HELLO CHINA TECH

The China Tech Field Guide

Six principles from a year of essays on China's AI, chips, robots, and EVs. Tested in public.

2026 EDITION

Contents

The framework in one page

1 **Capital structure beats technology**

The survival of Chinese technology companies is determined less by what they build than by who funds their losses and for how long.

2 **The state is the market's architect**

The government designs technology markets, funds their construction, plays in them, and decides when overcapacity has served its purpose.

3 **The layered market is the durable advantage**

Multiple technology maturity levels and demand tiers coexist, so a good enough product finds a paying customer before it is world-class.

4 **Pipes beat apps**

Value accrues below the application layer: memory, cloud inference, payment rails, and operating systems outlast the chatbots on top.

5 **Parallel substitution, not catch-up**

China is building a complete, domestically controlled stack at commercially acceptable performance, not racing to the frontier.

6 **The interface war decides the next decade**

Whoever owns the interface between user intent and service execution taxes everyone below it. This is the one open arena.

> Read every sector through three lenses at once: global technology state, the state's position in the ripen-then-glut cycle, and layered market demand. Omit one and the picture distorts.

What does a year of tracking China's technology economy actually reveal? Start with a number. In March 2026, China's National Data Administration reported that daily token consumption across the country had reached 140 trillion. Two years earlier, the figure was 100 billion. Growth exceeded a thousand-fold. The same month, Liu Liehong, the administration's director, gave tokens an official Chinese name, cíyuán (词元), and called them "the value anchor of the intelligent era" and "the settlement unit connecting technology supply with commercial demand." Tokens had become a cost line, not a metaphor. They were showing up in cloud invoices, in enterprise procurement budgets, in the cost structures of companies building AI products. The most consequential layer of China's AI economy turned out to be neither the chatbot nor the model. It was the cost infrastructure underneath.

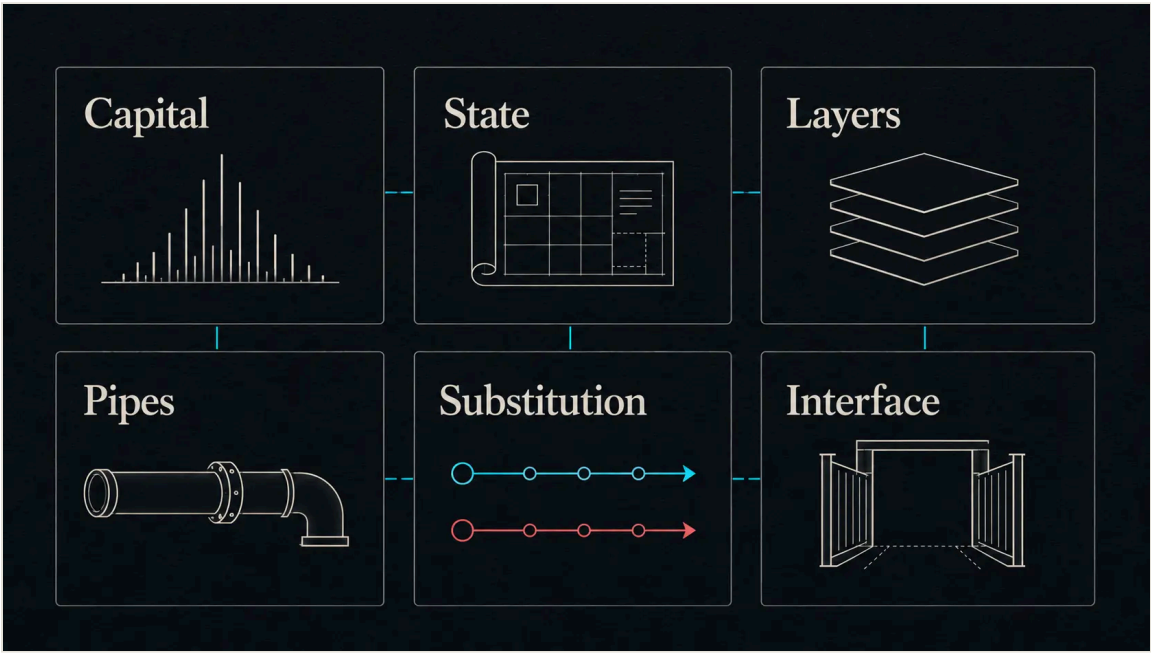
Tokens had become a cost line, not a metaphor.

