and continued technological innovation to keep developing frontier models with improving capability and rapid iteration.

EVIDENCE

1) Based on our analysis of the strategy of global leader Anthropic, **on the model layer**, it places a strong focus on coding capability and emphasises long-horizon tasks, where Claude Opus 4.6 ranks first globally in the task completion time horizon (~12 hours of human-equivalent task duration), according to METR. **On the engineering layer**, it has introduced innovations such as MCP, Skills/CoWorker and Claude Code (an AI coding agent), enhancing coding capability at the system level.

2) Zhipu has a similar strategic focus, with a strong emphasis on coding capability. **On the model layer**, its latest model, GLM-5.1, achieves about an eight-hour task completion time horizon (human equivalent), ranking first among global open-source models on METR, while closing the gap with leading closed-source models (eg, Claude Opus 4.6). **On the engineering layer**, it launched its first coding copilot product, CodeGeeX, in 2022, ahead of domestic peers.

3) Zhipu has a **strong R&D team** with deep roots at Tsinghua University. Its R&D personnel exceed 800, with 20+ papers published on arXiv (a primary preprint platform for AI/LLM research), exceeding leading domestic peers. **Key innovations** in its model development include Dynamic Sparse Attention (reducing training and inference costs while maintaining long-context fidelity), "slime" synchronous agent RL (enhancing post-training efficiency) and native agent-integrated design (agent + inference + coding).

4) On model iteration, Zhipu pursues **five paradigms for evolving coding capability** (AI coding assistant, vibe coding, agentic engineering, long-horizon tasks and autonomous OS). Over the past 10 months, Zhipu has launched six model upgrades (from GLM-4.5 to GLM-5.1), advancing its capability from AI coding copilot to solving long-horizon tasks.

5) **In third-party model performance rankings**, Zhipu's latest frontier model, GLM-5.1, ranks seventh globally in overall model intelligence, according to Artificial Analysis. It is first among all open-source models in agentic coding tasks involving multi-step reasoning and tool use, according to Code Arena.

WHAT'S PRICED IN?

We believe the market is sceptical that any model provider, including Zhipu, can maintain leadership in model capability, given the highly competitive and dynamic nature of LLM development, along with a shift in other companies' model development focus toward coding. Investors are also concerned about underlying computing power constraints, which may limit Chinese model providers' ability and pace to pursue model development.

Adopting a similar model development strategy as Anthropic

As a leading AI foundation model provider in China, Zhipu is dedicated to developing general-purpose large models with a continuous pursuit of frontier model intelligence. Founded in 2019, the company's goal is to explore the boundaries of AGI while serving China's industrial needs, and its strategy aims to enable AI to perform "more realistic and complex tasks" through further scaling. Zhipu's GLM framework was China's first proprietary pre-trained large model framework, according to the company, while its model portfolio covers multiple domains based on the GLM framework.

Figure 1: Zhipu's major model portfolio

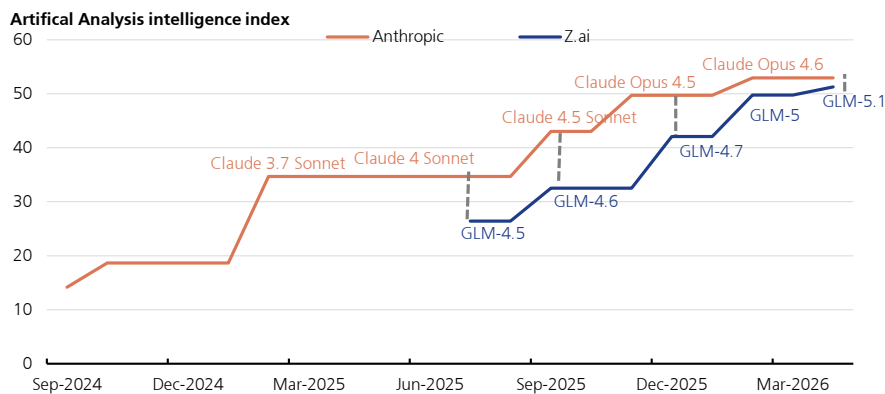
Model type	Models	Launched date	Key features
Language model	GLM-5.1	8-Apr-2026	Flagship text LLM (latest generation), with further enhanced coding and long-horizon task capabilities
	GLM-5-Turbo	16-Mar-2026	OpenClaw-optimised foundation model
	GLM-5.0	12-Feb-2026	Previously known as "PonyAlpha". Flagship text LLM with notable improvements in coding capability.
Image & video model	GLM-Image	14-Jan-2026	Text-to-image generation model
	GLM-4.6V	9-Dec-2025	Vision-language model
	CogVideoX-3	6-Aug-2025	Text-to-video generation model
Audio model	GLM-TTS	11-Dec-2025	Text-to-speech model

Source: Company data, UBS

We think Zhipu's model development strategy is closely aligned with that of global AI leader Anthropic:

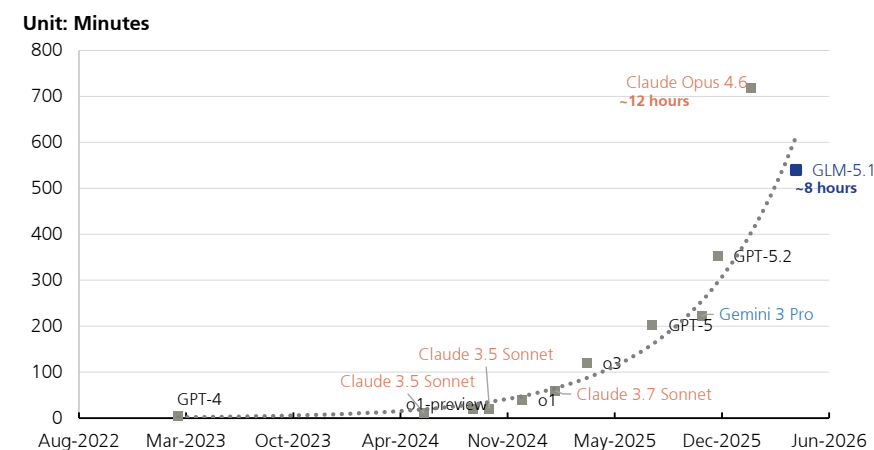
- **A specialized focus on coding capability at the model layer:** Both Zhipu and Anthropic have a clear focus on coding as the primary model capability frontier, aiming to achieve best-in-class performance for coding and developer workflows. Their models have consistently delivered top-tier performance in coding benchmarks. With continued model iteration, there is a narrowing gap between Zhipu and Anthropic's latest models (51.28 for GLM-5.1 vs. 52.95 for Opus 4.6, as measured by the Artificial Analysis index, or 58.4 vs. 57.3, respectively, as measured by SWE-bench Pro). Both are relatively less focused on multi-modality (ie, audio and video generation).
- **Within AI coding, emphasis on long-horizon task-solving capability:** We believe this is critical for real-world software development, which involves iterative, multi-step workflows requiring reliable execution of long, closed-loop tasks rather than single-turn generation. According to METR, Anthropic's Opus 4.6 achieves a ~12-hour task completion time horizon (50% success rate), while Zhipu's latest frontier model, GLM-5.1, reaches about eight hours, ranking first among global open-source models and ahead of most domestic peers.
- **Further enhanced agentic capability through engineering layer innovation:** Both Anthropic and Zhipu optimise their models for task completion through multi-step reasoning and tool use, enabling the translation of model intelligence into task execution. Anthropic has introduced system-level innovations such as MCP, Skills/CoWorker and Claude Code to enhance agentic workflows. Similarly, Zhipu launched its first coding copilot product, CodeGeeX, in 2022, ahead of domestic peers. Notably, Zhipu also leverages advantages in coding to develop more general-purpose AI agent products, including AutoGLM and AutoClaw.

Figure 2: Zhipu's model intelligence is improving, with a narrowing gap to Anthropic



Source: Artificial Analysis, UBS

Figure 3: Anthropic's Claude Opus 4.6 leads in long-horizon coding tasks globally (~12 hours), while GLM-5.1 reaches ~8 hours



Source: Model Evaluation & Threat Research, company data. Note: Time horizon measures the length of tasks (human-equivalent duration) that AI agents can complete with a 50% success rate.

Strong R&D capability deeply rooted in leading academic institution

We believe Zhipu's **close connections with a leading academic institution** (ie, Tsinghua University) serve as a key foundation for its R&D capability. All five of Zhipu's founders are Tsinghua alumni, and the university's asset management arm holds a 3.53% stake in the company, reflecting its strong academic linkage. Zhipu also maintains a close partnership with the Knowledge Engineering Group at Tsinghua University, enabling systematic technological exchange.

It is further supported by its **sizeable R&D team**. We estimate Zhipu's R&D headcount exceeds 800, higher than other domestic emerging AI labs.

With a **proven academic track record**, Zhipu has released 20+ papers on arXiv (a primary preprint platform for AI/LLM research), ahead of domestic peers. As of 30 June 2025, Zhipu's research and academic advisory teams published approximately 500 highly influential papers, demonstrating its strong R&D capability.

We believe Zhipu's continued model iteration and capability improvement are backed by **solid innovations across multiple dimensions in AI model development**, supporting Zhipu's pursuit of the upper bound of model intelligence, including: **1) Model architecture:** Dynamic Sparse Attention, a DeepSeek-like attention mechanism applied in GLM-5 training, reducing training and inference costs while maintaining long-context fidelity. **2) Training methodology:** "Slime" asynchronous reinforcement

learning infrastructure, enhancing post-training efficiency. **3) Agent framework design:** Native integration of agent, inference and coding capability (the ARC framework), improving agentic performance in complex, long-horizon tasks.