Token billing is the most visible example. The deeper pattern spans six dimensions. Value accruing below the application layer is one of six principles in this document. The others cover capital structure, state-directed markets, layered demand, parallel technology substitution, and the contest over who controls the digital interface. Together they form a working framework for reading China's AI, semiconductor, robotics, and EV sectors. The framework draws on the essays published in Hello China Tech between July 2025 and July 2026, each linked where it appears. This is the 2026 Edition, the first in an annual series.

A note on what this is not. It offers no investment advice and no US-versus-China scorecard. The principles describe how China's technology economy operates. They do not prescribe positions or predict winners.

One commitment separates this from most frameworks distributed as free downloads: it has been tested in public. The principles below were not written first and verified later. They emerged from 12 months of published analysis, each essay staking a specific claim that subsequent events could confirm or contradict. This document consolidates those claims into a framework for the first time. Part I presents the six principles. Part II scores them against the evidence, documenting where the analysis called the outcome correctly and where it missed. Part III provides the falsifiability conditions: the specific, observable facts that would require each principle to be revised or abandoned.

A framework that cannot be wrong is not a framework. It is a slogan. The value of this one rests on whether it has been useful, and on how honestly it accounts for its errors.


Capital
State
Layers
Pipes
Substitution
Interface

Part I: Six Principles

PRINCIPLE 1

Capital structure beats technology

The survival of Chinese technology companies is determined less by what they build than by who funds their losses and for how long. In sectors where the state sets the pace of development, the question that predicts which firms endure is not who has the best product. It is whose losses are underwritten.

This principle operates through capital concentration. In 2025, 764 AI companies in China raised a combined Rmb 83bn. The top 20 captured 30 per cent of that total, with a median raise of Rmb 1bn each. Eleven were already unicorns. The remaining 744 averaged Rmb 78m apiece. The distribution resembles a spike, not a pyramid: a small cohort of anointed firms surrounded by hundreds competing for what remains. RMB-denominated funds now account for 95 per cent of all venture deals, which means the capital flowing into Chinese AI is overwhelmingly domestic, state-adjacent, and denominated in a currency whose allocation channels the government influences directly. (China's state-engineered AI boom)

The mechanism extends beyond venture rounds. State capital, or guozi (国资), has become the dominant force in the larger funding rounds that confer unicorn status. Private venture capitalists may identify early-stage opportunities, but it is state-backed funds that select which companies reach the scale where survival becomes structurally secure. The result is a Capital Intensity Multiplier: Chinese chip companies, for instance, appear to require 10 to 20 times more capital than established global players to reach comparable market positions. That multiplier reflects the simultaneous cost of reinventing existing technologies, building software ecosystems from scratch, competing for scarce talent, and localizing supply chains. MiniMax, another model lab whose cap table once featured elite dollar-denominated VCs, recently took a $300m strategic cheque from Shanghai's state asset manager.

The question that predicts which firms endure is not who has the best product. It is whose losses are underwritten.

The Cambricon and Metax contrast illustrates the principle in practice. Cambricon, a domestic AI chip designer backed by deep state capital and major tech investors, reported Q1 2025 revenue of Rmb 11.11bn, a 4,230 per cent year-over-year increase, and Rmb 3.55bn in net profit. Its market capitalisation reached approximately Rmb 579bn. Metax, operating at a different scale, raised capital at a Rmb 21.07bn valuation while posting a Rmb 1.41bn loss in 2024 on revenue of Rmb 743m from over 25,000 shipped GPUs. Both companies build AI chips. The divergence in trajectory points toward differences in capital access and the patience of their respective backers, not solely toward differences in silicon. (Cambricon's AI chip turning point)

At the startup layer, the Darwin Moment, the point at which capital scarcity forces startups into an existential selection event, has arrived. Zhipu AI, a Tsinghua University spinout backed heavily by state funds, generated Rmb 300m in revenue against a Rmb 2bn loss in 2024, pushing toward an IPO at roughly Rmb 40bn. Moonshot, a consumer AI startup, saw its monthly advertising budget collapse from Rmb 100m in January 2025 to Rmb 400,000 by June, squeezed out by incumbents whose war chests dwarf anything venture capital can sustain. ByteDance alone allocated over $20bn for AI chips and data centres in 2025. The startups that survive this period will be those that answered one question honestly: what can we actually build, sell, and defend given who backs us and what our market will pay for?

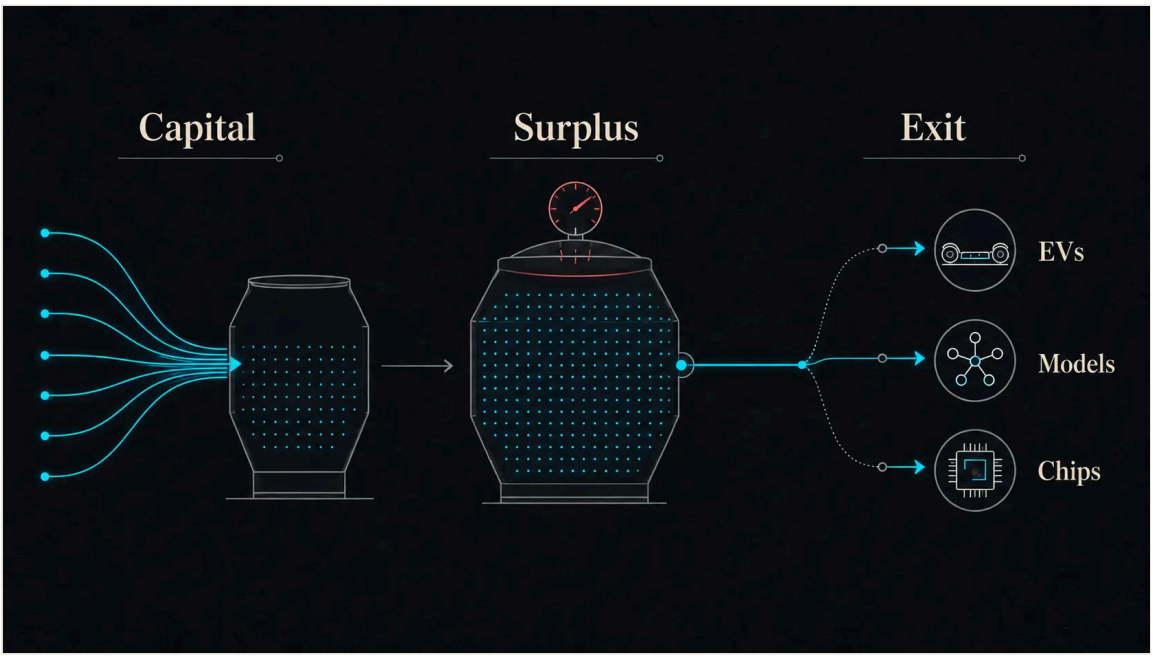
That capital structure is itself a product of deliberate design.

PRINCIPLE 2

The state is the market's architect

The Chinese government does not referee technology markets. It designs them, funds their construction, plays in them, and decides when overcapacity has served its purpose. The recurring pattern is the ripen-then-glut cycle: the state accelerates an industry's growth through directed capital and policy support, rides the resulting overcapacity, then channels the fallout differently depending on the industry's structure.

The mechanism has three phases. In the first, state capital floods a sector to compress the timeline from laboratory to market. The central government launched an $8.2bn national AI investment fund in January 2025 and followed with an "AI+ Action Plan" in August 2025, requiring local governments to produce measurable implementation schemes. In the second phase, the capital achieves its purpose: capacity exceeds demand. EVs have already crossed this threshold. AI models, chips, and humanoid robotics remain in what might be called the ripening phase, with early signs of surplus appearing in models and policy-backed procurement flowing into robotics. In the third phase, the exit route depends on the product. Manufacturing industries like EVs export their overcapacity. AI models, which cannot be shipped like cars, deflate into the domestic enterprise market as costs fall. Chips consolidate under continued policy direction. Each industry follows the same cycle but exits through a different door.



The Chinese government does not referee technology markets. It designs them, funds their construction, plays in them, and decides when overcapacity has served its purpose.