Figure 4: Comparison of R&D teams and publications among leading Chinese AI labs

	Z.ai	Moonshot AI	MiniMax
R&D as % of total employees	74%	-	74%
Number of total employees	1094	~300	428
Number of R&D employees	~800	~200	~300
Number of papers on arXiv platform	>20	>8	>5

Source: Company data, arXiv, UBS estimates. Note: Zhipu and MiniMax's employee numbers as of 31 Dec 2025.

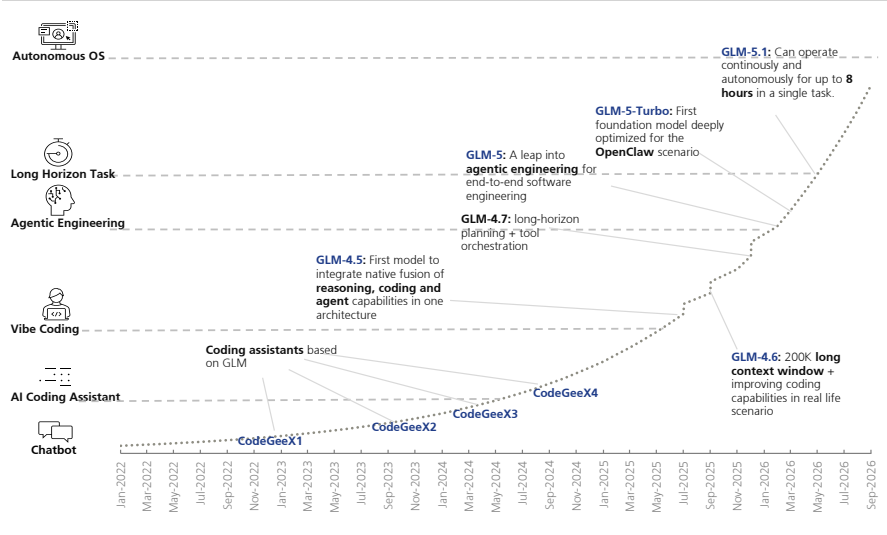
Continued model iteration with notable intelligence improvement

In model iteration, **Zhipu pursues five paradigms for evolving coding capability** (AI coding assistant, vibe coding, agentic engineering, long-horizon tasks and autonomous OS). Specifically:

- 1. As early as 2022, Zhipu launched its **AI coding assistant** product series, CodeGeeX, based on the GLM model, serving as a copilot tool to assist developers with basic coding tasks (eg, code completion and annotation).
- 2. GLM-4.5 integrates reasoning, coding and agent capability, marking a paradigm shift from copilot to **vibe coding**.
- 3. Subsequently, GLM versions 4.6-5.0 further improved in areas such as long context and planning, advancing toward the **agentic engineering** paradigm.
- 4. GLM-5.1 represents further progression to the **long-horizon task** paradigm, achieving about an eight-hour task completion time horizon (human equivalent).

Over the past 10 months, the company has launched six model upgrades (from GLM-4.5 to GLM-5.1), advancing its capability from AI coding toward long-horizon task solving. This represents acceleration compared with the upgrade pace of the CodeGeeX series during 2022-24.

Figure 5: Zhipu's model iteration is in line with the five paradigms



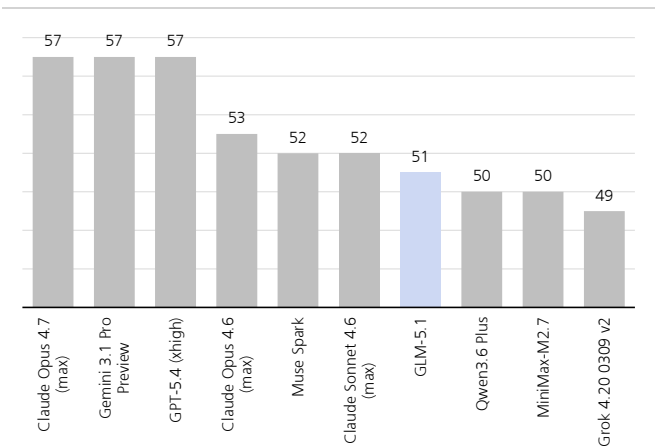
Source: Company data, UBS

Measured by external benchmarks, Zhipu has demonstrated strong model performance, particularly in AI coding and agentic capability. Zhipu's past model upgrades have a solid track record, with each iteration delivering meaningful model intelligence improvement. We note a material step change especially after the launch of GLM-5.0 (also known as PonyAlpha), which received [positive feedback in the AI community](#) for its exceptional coding capability, logical reasoning and agentic workflow.

On 7 April 2026, Zhipu further upgraded its flagship model to GLM-5.1, advancing the upper bound of its model performance. Regarding to GLM-5.1:

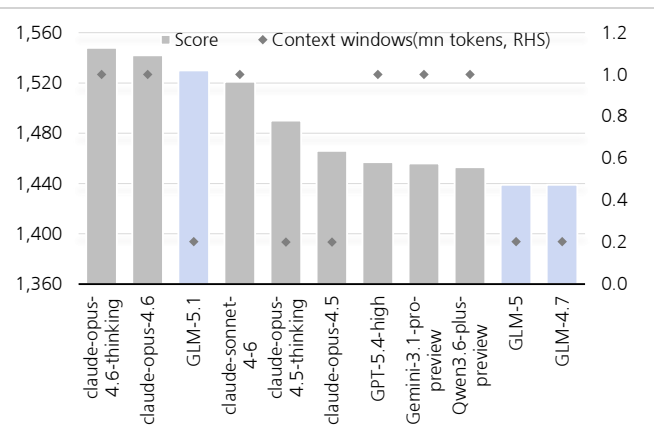
- In the overall intelligence index, it ranked **seventh**, according to Artificial Analysis.
- In complex agentic coding tasks involving multi-step reasoning and tool use, it ranked **third** among all frontier models, according to Code Arena.
- In addition, its score on agentic coding and terminal-use capability reached 42%, ranking **sixth** globally, according to Artificial Analysis.
- On SWE-bench Pro, it achieved a score of 58.4, ranking **second** globally, suggesting its strong real-world coding capability.

Figure 6: In overall model intelligence, GLM-5.1 ranked seventh among all frontier models



Source: Artificial Analysis. Note: As of 17 April 2026, the Artificial Analysis Intelligence Index v4.0 includes: GDPval-AA, ²-Bench Telecom, Terminal-Bench Hard, SciCode, AA-LCR, AA-Omniscience, IFBench, Humanity's Last Exam, GPQA Diamond and CritPt.

Figure 7: In coding capability, GLM-5.1 ranked third among all frontier models



Source: Code Arena. Note: As of 10 April 2026 a view of overall rankings across AI models in terms of agentic coding tasks, involving multi-step reasoning and tool use.

Figure 8: In Agentic terminal use cases, GLM-5.1 delivers top-tier performance globally

Rank	Company	Model	Terminal-Bench Hard
1	OpenAI	GPT-5.4 (xhigh)	58%
2	Google	Gemini 3.1 Pro Preview	54%
3	Anthropic	Claude Sonnet 4.6 (max)	53%
4	OpenAI	GPT-5.4 mini(xhigh)	52%
5	Anthropic	Claude Opus 4.6 (max)	46%
6	Z.ai	GLM-5.1	42%
7	xAI	Grok 4.20 Beta 0309	41%
8	Xiaomi	MiMo-V2-Pro	41%
9	Alibaba	Qwen3.5 397B A17B	41%
10	MiniMax	MiniMax-M2.7	39%

Source: Artificial Analysis. Note: as of 8 April 2026, Terminal-Bench Hard evaluates a model's terminal operation, autonomous debugging and long-horizon sustained execution capability via challenging multi-step system and engineering tasks in Linux terminal environments,

Figure 9: In agentic engineering capability, GLM-5.1 ranked second among all frontier models

Rank	Company	Model	SWE-Bench-Pro
1	Anthropic	Claude Mythos Preview	77.80
2	Z.ai	GLM-5.1	58.40
3	OpenAI	GPT-5.4	57.70
4	OpenAI	GPT-5.3 Codex	56.80
5	Alibaba	Qwen 3.6 Plus Preview	56.60
6	MiniMax	MiniMax-M2.7	56.20
7	OpenAI	GPT-5.2	55.60
8	MiniMax	MiniMax-M2.5	55.40
9	OpenAI	GPT-5.4 mini	54.40
10	Google	Gemini 3.1 Pro Preview	54.20

Source: DataLearnerAI. Note: As of 8 April 2026, requiring models to locate and fix highly complex engineering bugs in real Github repositories, representing the closest simulation to real-world software development.