The Hefei Model offers the clearest case study. In 2008, the city government committed approximately $1.3bn to help BOE Technology, a domestic display maker, build a next-generation production line. That sum approached half of the city's entire fiscal revenue for the year. The government acted as a strategic anchor investor, de-risking the project for private capital. BOE raised Rmb 12bn by June 2009. The factory cluster now produces 20 per cent of the world's LCDs. The government exited its initial stake at a profit estimated at roughly 300 per cent. (Inside China's government VC model)

Hefei replicated the playbook twice more. In April 2020, the city rescued NIO, a premium EV maker reportedly rejected by 18 other cities, with a Rmb 7bn injection for a 24 per cent equity stake. Hefei's investors earned a substantial profit within 18 months. Then came CXMT, China's national champion in DRAM memory chips, which the city incubated with approximately 75 per cent of initial capital. CXMT's factory was built in 10 months. Its pending IPO carries a valuation of roughly Rmb 150bn ($21bn). Three industries, three time horizons, one playbook: the government provides the patient capital that compresses development timelines, absorbs early losses, and creates the conditions for a domestic industry to reach self-sustaining scale.

The venture capital data confirms the pattern at the national level. The same spike visible in AI funding, where the top 20 firms capture 30 per cent of all capital while 744 competitors split the remainder, reflects a system where state capital selects a small number of firms for strategic cultivation. Private venture investors identify candidates. The state decides which ones scale.

The market these architects are building has a specific shape.

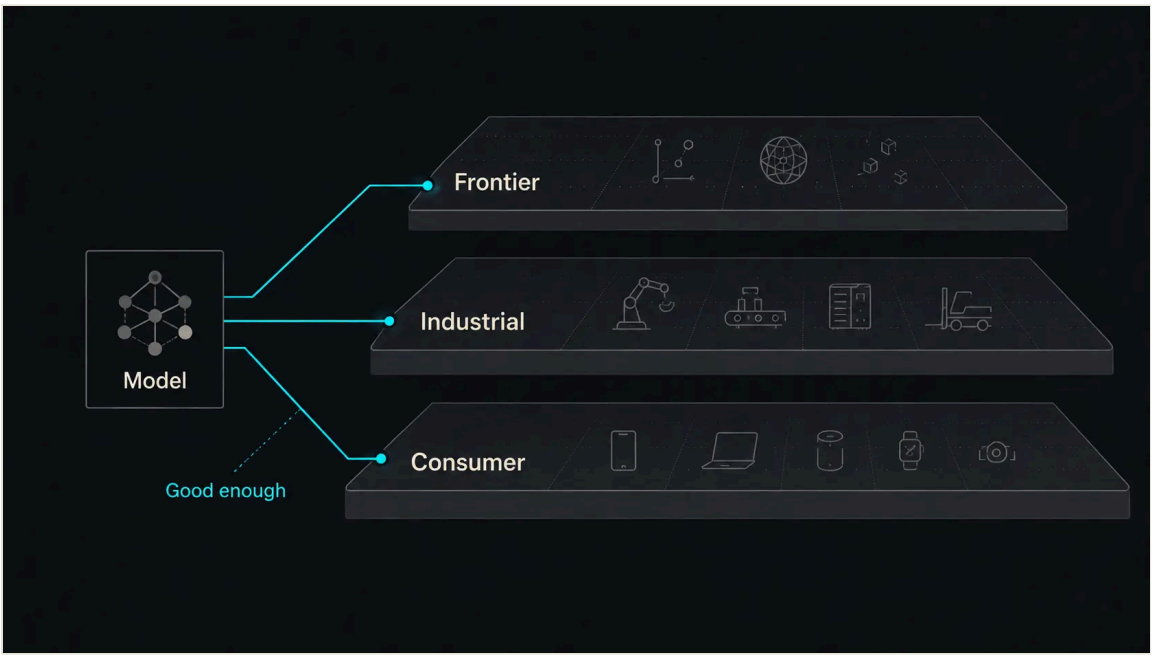
PRINCIPLE 3

The layered market is the durable advantage

China's market advantage rests less on size than on the simultaneous coexistence of multiple technology maturity levels and demand tiers, so a product that qualifies as good enough can find a paying customer before it qualifies as world-class. This is a demand-side structural feature, distinct from the supply-side forces of capital and policy described above. The better

term may be right-sized intelligence: a model does not need to be frontier-grade if the task, price point, and deployment environment only require practical AI. Application innovation persistently outruns frontier innovation because the layered market gives each iteration somewhere to land.