Well positioned against competition, with early mover advantages

While the industry is converging toward advancing coding and agentic capability among model providers, we are constructive on Zhipu's early-mover advantages in this area. As one of the earliest Chinese AI labs to launch an AI coding assistant product, CodeGeeX (Sep 2022), Zhipu has maintained a consistent strategic focus on coding, supported by continued model iterations. Since GLM-5, the company has further emphasised long-horizon task solving, which is increasingly viewed as an important dimension of coding capability. Model performance improvement in this area

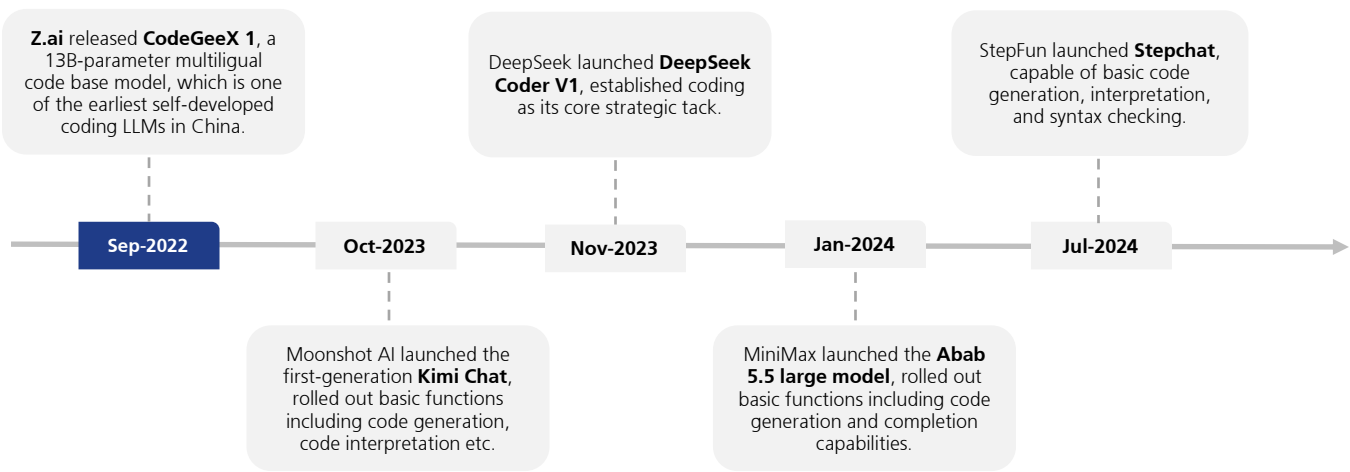
has been reflected in external benchmarks. Our industry checks also indicate positive enterprise feedback and strong user mindshare for GLM models as among the top choices for coding-related tasks.

Based on research papers released by Zhipu, we think one of its core technological differentiators lies in its asynchronous RL framework (“slime”), which reconstructs the training paradigm and enhances stability and scalability in long-horizon agent tasks (particularly coding tasks) and large-scale MoE models.

Despite the increasing focus on coding across the industry, we also note **the model development strategies and areas of specialisation among leading AI labs remain somewhat differentiated:**

- **Moonshot AI’s** Kimi series focuses on multi-agent coordination and complex task decomposition.
- **MiniMax** emphasises full-stack multimodal capability, alongside strong inference efficiency and cost control.
- **DeepSeek** focuses on mathematical reasoning, logical inference and long-context processing.
- **StepFun** has a relative advantage in lightweight on-device deployment and integration with domestic computing infrastructure.

Figure 10: Timeline of model launches focused on coding capability among China's AI labs



Source: Company data, UBS

PIVOTAL QUESTIONS

Q: What is Zhipu's commercialization potential in open platform and on-premise deployment?

UBSVIEW

We forecast a 231% revenue CAGR in 2025-27, with the open platform portion rising from 26% in 2025 to 78% in 2027. We see significant monetization potential for China's LLM API services, driven by robust enterprise AI demand, while improving model capability supports user growth and token consumption, providing meaningful upside for Zhipu's open platform business. We also expect superior model intelligence to drive demand and pricing power, supporting volume and price growth in API services, although tight computing power availability could constrain upside. Meanwhile, we expect on-premise deployment to continue growing, supported by penetration among government and enterprises.

EVIDENCE

1) Based on government tendering data in the public service sector, Zhipu secured 57 LLM projects with a disclosed contract value of Rmb254m in 2025, ranked fifth among peers. The gross margin for on-premise deployment was 54%/68%/66%/49% between 2022-25, while days of sales outstanding (DSO) increased from 107 in 2024 to 153 in 2025.

2) Zhipu's model has well established collaborations with leading internet companies (with nine out of China's top 10 as key customers). It has also cooperated with leading domain players such as Didi, indicating it has strong capability to reach the enterprise software market through coding. Zhipu's ARR has surged 6.4x to US\$250m in March 2026 from December 2025, higher than Anthropic's growth rate (3.3x from December 2025 to April 2026).

3) Continued model iteration could drive a pricing uplift. The release of GLM-5.0 led to significantly increased input/output pricing of >100% vs. GLM-4.7. Subsequently, GLM-5-Turbo and GLM-5.1 were each priced ~20% higher than the previous version. Zhipu also implemented direct price adjustments, announcing a 30% price hike for its coding plan in February 2026.

4) Following the price adjustments, Zhipu's total token usage still surged ~ 3x MoM in March 2026, according to OpenRouter data. Meanwhile, although GLM-5.0 was the most expensive model in its portfolio at that time, it accounted for the highest share of usage in March 2026.

5) For frontier models, Zhipu's GLM-5.1 and GLM-5.0 achieve over 90% of Anthropic's performance metrics, while priced at only 20-30% of Anthropic's levels. At the entry level, GLM-4.7 slightly outperforms Claude 4.5 Haiku on model intelligence, while priced at roughly half of the latter.

WHAT'S PRICED IN?

We believe investors lack conviction about the growth upside potential that Zhipu can achieve in its open platform and API services, despite favourable industry trends, given intense competition and tight computing power capacity. Investors are also concerned about the sustainability of Zhipu's premium model pricing and that its models may commoditize over time, with price competition as an increasing risk.

On-premise deployment as revenue anchor

On-premise deployment historically has accounted for the majority of Zhipu's total revenue (74% in 2025). This business primarily involves localised deployment (ie, on customers' own infrastructure) of LLMs alongside necessary on-site services (eg, model training, fine-tuning and related implementation support), targeting large enterprises (to-B) and government clients (to-G).

The business generally charges on a project or annual basis, and its pricing may vary by model type, computing resource requirements and implementation complexity. Industries such as finance and legal tend to have stringent data security requirements and require extensive fine-tuning using proprietary data and domain-specific knowledge, bringing demand for on-premise solutions.

We believe Zhipu is well positioned in China's to-G/to-B market, underpinned by the following:

- **Established access to local governments and large enterprises**, backed by its solid shareholder profile, including multiple local government-backed investment platforms (eg, Hangzhou Urban Construction Investment Group), as well as leading internet companies such as Meituan and Ant Group. We believe this provides Zhipu with access to public sector and enterprise project opportunities.
- **Zhipu's technological stack and localisation efforts**, which are aligned with demand from certain central SOEs and state-owned enterprises for technological self-reliance and controllability. For example, GLM-5 is fully adapted and optimised for China's GPU ecosystems across seven mainstream domestic chip providers, spanning from underlying kernels to upper-layer inference frameworks.

Figure 11: Zhipu's flagship on-premise deployment projects with strategic shareholders

Project	Counterparty	Date	Equity stake in Z.ai - after IPO (%)	Key highlights
To-G projects				
AI Industry Large Model Project (Phase I)	Hangzhou Urban Construction Investment Group	Mar-2025	1%	The first full-scenario localized deployment benchmark in China (contract value 61.3 mn RMB)
City-level GLM Large Model Project	Zhuhai Huafa Group	Mar-2025	2%	The first city-level large model on-premise deployment benchmark in the Guangdong-Hongkong-Macao Greater Bay area
Zhuge Large Model National Headquarters Project	Chengdu Hi-Tech Zone	Mar-2025	1.13% (via Hi-Tech Orinno)	The first Provincial-level base large model R&D Center
Zhongguancun (multiple projects)	Zhongguancun Science City	Sep-2024	2%	Strategic partnership
Moli Community - Zhipu Large Model MaaS Platform Project	Shanghai Pudong Venture Capital Group, Zhangjiang Group etc.	Jun-2025	1.88% (via Zhihui Linghang)	A city-level on-premise deployment benchmark in Pudong, launches the MaaS Platform.
Projects for large enterprises				
Strategic partnership	Meituan(Tianjin Sankuai Technology)	Jul-2023	4%	
	Shanghai Feiya/ Shanghai Yunya	Aug-2023	4%	
	Legend Capital	Sep-2022	6%	

Source: Company data, UBS

Based on government tendering data in the public service sector, Zhipu has reached the scale typically seen among large internet incumbents. In 2025, the company secured 57 LLM projects, with a disclosed contract value of Rmb254m. Companies reaching a similar scale are primarily established players such as Alibaba Cloud, Tencent Cloud and Volcano Engine. Notably, Zhipu's projects span a broad range of industries, including energy, government affairs, telecom and education, with the Hangzhou Urban Investment Group's industrial LLM project (Rmb62m) representative and among the largest.

Zhipu has also begun to gain traction in overseas markets. The company is working with the Malaysian government to build a national sovereign AI platform based on its foundation models, and it plans to replicate similar projects across Southeast Asia through partnerships with governments, universities and local technology ecosystems. Revenue from Southeast Asia on-premise deployment projects reached Rmb17.9m as of 30 June 2025, increasing significantly from Rmb1.1m in 2024.

Figure 12: Zhipu fifth among domestic LLM vendors in public sector project bidding (2025)

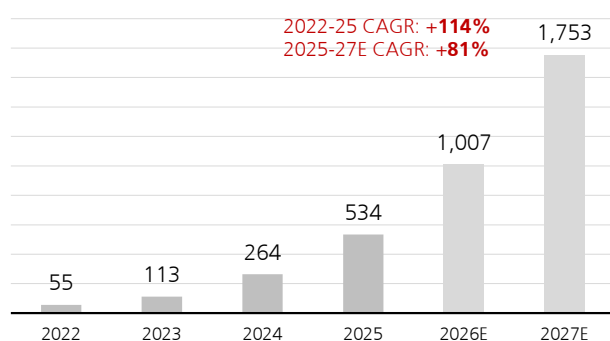
Rank	Vendors	Number of projects	Total contracts value (RMB million)
1	iFlytek	210	2,316
2	Baidu	110	890
3	Volcano Engine	83	518
4	Alibaba Cloud	69	402
5	Z.ai	57	254
6	Tencent Cloud	44	123

Source: China Tendering & Bidding Public Service Platform. Notes: Ranking based on awarded project count; total contract value is secondary. Some contracts are undisclosed.