The mechanism works across two dimensions. The first is temporal: technologies at different stages of maturity operate in the same economy at the same time. A frontier reasoning model serves a cloud provider's most demanding enterprise clients. A 67-billion-parameter model, far from the frontier, handles government paperwork in Shenzhen's Futian district at 95 per cent accuracy while cutting review times by 90 per cent. Both deployments generate revenue. The second dimension is spatial: top-tier cities, smaller-city markets, and industrial settings have structurally different needs, price sensitivities, and infrastructure constraints. A product designed for a Tier 1 cloud customer and a product designed for a factory floor in a smaller city may use the same underlying model family at different scales. The layered market supports both without requiring either to dominate.



The consequence is that winner-take-all dynamics, familiar from American platform markets, appear less likely to hold in China. The global small language model market is projected to grow from $0.93bn in 2025 to $5.45bn by 2032. Chinese companies are capturing disproportionate share through three factors: data sovereignty requirements that favour local deployment, cost sensitivity in manufacturing, and infrastructure realities in smaller cities. The good enough threshold turns out to be commercially productive, not merely a consolation prize.

China's AI toy industry illustrates how layered demand creates entire product categories. The market reached roughly Rmb 25bn ($3.5bn) in 2024, accounting for 38 per cent of the global total, with projections pointing toward Rmb 100bn by 2030. Over 1,500 AI toy companies were operating in China by late 2025. On JD.com, AI toy sales for children aged 3 to 6 surged six-fold in February 2025. BubblePal shipped over 200,000 units. FoloToy reached 20,000 units in Q1 2025, projects 300,000 for the year, and sells across 10 countries. These products run on open-source models and commodity voice modules manufactured minutes away from AI chip suppliers in Shenzhen and Dongguan. The layered market provided the demand tiers, the supply chain provided the iteration speed, and cheap inference provided the economics. Western toy companies were still debating the ethics of conversational AI for children while Chinese manufacturers were shipping millions of units. (China's $3.5bn AI toy industry)

A common misreading: Western analysts frequently apply winner-take-all frameworks developed in American platform markets to Chinese technology sectors. The expectation that a single cloud provider, a single model lab, or a single chip vendor will consolidate each category the way their US equivalents have tends to underestimate the structural persistence of layered demand. The layered market structure suggests that high concentration is harder to achieve and sustain. The layered market does not prevent dominant players from emerging, but it suggests that the path to dominance runs through serving multiple tiers rather than collapsing them into one.

If the market is layered, the question becomes which layer captures the most value.

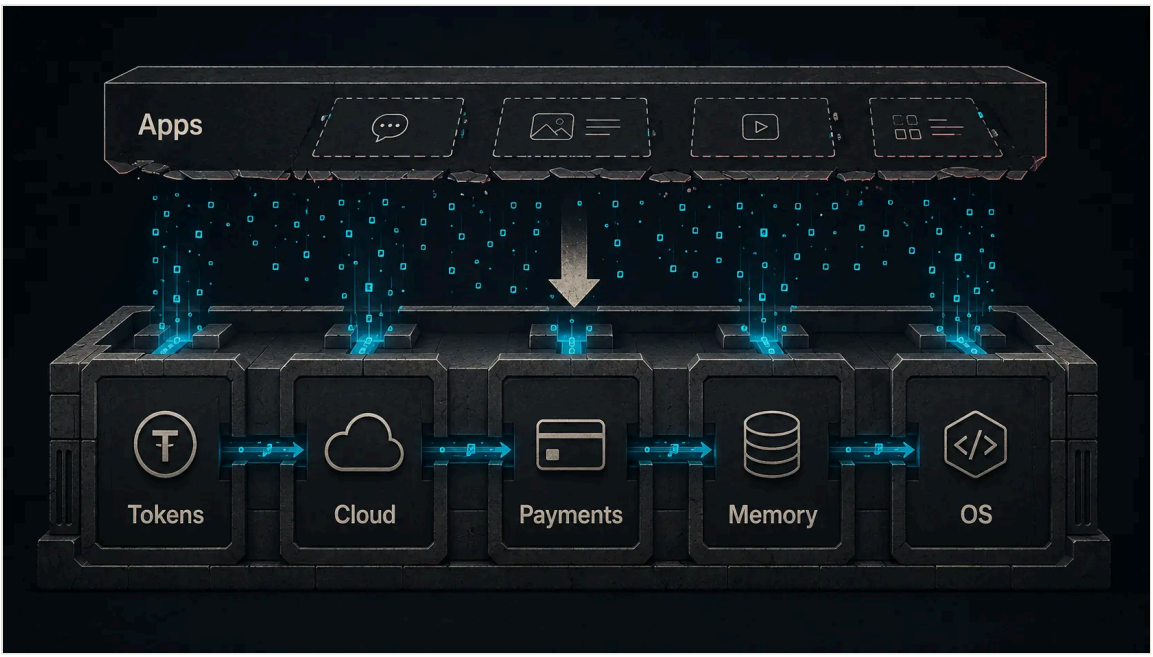
PRINCIPLE 4

Pipes beat apps

Value in China's AI economy accrues below the application layer. Memory chips, cloud inference platforms, payment rails, and device operating systems capture margin more durably than the chatbots and consumer apps that sit on top. The model layer commoditises fastest precisely where the layered market is strongest. In China's AI economy, the scarce asset is not the chatbot. It is the metered, billable infrastructure that makes the chatbot cheap enough to use at scale.

The token economy illustrates the mechanism. The 140 trillion daily tokens described in the opening are not merely a volume statistic. They represent billing events flowing through cloud infrastructure. Agent workloads consume tokens at 10 to 100 times the rate of conversational chat. That ratio shifts the economics of the entire stack. ByteDance's Volcano Engine, Alibaba Cloud, and Tencent Cloud are each pricing and measuring AI infrastructure around token throughput. ByteDance's daily cloud model throughput reached 120 trillion tokens by April 2026, growing a thousand-fold from its May 2024 launch. JPMorgan

estimates that China's inference token consumption will grow roughly 370-fold between 2025 and 2030. At this scale, whoever operates the most efficient token delivery infrastructure holds a structurally advantaged position. The signal extended beyond China's borders. In the week of March 16, Chinese AI models generated 7.36 trillion tokens on OpenRouter, 36 per cent of global volume, surpassing American models for the third consecutive week. IDC put China's MaaS market at Rmb 710m in 2024; the category is projected to exceed 30 per cent of total cloud income. The pipes, not the applications running through them, capture the rent. (The token economy China is building)



The pipes, not the applications running through them, capture the rent.

The same logic extends to hardware. When Samsung exited China's TV and home-appliance market in May 2026, investors rewarded the move. The semiconductor division generated 61 per cent of total revenue and 94 per cent of operating profit. Samsung holds approximately 37 per cent of global DRAM and 28 per cent of NAND flash. Chinese brands had captured 94.1 per cent of domestic TV shipments. Samsung lost the consumer-facing layer where Chinese companies hold decisive advantages in assembly, channel management, and price competition. It concentrated on the upstream layer where it holds pricing power: advanced memory fabrication. The market treated the retreat as overdue housekeeping because the value had already migrated to the pipes. (Samsung's value migration to memory)

The infrastructure thesis is also visible in China's agent economy. Alipay AI Pay surpassed 100 million users and processed over 120 million transactions in a single week. Xiaomi committed Rmb 60bn ($8.7bn) to AI over 3 years, with over 800 million connected IoT devices providing a surface for system-level agent deployment. Baidu proposed measuring Daily Active Agents, or DAA, as the output metric for an era where agents, not humans, generate the bulk of compute demand. The competitive question these companies are answering is not who builds the best agent. It is who builds the plumbing that agents run on: the inference infrastructure, the payment authorisation layer, the device operating system, the orchestration platform that minimises wasted tokens.

One executive framed the implication: the agent era does not belong to whoever burns the most compute. It belongs to whoever uses it wisely. That distinction between volume and efficiency points toward a new pricing problem. Alibaba Cloud's Pro coding plan, priced at Rmb 200 per month, sells out by 9:30 each morning. One Anthropic Claude Max subscriber generated $5,600 in equivalent API costs on a $100 monthly plan. Both cases reflect the same structural issue: subscription pricing built for chat-era consumption cannot absorb agent-scale workloads. The companies solving that problem operate below the application layer.