We expect on-premise revenue to maintain a solid growth trajectory. Historically, the segment has grown rapidly, with revenue increasing from Rmb55m in 2022 to Rmb534m in 2025 (~9.7x over three years), and it is primarily concentrated in the internet and technology, public service, and telecommunications sectors, which combined accounted for 81.3% of total revenue as of 1H25. We forecast an 81% CAGR in on-premise revenue in 2025-27, driven by continued customer expansion alongside improving AI penetration, as well as higher revenue per customer, with AI increasingly integrated into productivity-related use cases.

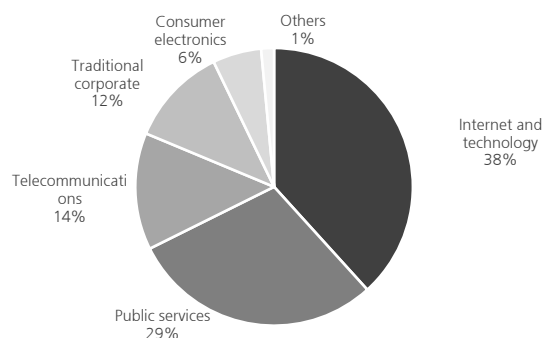
Figure 13: We forecast an 81% CAGR for on-premise revenue in 2025-27

Unit: RMB m



Source: Company data, UBS

Figure 14: On-premise deployment revenue by industry (1H25)

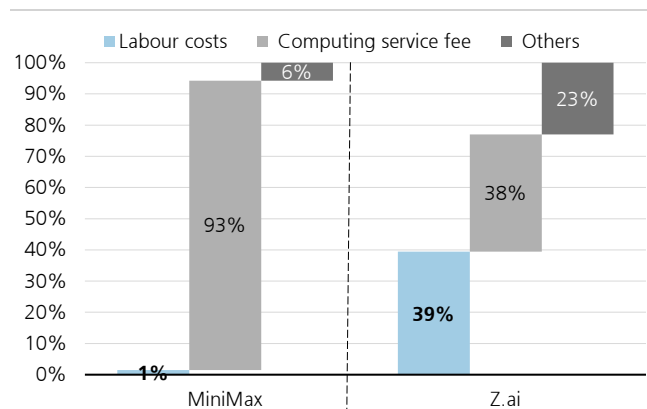


Source: Company data, UBS

Although on-premise deployment involves heavy labour cost, the gross margin for this segment has been healthy. On-premise deployment for enterprises and government is generally accompanied by in-depth customization requirements, including fine-tuning, integration with an enterprise's internal database, etc., requiring dedicated on-site support staff. As a result, payroll accounted for 39% of Zhipu's total cost of revenue in 1H25, mainly related to on-premise deployment. Nevertheless, the segment's gross margin remained healthy, at roughly 50-70%, in 2022-25 (a relatively lower 49% in 2025), suggesting enterprise customers can offer price premiums for Zhipu's quality model capability and services, in our view.

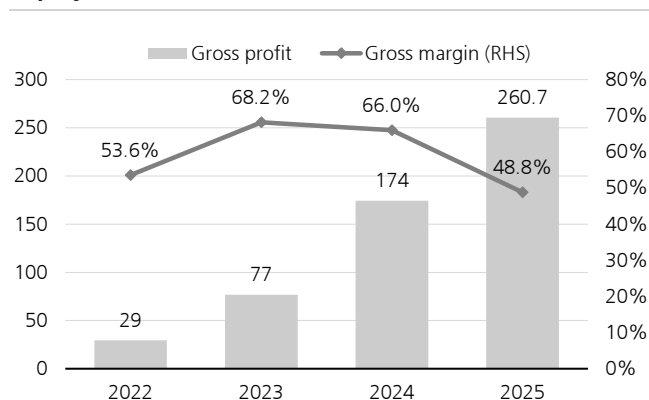
We note a moderation in Zhipu's cash collection efficiency, with DSO rising to 153 in 2025 (vs. 107 in 2024), alongside an increase in trade receivables from Rmb100m in 2024 to Rmb339m in 2025, which we view as acceptable at the current stage of commercialization, though worth monitoring. Zhipu's credit loss coverage ratio (i.e., total loss allowance/trade receivables) was 10.6% in 2025, broadly stable compared with the previous two years.

Figure 15: COGS breakdown – Zhipu vs. MiniMax



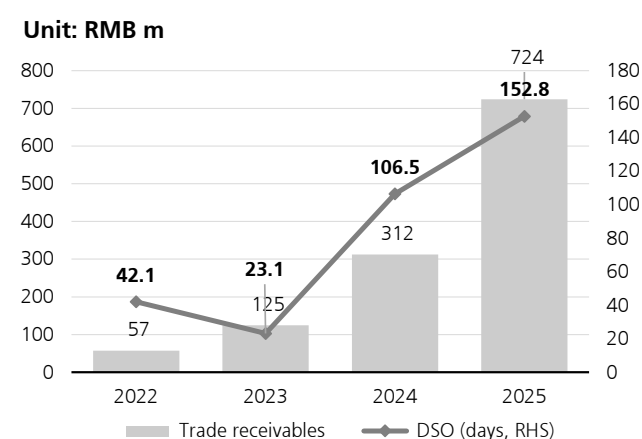
Source: Company data, UBS. Note: MiniMax data as of Sept 2025. Zhipu data as of June 2025.

Figure 16: Zhipu maintains healthy GPM for on-premise deployment business



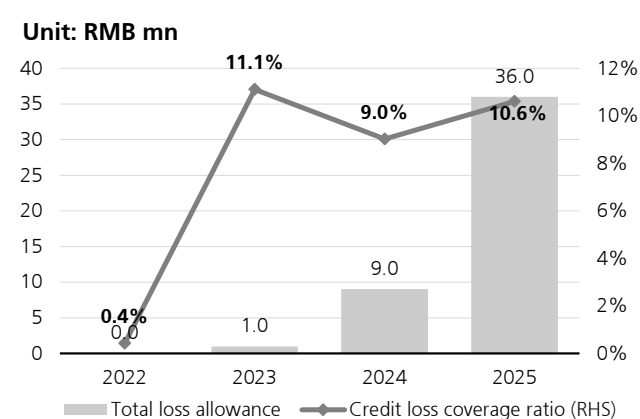
Source: Company data, UBS

Figure 17: Zhipu's DSO trend during 2022-25



Source: Company data, UBS

Figure 18: Zhipu's credit loss coverage ratio remained stable



Source: Company data, UBS

Unlocking scalable revenue upside from open platform and API business

We are more positive on the growth trajectory and revenue potential of open platform and API services (ie, cloud-based deployment), which are hosted on a scalable cloud infrastructure sensible for businesses seeking agility and ease of implementation. Revenue is recognized over the contract term ratably for subscription-based contracts and based on a customer's volume of tokens consumed for usage-based contracts.

Upon the release of its API product matrix, Zhipu's open platform business has ramped up significantly. Besides basic API services, Zhipu launched a coding plan on 1 September 2025 to address surging demand for high-quality AI coding capability at reasonable prices, along with a claw plan on 16 March 2026 (concurrent with GLM-5-Turbo), amid rising demand for emerging agentic frameworks (eg, OpenClaw), while enhancing security protections in enterprise use cases.

According to the company, paid token consumption surged **15x** over the past six months from October 2025 to March 2026. The number of users subscribed to the claw plan reached 100,000/400,000 within two/20 days post launch. As of March 2026, total users on its open platform reached 4m.

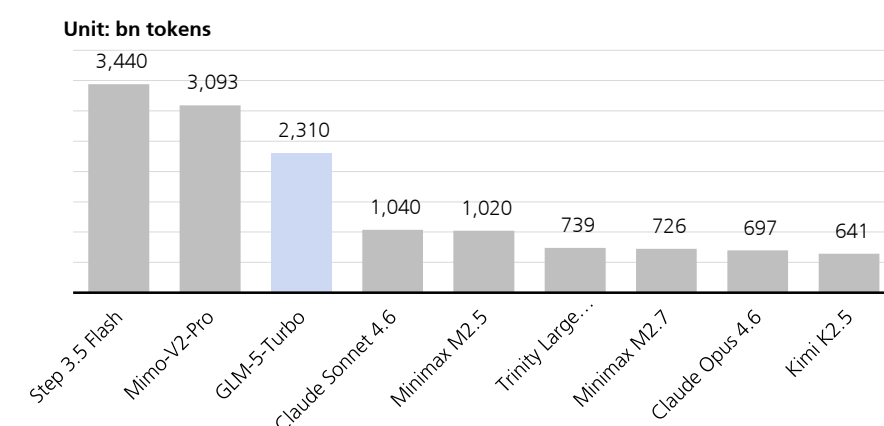
Figure 19: Zhipu's coding plan and claw plan pricing and usage

	Plan	Pricing (Rmb/month)	Usage	Key highlights
Coding Plan	Lite	49	1x base usage	- 4mn Open Platform users (as of Mar 2026); - Paid token consumption surged by 15x over the past 6 months
	Pro	149	5x lite usage	
	Max	469	20x lite usage	
Claw Plan	Basic	39	35mn tokens for GLM-5-Turbo	Post launch: - 100k new users in 2 days; - 400k new users in 20 days
	Pro	99	100mn tokens for GLM-5-Turbo	

Source: Company data, UBS

The recent rising popularity of OpenClaw and agentic workflows has further boosted API usage for the industry, and we view Zhipu as a notable beneficiary of this trend. The fast adoption of OpenClaw is driving rapid expansion of the overall AI model API market globally. In China, [internet leaders are integrating OpenClaw in their ecosystems](#), with most providing access to Zhipu's GLM model. Moreover, to capture this trend, Zhipu launched AutoClaw to lower the deployment barrier for OpenClaw. The recently launched GLM-5-Turbo model also provides further optimization to OpenClaw use cases with agent-native design. According to OpenRouter, GLM-5-Turbo ranked third in terms of tokens used by OpenClaw as of 2 April 2026.

Figure 20: Model token usage by OpenClaw over past-one-month period (as of 2 April 2026)



Source: OpenRouter, UB

Beyond robust volume growth, **we also note meaningful increases in Zhipu's API pricing**, underpinned by:

- 1. Continued model iteration driving a pricing uplift.** Zhipu implemented a **major upgrade** of its GLM model series in 2026, transitioning from the 4.x to the 5.x series. For example, the initial release of GLM-5.0 had significantly increased input/output pricing of **>100%** vs. GLM-4.7. Subsequently, Zhipu introduced **incremental upgrades** within the 5.x series, including GLM-5-Turbo and GLM-5.1, with each iteration priced **~20%** higher than the previous version.
- 2. Direct price hikes.** In addition to pricing increases through model upgrades, Zhipu implemented direct price adjustments. For example, in Feb 2026, the company raised the price of its coding plan **30%**, while introducing tighter constraints on usage and model selection.

Despite these price increases, the higher pricing did not discourage model usage, suggesting resilience in user demand and willingness to pay for improved model intelligence:

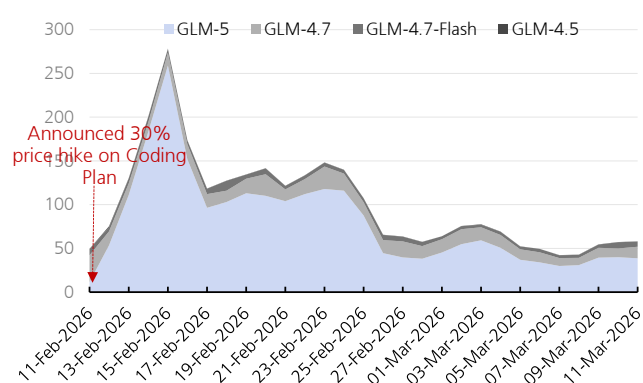
- Following the price adjustments, Zhipu's total token usage increased by approximately 3x MoM in March 2026, according to OpenRouter data.
- Also, although GLM-5.0 was the most expensive model in its portfolio at that time, it accounted for the highest share of usage in March 2026.

Figure 21: API price trend of Z.ai key model upgrades

Model	Launch date	Input price (RMB/m tokens)	Growth vs. last version	Output price (RMB/m tokens)	Growth vs. last version
GLM-4.5-Air	Jul/2025	0.8		2.0	
GLM-4.7	Jul/2025	2.0	▲ 150%	8.0	▲ 300%
GLM-5	Feb/2026	4.0	▲ 100%	18.0	▲ 125%
GLM-5-Turbo	Mar/2026	5.0	▲ 25%	22.0	▲ 22%
GLM-5.1	Apr/2026	6.0	▲ 20%	24.0	▲ 9%

Source: Company data, UBS

Figure 22: GLM's token usage trend on OpenRouter



Source: OpenRouter, UBS

We believe model intelligence is the most critical factor in driving commercial upside. Models with higher intelligence levels, characterized by superior capability, high precision and robustness, and the ability to complete high-value tasks, should be able to open up new use cases and expand the addressable market ([see our note](#)). Enterprises and developers also exhibit higher willingness to pay with less price sensitivity for models that can execute real-world complex workflows, reducing the risk of price competition for advanced model providers.

Compared with Anthropic, Zhipu's models offer a higher cost-to-performance ratio, suggesting pricing upside potential over time. Anthropic's model portfolio has three pricing tiers, with Opus at the high end; Sonnet in the mid-range, with a balance of performance, speed and cost; and Haiku focusing on cost efficiency. Our comparison of their comparable model series indicates the following:

- **For frontier models**, Zhipu's GLM-5.1 and GLM-5.0 achieve **over 90%** of Anthropic's performance metrics, while priced at only **~20-30%** of Anthropic's levels.
- **At the entry level**, GLM-4.7 slightly **outperforms** Claude 4.5 Haiku on model intelligence, while priced at **roughly half** of Haiku.

That said, Anthropic has built a more advanced ecosystem beyond model capability, including products such as Claude Code, Skills and Coworker. These offerings enhance user stickiness and workflow integration, allowing Anthropic's models to remain competitive with strong pricing power.

Figure 23: API price comparison between Zhipu's and Anthropic's models with similar intelligence

Provider	Model	Release Date	AA Intelligence Index	Input Price (\$/M tokens)	Output Price (\$/M tokens)	Overall price (\$/M tokens)
Anthropic	Claude Opus 4.6 (Adaptive Reasoning, Max Effort)	Feb-2026	53	\$5.0	\$25.0	\$9.0
Zhipu	GLM-5.1 (Reasoning)	Apr-2026	51.4	\$1.4	\$4.4	\$2.0
Zhipu as % of Anthropic			97%	28%	18%	22%
Anthropic	Claude Sonnet 4.6 (Adaptive Reasoning, Max Effort)	Feb-2026	51.7	\$3.0	\$15.0	\$5.4
Zhipu	GLM-5 (Reasoning)	Feb-2026	49.8	\$1.0	\$3.2	\$1.4
Zhipu as % of Anthropic			96%	33%	21%	27%
Anthropic	Claude 4.5 Haiku (Reasoning)	Oct-2025	37.1	\$1.0	\$5.0	\$1.8
Zhipu	GLM-4.7 (Reasoning)	Dec-2025	42.1	\$0.6	\$2.2	\$0.9
Zhipu as % of Anthropic			113%	60%	44%	51%

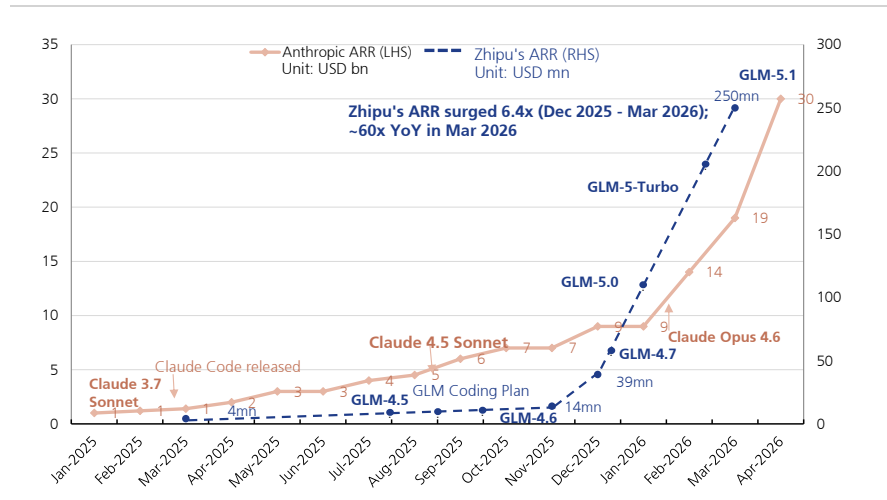
Source: Artificial Analysis, UBS

By examining Anthropic's ARR ramp-up trajectory, we see accelerating monetization potential as model intelligence reaches an inflection point:

- Upon the release of Claude 4.6 Sonnet and Opus 4.6, along with blockbuster agent product Claude Code, Anthropic's ARR increased from US\$9bn in December 2025 to US\$30bn in April 2026, with accelerated growth of over 3x over four months (vs. 6.4x over nine months previously).

- Zhipu has achieved even faster ARR growth, albeit off a much smaller base, up 60x YoY to reach US\$250m in March 2026 (or 6.4x vs. US\$39m in December 2025), due to its improved model intelligence and coding capability, aligning with the viral adoption of OpenClaw-like products.

Figure 24: AI model releases and ARR trends for Zhipu and Anthropic



Source: The Information, company data, UBS estimates. Note: The above models may not be exhaustive.

With coding as a proven use case with rapidly growing commercial value, we expect continued penetration of Zhipu's models into the China and overseas enterprise IT and software markets. As of March 2026, nine of the top 10 internet companies in China are Zhipu's paying customers, with a growing wallet share, an expanding user base and deepening product integration. Assuming all staff at the top 10 Chinese internet leaders are equipped with Zhipu's coding plan (Rmb375 per month for the max plan with an annual auto-renewal subscription) already suggests at least a Rmb4.6bn revenue opportunity annually. If we assume c5% of Chinese software developers' annual salaries are spent on AI coding, the addressable market for model providers could further grow to US\$14.2bn.

Additional revenue may come from leveraging Zhipu's models to support those companies' AI-powered features or AI native products, with a growing user base and usage requests (eg, GLM-5/5.1 to support Tencent's ima), or from expanding collaborations with vertical platforms (eg, announced cooperation with Didi in mobility assistants and corporate mobility agent co-design). That said, there is downside risk from the loss of KA customers, which are also leading internet companies, if their in-house model capability catches up.