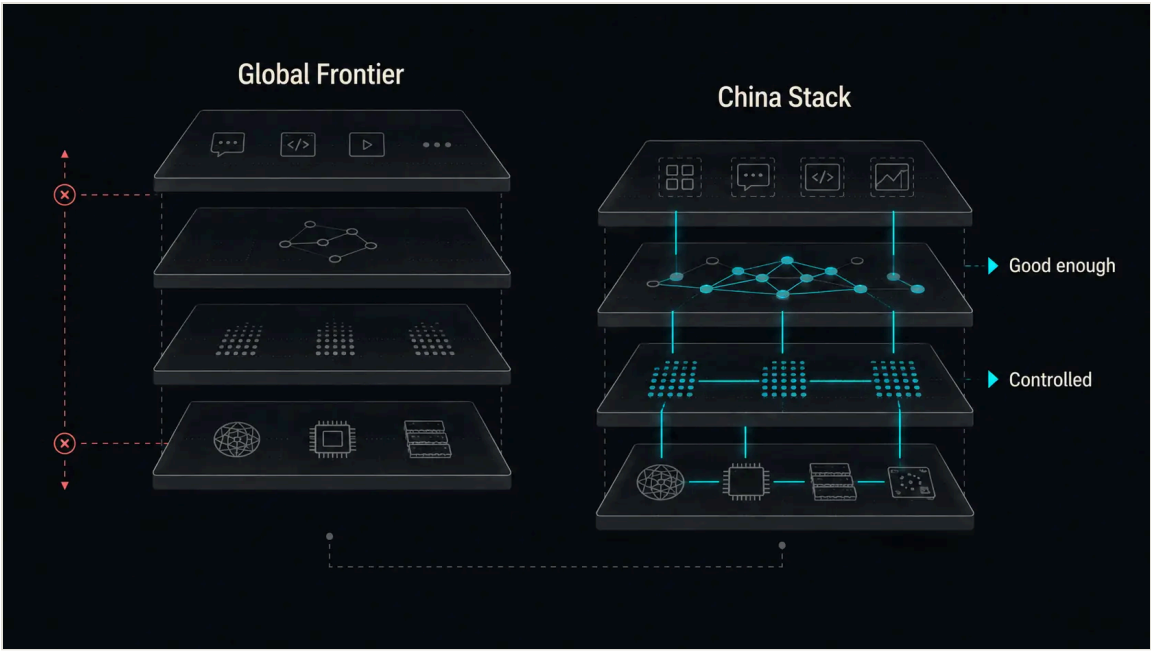
The infrastructure layer faces its own constraint: much of it must be built without access to the global frontier.

PRINCIPLE 5

Parallel substitution, not catch-up

China's structural disadvantage in semiconductors and its lack of a stable technology alliance with the West appear to be durable conditions under current geopolitical arrangements rather than transitional ones. The operating strategy is the construction of a parallel good enough stack: a complete, domestically controlled technology chain that functions at commercially acceptable performance, even where it trails the global state of the art. Company moves that look like grand strategy are better read as forced adaptation to this constraint.

The logic begins with the chip gap. Metax, in a regulatory filing of unusual candour, acknowledged that the vast majority of domestic manufacturers maintain a 1 to 2 generation gap with Nvidia's latest products. Its C500 series benchmarks against the A100, a chip Nvidia released in 2020. Its C700 targets H100 performance from 2022. Metax's software ecosystem has roughly 15,000 registered developers, compared with millions for Nvidia's CUDA platform, a gap of approximately 3,000 to 1. These constraints appear structural given current manufacturing access and are unlikely to close through engineering effort alone.



But the operating logic points away from closing the gap on Nvidia's terms. Chinese model labs are instead reshaping workloads around domestic hardware constraints. DeepSeek's V4 technical report reads less as a model benchmark and more as a compute-shaping document. V4-Pro uses 27 per cent of the single-token inference FLOPs and 10 per cent of the KV cache of its predecessor in a one-million-token context setting. The report goes further, offering specific proposals to hardware vendors on power headroom, cross-GPU signalling latency, and activation function design. Frontier model labs are beginning to tell the hardware industry what useful compute should look like. DeepSeek's API pricing page notes that V4-Pro throughput is constrained by high-end compute availability, with an expected price reduction once Huawei's Ascend 950 series reaches volume production in the second half of 2026. The model's commercial roadmap is explicitly timed to a domestic chip delivery schedule. (DeepSeek's compute-shaping bet)

Geopolitics shapes the constraint. Commerce shapes the response.