Figure 25: Estimated revenue opportunity of coding plan sales among top 10 China internet leaders

Assume only penetrate into R&D employees					Assume penetrate into all employees	
Company	In-house coding agent product	Coding Plan cost (Rmb/year)	R&D employees (assume 50% of total)	Potential revenue (Rmb m)	Total employees	Potential revenue (Rmb m)
Bytedance	Trae	4,502	75,000	338	150,000	675
Tencent	CodeBuddy		57,925	261	115,849	522
Alibaba	Qoder		64,099	289	128,197	577
Meituan	CatPaw AI		54,450	245	108,900	490
PDD	-		11,500	52	23,000	104
BIDU	Comate		16,750	75	33,500	151
NetEase	CodeWave		13,014	59	26,028	117
Kuaishou	CodeFlicker		12,500	56	25,000	113
JD	JoyCode		90,000	405	180,000	810
Huawei	CodeArts Snap		110,000	495	220,000	991
Total			505,237	2,275	1,010,474	4,550

Source: Company data, UBS

Figure 26: Vibe coding TAM in terms of software developers' spending in US/China

	Number of Software developers	Coding Plan cost (USD/year)	Number of coding plans for each person	Total addressable market (USD bn)
US	4,400,000	2400	1	10.6
China	9,400,000	655	1	6.2

	Medium salaries(USD)- as of 2025	% of salaries for subscription	Annual subscription fee (USD)	Total addressable market (USD bn)
US	131,450	4.6%	6,000	26.4
China	33,000	4.6%	1,506	14.2

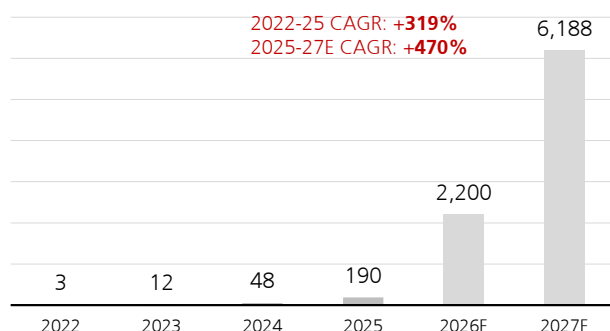
Source: Company data, JetBrains, Ministry of Industry and Information Technology

We expect Zhipu to deliver robust growth in 2025-27, driven by rapid improvements in model intelligence and expanding enterprise adoption. We forecast open platform revenue to have a 470% CAGR, from Rmb190m in 2025 to Rmb6,188m in 2027, supported by rising token consumption amid broader use cases, as well as higher pricing with model upgrades. As a result, we expect total revenue to have a 231% CAGR over the same period.

Open platform gross margin has improved recently. Amid industry-wide price competition in China's LLM sector, Zhipu reported an API gross margin of -0.4% in 1H25 but achieved a notable recovery in 2H25, supported by the launch of API offerings such as the coding plan and higher pricing from continued model upgrades, bringing the 2025 gross margin to 18.9%. Following this turnaround, we expect the gross margin to continue trending up, supported by further pricing upside and improving inference efficiency, although rising computing costs may create near-term pressure.

Figure 27: We forecast a 470% revenue CAGR for Zhipu's Open Platform and API business in 2025-27

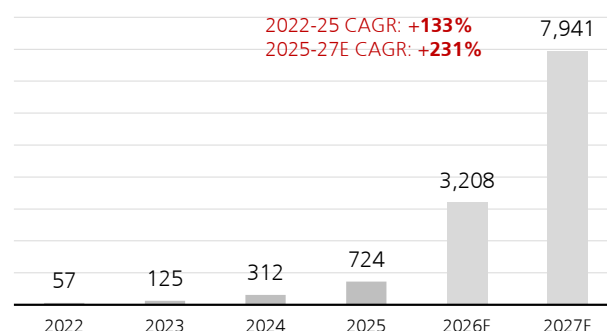
Unit: RMB m



Source: Company data, UBS estimates

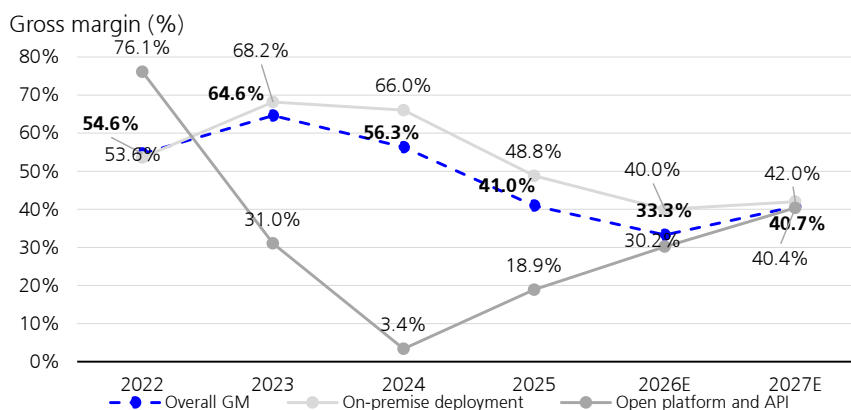
Figure 28: We forecast a 231% revenue CAGR for Zhipu's total revenue in 2025-27

Unit: RMB m



Source: Company data, UBS estimates

Figure 29: Zhipu's group and by segment GPM

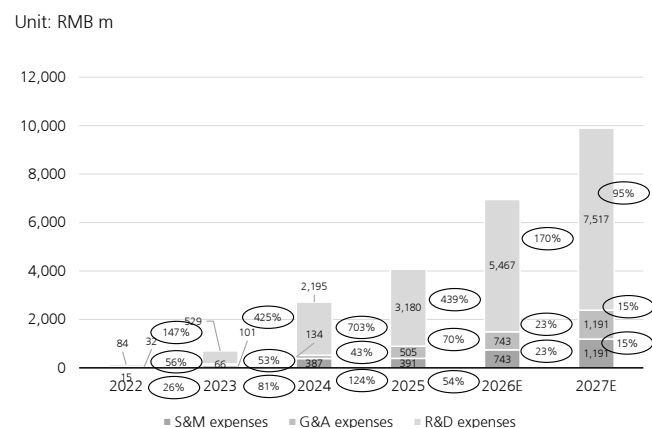


Source: Company data, UBS estimates

Net losses may narrow as revenue scales

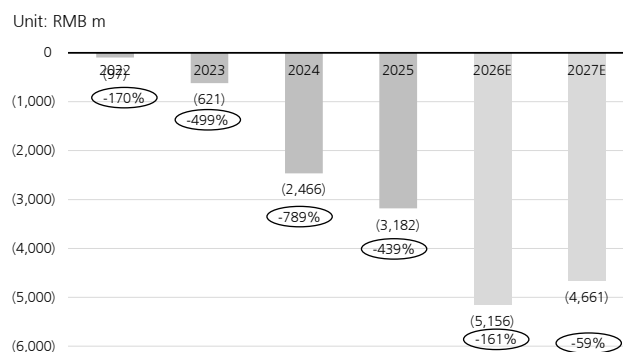
We expect significant losses over the next few years, as Zhipu continues to invest heavily in R&D while revenue takes time to scale. However, we expect narrowing losses from 2026 onward due to operating leverage. As of 31 Dec 2025, we estimate Zhipu had total cash liquidity of Rmb7.8bn, including cash and cash equivalents, short-term investments measured at FVPL and available unutilised bank facilities (Rmb5.2bn as at 31 Dec 2025, which can be utilised by Zhipu to fulfil its liquidity requirements when necessary), while we estimate IPO proceeds could add another Rmb4.5bn. Given its loss-making status, we expect Zhipu to conduct further financing activity in the near future, likely via an A-share listing or placement in Hong Kong, to further strengthen its cash position.

Figure 30: Zhipu's operating expenses as % of total revenue



Source: Company data, UBS estimates

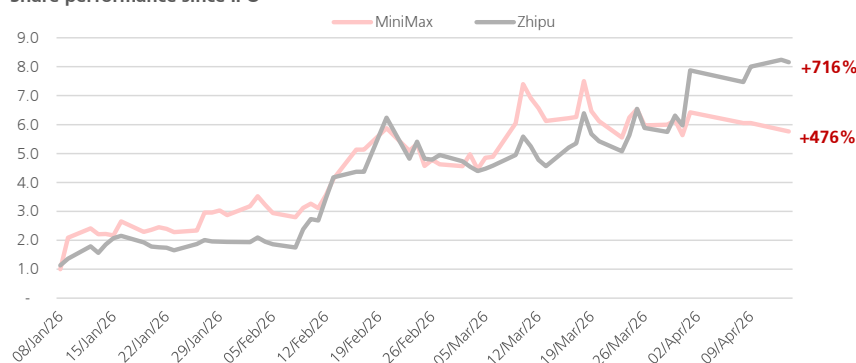
Figure 31: Zhipu's adjusted NP/(loss) and NPM



Source: Company data, UBS estimates

WHAT'S PRICED IN?

Share performance since IPO



Source: Wind, UBS. Note: Prices indexed to the IPO price (IPO price = 1). Data as of 17 April 2026.

Zhipu's stock has had strong price performance post-listing, driven by improving model capability and accelerating monetisation, as widely recognized by the market. Since its IPO on 8 Jan 2026, Zhipu's share price has increased nearly 8x and outperformed closest peer, MiniMax. We believe this has been largely driven by investors' improving recognition of its model capability (especially in coding and agentic tasks after the launch of GLM-5), as well as accelerating monetisation (ARR reached US\$250m in March 2026), driven by rapidly evolving agentic frameworks (ie, OpenClaw, etc.), bringing significant upside potential for to its cloud-based revenue, in our view.

We adopt a **price-to-sales (P/S) multiple** as our primary valuation methodology for Zhipu, considering the following:

- **More intuitive, better comparability:** While we recognise the advantages of P/ARR in capturing rapid technological iteration and evolving monetisation dynamics in the foundation model industry, and note its wide adoption in private market AI valuations, as well as in our valuation of listed peer MiniMax, investor feedback suggests P/S is more intuitive and allows for better comparability with listed peers across subsectors.
- **The current revenue mix is skewed toward project-based:** Zhipu's total revenue gets a predominant contribution from on-premise deployment (c73% in 2025), which is project-based in nature and may exhibit volatility due to seasonality.

We apply **145x P/S** to Zhipu's 2026E revenue of Rmb3,208m, deriving our valuation of HK\$532bn and price target of **HK\$1,160**, based on a 2025-27E revenue CAGR of 231% and 0.6x 2026E price/sales growth (PSG). Our target multiple is supported by the following considerations:

- Our price target for Zhipu is based on 145x 2026E P/S, implying 59x 2027E P/S, on a 231% revenue CAGR in 2025-27E. Compared with our price target for MiniMax, which implies 180x/49x 2026E/2027E P/S, on a 220% revenue CAGR in 2025-27E, our price target for Zhipu suggests slightly higher 2027E P/S on a higher growth profile. On 2026E P/S, we also see upside to our 2026E revenue for MiniMax, which may further digest its P/S multiple to a similar level as Zhipu.
- Our price target implies a comparable PSG with listed domestic AI chip companies, which, on average, are trading at c0.7x consensus PSG in 2026 (based on 69x 2026 P/S and a 105% revenue CAGR in 2025-27).
- While listed AI foundation model companies remain limited globally, we believe Palantir may serve as a reference for Zhipu, considering its AI-driven business model and customer profile (ie, large exposure to enterprise and government customers). Palantir is trading at 0.8x 2026 consensus PSG (based on 48x 2026 P/S and a 61% revenue CAGR in 2025-27).

Figure 33: Target valuation (12 month forward)

UBS projections for Z.ai's revenue (Rmb mn)				
	2023	2024	2025	2026E
Revenue	125	312	724	3,208
YoY growth		151%	132%	343%
Target valuation (12-month forward)				
Revenue (Rmb mn) - 2026E				3,208
Target P/S multiple				145x
Market cap (Rmb mn)				465,117
HKD/RMB				0.87
USD/RMB				6.89
Market cap (HK\$ mn)				532,780
Market cap (US\$ mn)				67,506
No. of shares (mn)				459
Share price (HK\$/share)				1,160

Source: UBS estimates

Figure 34: Comparable valuations

Ticker	Company name	Current Price (LC)	Mkt Cap (USD, m)	P/S (x)		Rev. CAGR	PS/G		EV/EBITDA (x)		EBITDA CAGR	P/E (x)		EPS CAGR	PEG	
				2026E	2027E	25-27E	2026E	2027E	2026E	2027E	25-27E	2026E	2027E	25-27E	2026E	2027E
AI foundation model companies																
0100.HK	MiniMax	859.00	34,403	170	59	220%	0.8	0.3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2513.HK	Zhipu Huazhang	891.00	50,771	145	61	231%	0.6	0.3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Average				157	60	226%	0.7	0.3								
Domestic AI chips																
688795.SS	Moore Threads	656.62	45,248	119	66	83%	1.4	0.8	n.a.	n.a.	n.a.	>100	>100	n.a.	n.a.	n.a.
688256.SS	Cambricon	1334.00	82,472	38	20	104%	0.4	0.2	109.4	63.5	94%	117.7	66.6	101%	1.2	0.7
688802.SH	MetaX	721.00	42,293	73	37	136%	0.5	0.3	n.a.	>100	n.a.	n.a.	>100	n.a.	n.a.	n.a.
6082.HK	Shanghai Biren	47.80	14,900	45	16	95%	0.5	0.2	n.a.	>100	n.a.	n.a.	>100	n.a.	n.a.	n.a.
Average				69	34	105%	0.7	0.3	>100	>100	94%	>100	>100	101%	n.a.	n.a.
Global companies with AI exposure																
PLTR.OQ	Palantir	146.39	349,970	48	34	61%	0.8	0.6	83.1	57.0	67%	111.5	79.0	60%	1.9	1.3

Source: Company data, Refinitiv . Note: The consensus estimates are based on Refinitiv, as of 17 April 2026.

SOTP valuation

We also conducted sum-of-the-parts (SOTP) based valuation as a cross-check, separating Zhipu's business into on-premise deployment and cloud-based open platform and API.

- **For the on-premise deployment segment**, which is project-based, we apply 60x P/S to 2026E revenue of Rmb1,007m, based on an 81% revenue CAGR in 2025-27E and 0.8x PSG (in line with Palantir's), deriving a value of HK\$69.2bn (HK \$151/share).
- **For the open platform and API segment**, which is subscription- or usage-based, with stronger scalability, we apply 100x P/ARR to our Dec 2026E ARR of US \$600m, deriving a value of HK\$474bn (HK\$1,031/share).

In aggregate, our SOTP valuation yields a total equity value of HK\$543bn, or **HK\$1,182/share**, broadly in line with our primary P/S-based valuation, providing additional support to our price target.

Figure 35: SOTP valuation as cross-check

Segment	Business model	Valuation method	Value (HK\$ mn)	Value Per share (HK\$/share)	Justification
On-premise deployment	Project-based	P/S	69,235	151	Apply 60x P/S to 2026E revenue from on-premise deployment business, with +81% revenue CAGR in 2025-27E and 0.8x PSG.
Open platform and API	Subscription-based or usage-based contracts	P/ARR	473,540	1,031	Apply 100x P/ARR to December 2026E ARR of US\$600mn for Open platform and API segment.
Total			542,775	1,182	

Source: UBS estimates

How we frame our upside scenario

As Zhipu is an AI foundation model company, we believe model iteration is the core driver of its revenue growth, with major model upgrades typically delivering stronger intelligence and, in turn, supporting accelerating commercialisation. We view GLM 5.0 as a key milestone in Zhipu's model iteration (launched in March 2026), given its notable improvement in coding and agentic capability.

Using the launch time of GLM 5.0 as the starting point (T+0), we assess the ARR potential of Zhipu's open platform and API (cloud-based) business on Dec 2026 (T+9), considering the following:

- Zhipu's Open platform and API ARR reached **US\$250m** as of March 2026.
- As a benchmark, Anthropic's ARR grew **6.4x** to US\$9bn within nine months following the launch of its key coding model, Claude 3.7 Sonnet.

In our base case, our revenue estimate of Rmb2,201m implies this segment's **ARR reaches US\$600m by Dec 2026**, representing 2.4x growth over 9 months (from Mar 2026), which we view as conservative, considering:

1. **This growth is below Anthropic's recent ramp**, with ARR increasing 3.3x within four months in 2026 YTD, reaching US\$30bn in April 2026 (vs. US\$9bn in Dec 2025).
2. **Zhipu historically has demonstrated even steeper ARR ramp-up**, albeit off a smaller base, with ARR increasing 6x over the past four months (US\$250m in March 2026 vs. US\$39m in Dec 2025) compared with Anthropic taking about eight months to achieve a similar ramp in 2025.

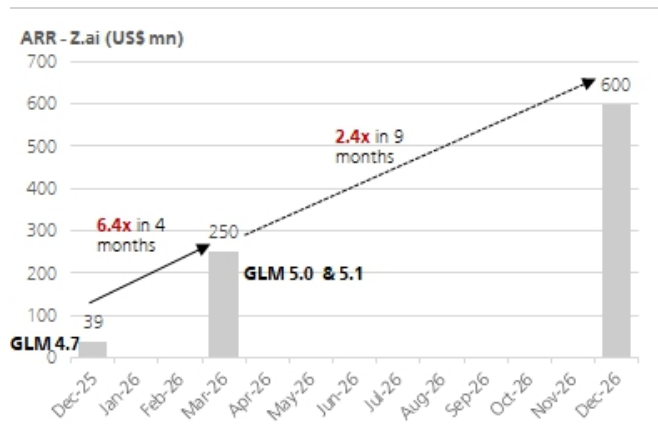
In our upside scenario, we assume Zhipu's open platform and API **ARR reaches US \$1.2bn by Dec 2026** (above management's internal target of US\$1bn). While we acknowledge this is ambitious, we believe it is viable, as follows:

1. It implies ~**5x** growth from T+0, close to ~**6x** achieved by Anthropic's comparable model (ie, Claude 3.7 Sonnet) over the same period.
2. Potential further model upgrades could boost ARR growth, with enhancing capability in long-horizon tasks, which we view as a key feature to allow coding to expand into a broader user base and more productivity-driven use cases, as evidenced by Anthropic's model iteration roadmap (eg, Claude Opus series).

With our upside ARR estimates, while keeping our other assumptions unchanged (ie, on-premise segment revenue and P/S), the current share price of HK\$891 implies Zhipu's open platform and API business is valued at **36x P/ARR**, near the lower end of the **34-54x range** for Anthropic and OpenAI at similar stages in the private market. We view this valuation as attractive, given better liquidity in the secondary market.

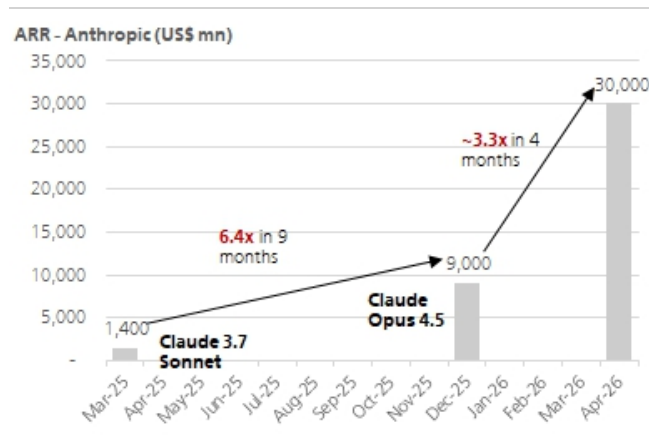
Additionally, under a US\$1.2bn ARR assumption and applying the high-end 54x P/ARR implies upside valuation of **HK\$1,274**, representing 43% potential upside from the current share price.

Figure 36: ARR trend of Z.ai's cloud-based business



Source: Company data, UBS estimates

Figure 37: ARR trend of Anthropic



Source: The Information, Company data

Figure 38: ARR assumptions in our base case and upside scenarios, benchmarking to Anthropic

Unit: US\$ mn

Z.ai

Key model	Launch time (T+0)	ARR, T+0 (Mar 2026)	ARR, T+9 (Dec 2026E)	Growth vs. T+0	Notes
GLM 5.0	Mar-26	250	600	2.4x	Base
			1200	4.8x	Upside

Anthropic

Key model	Launch time (T+0)	ARR, T+0 (Mar 2025)	ARR, T+9 (Dec 2025)	Growth vs. T+0
Claude 3.7 Sonnet	Late Feb 2025	1,400	9,000	6.4x

Source: The Information, company data, UBS estimates

Figure 39: Revenue projection – base case

In Rmb mn	2023	2024	2025	2026E
Total revenue	125	312	724	3,209
YoY growth		151%	132%	343%
On-premise deployment	113	264	534	1,007
YoY growth		134%	102%	89%
Open platform and API	12	48	190	2,201
YoY growth		307%	293%	1056%

In USD mn	2023	2024	2025	2026E
Total revenue	18	45	105	466
On-premise deployment	16	38	77	146
Open platform and API	2	7	28	320

ARR projection:	Dec-2025	Mar-2026	Dec 2026E
Open platform and API	39.0	250.0	600
Growth vs. Mar 2026			2.4x
YoY growth			1438%

Source: UBS estimates

Figure 40: Revenue projection – upside scenario

In Rmb mn	2023	2024	2025	2026E
Total revenue	125	312	724	5,276
YoY growth		151%	132%	628%
On-premise deployment	113	264	534	1,007
YoY growth		134%	102%	89%
Open platform and API	12	48	190	4,268
YoY growth		307%	293%	2142%

In USD mn	2023	2024	2025	2026E
Total revenue	18	45	105	766
On-premise deployment	16	38	77	146
Open platform and API	2	7	28	620

ARR projection:	Dec-2025	Mar-2026	Dec 2026E
Open platform and API	39.0	250.0	1,200
Growth vs. Mar 2026			4.8x
YoY growth			2977%

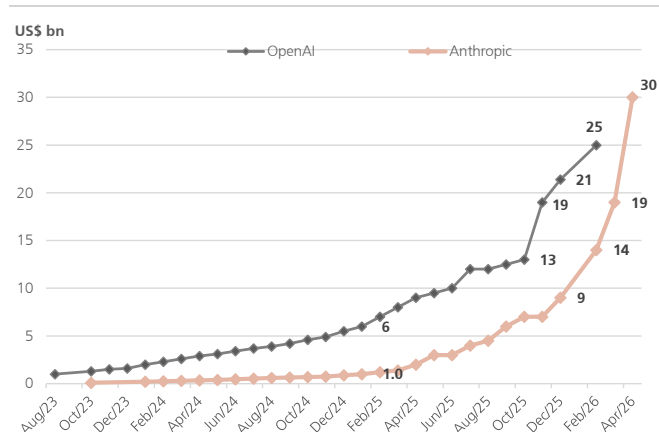
Source: UBS estimates

Figure 41: In our upside scenario, the current share price implies 36x P/ARR for the cloud-based segment

On-premise deployment	
Revenue, 2026E (Rmb m)	1,007
Revenue CAGR, 2025-27E	81%
Multiple	60x
Value (HK\$ m)	69,235
Value Per share (HK\$)	151
Open platform and API	
Revenue, 2026E (Rmb m)	4,268
ARR, Dec 2026E (US\$ m)	1200
Implied P/ARR multiple	36x
Value (HK\$ m)	339,928
Value Per share (HK\$)	740
Total market cap (HK\$ m)	409,163
No. of shares (m)	459
Current share price (HK\$/share)	891

Source: UBS estimates. Note: Price data as of 17 April 2026

Figure 42: ARR trends of OpenAI and Anthropic



Source: The Information, company data. Note: Data as of 7 April 2026

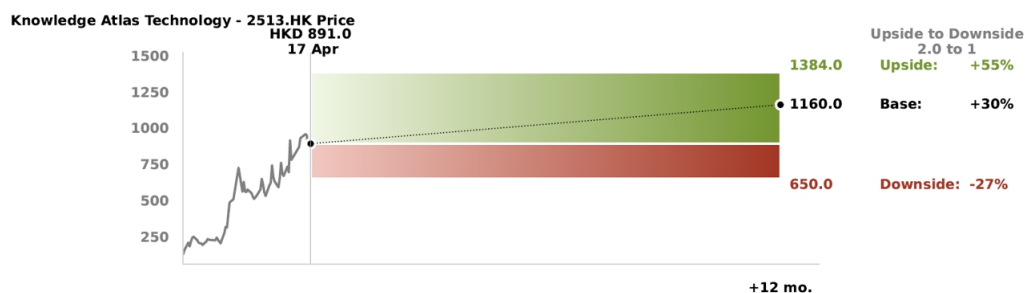
Figure 43: Anthropic and OpenAI valued at 34-54x P/ARR at much larger revenue scale

OpenAI	Apr-2023	Jan-2024	Oct-2024	Mar-2025	Oct-2025	Mar-2026
Valuation (US\$ bn)	29	86	157	260	500	730(pre-money) 850(post-money)
ARR (US\$ bn)	<1	1.6	3.7	6	13	25
P/ARR	>50x	54x	42x	43x	38x	29x/34x

Anthropic	May-2023	Apr-2024	Mar-2025	Aug-2025	Feb-2026	Mar-2026
Valuation (US\$ bn)	4.65	15.35	58	170	350(pre-money) 380(post-money)	
ARR (US\$ bn)	<0.1	0.3	1.4	5	14	19
P/ARR	>50x	51x	41x	34x	25x/27x	

Source: UBS; valuation numbers are from The Information.

UPSIDE/DOWNSIDE SPECTRUM



Zhipu is trading at HK \$891 (as of 17 April 2026).

Value drivers (2026E)	Total rev	On-premise deployment rev	Open platform and API rev	GPM	Adjusted OPM
HK\$1,384 upside	Rmb5,443m	Rmb1,175m	Rmb4,268m	40%	-120%
HK\$1,160 base	Rmb3,208m	Rmb1,007m	Rmb2,200m	33%	-161%
HK\$650 downside	Rmb2,313m	Rmb801m	Rmb1,512m	30%	-240%

Source: UBS estimates

Risk to the current share price is skewed to the upside (2.0:1)

Zhipu is trading at HK\$891 (as of 17 April 2026).

UPSIDE (HK\$1,384): In our upside scenario, we expect Zhipu to further strengthen its leadership in coding and agentic capability, driven by continued model iteration and faster adoption of agentic frameworks, supporting faster ramp-up in cloud-based monetisation. We assume open platform and API revenue reaches Rmb4,268m in 2026, implying ARR scales toward ~US\$1.2bn by Dec 2026, supported by stronger developer adoption, broader agentic use cases and improved monetisation. On-premise deployment revenue grows 120% YoY to Rmb1,175m, with a 100% CAGR in 2025-27. We expect operating leverage from scale and inference optimisation, supporting a gross margin of 40% and an adjusted operating margin of -120%. We apply 100x 2026E P/S on total revenue of Rmb5,443m (implying 54x P/ARR on US\$1.2bn Dec 2026E ARR and 80x P/S on on-premise revenue in 2026E), deriving upside valuation of HK\$1,384.

BASE (HK\$1,160): In our base case, we assume Zhipu sustains its current trajectory of model iteration and commercialisation, with continued improvement in coding and agentic capability supporting revenue growth and gradual scaling of cloud monetisation. We forecast total revenue of Rmb3,208m in 2026 (+342% YoY), with on-premise revenue of Rmb1,007m (+89% YoY) and Open platform and API revenue of Rmb2,200m (+1,056% YoY), driving a mix shift toward cloud. On-premise deployment revenue grows 89% YoY (81% CAGR in 2025-27). Our estimates imply ARR of ~US\$600m by Dec 2026 (vs. US\$250m in March 2026), or c.2.4x growth, which we view as conservative. We expect gross margin of 33% and an adjusted operating margin of -161%. We apply 145x 2026E P/S (implying 100x P/ARR on US\$600m Dec 2026E ARR and 60x P/S on on-premise revenue in 2026E), deriving a price target of HK\$1,160.

DOWNSIDE (HK\$650): In our downside scenario, we assume slower model iteration and a more competitive commercialisation environment, leading to weaker cloud adoption and delayed ramp-up in agentic use cases. Open platform and API revenue reaches Rmb1,512m in 2026, with ARR only reaching US\$400m in Dec 2026 due to weaker adoption, pricing pressure and slower monetisation. On-premise deployment revenue grows 50% YoY to Rmb801m, with a 30% CAGR in 2025-27. Total revenue reaches Rmb2,313m, with limited operating leverage. We expect gross margin of 30% and an adjusted operating margin of -240%. We apply 110x 2026E P/S (implying 80x P/ARR on US\$400m Dec 2026E ARR and 50x P/S on on-premise revenue in 2026E), deriving downside valuation of HK\$650.