The training stack has also shifted. For 2 years, the working assumption held that Chinese chips could handle inference but not the far more demanding work of training models from scratch. In late June 2026, Meituan released LongCat-2.0, a trillion-parameter model pre-trained on a 50,000-chip domestic cluster, ingesting over 35 trillion tokens without a single rollback. Separately, a research consortium completed full-parameter post-training of DeepSeek-V4-Pro on over 1,000 Ascend 910C chips with zero interruptions. Huawei's own

Pangu Ultra 135B pre-trained 13.2 trillion tokens on 8,192 Ascend NPUs. Chinese domestic chips have entered the training stack. The pace of entry appears to be increasing. The constraint is not only chip design. SMIC's 7nm capacity is expanding from roughly 45,000 wafers per month at the end of 2025 to approximately 60,000 in 2026, while Huawei targets roughly 750,000 Ascend 950PR units in the same period. Bernstein analysts project that Chinese domestic AI chip suppliers' market share will rise to 55 per cent by 2027. The gap is not disappearing. It is becoming a question of system-level economics rather than a binary of can or cannot.

The more useful reading is commercial rather than ideological. External observers often treat supply chain localisation as nationalist strategy. In practice, many procurement decisions follow a simpler logic: when Nvidia chips become unavailable or restricted, firms that need compute source alternatives. When domestic chips cross the good enough threshold, procurement can shift for commercial reasons, not only political ones. Geopolitics shapes the constraint. Commerce shapes the response. Conflating the two obscures the mechanism that actually drives company behaviour.

The parallel stack describes how China is rebuilding the supply side under constraint. The unresolved question is who controls the demand side: the interface through which users, agents, and services meet.

PRINCIPLE 6

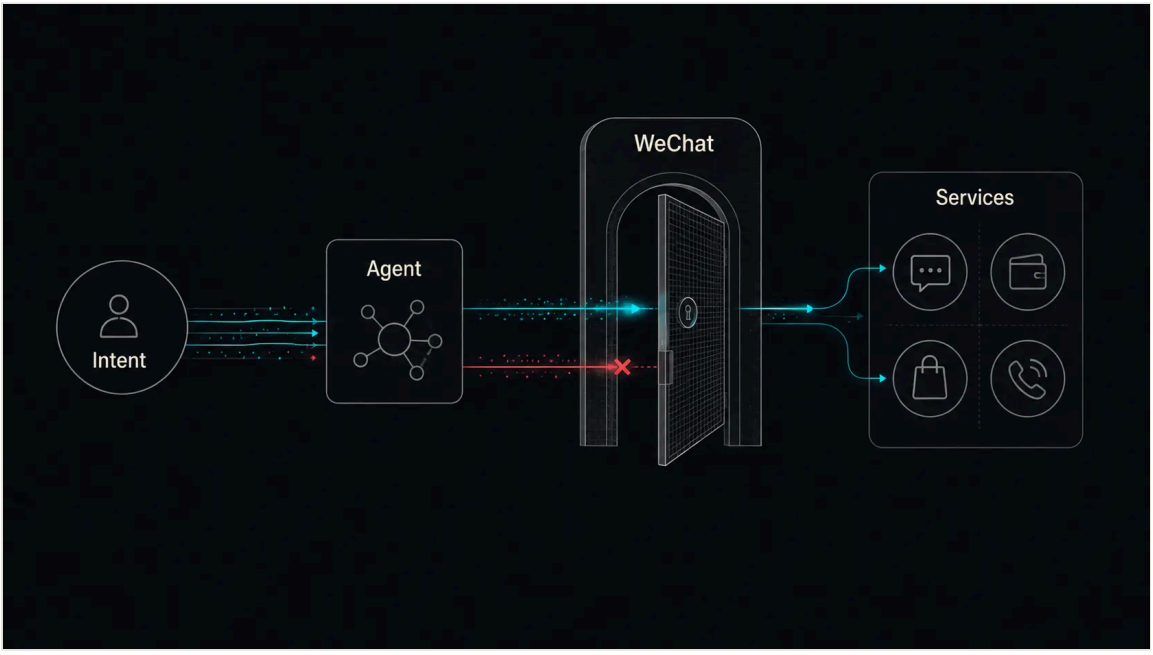
The interface war decides the next decade

Whoever owns the interface between user intent and service execution taxes everyone below it. China's technology giants are fighting to replace the search box and the app grid with a conversational layer so deeply embedded in daily life that switching becomes prohibitively costly. This is the one arena covered by these principles where the outcome remains genuinely open.

The fight has two fronts. The first is the Do-Engine: an AI-powered interface that does not merely find information but synthesises, plans, and executes. Alibaba's Quark and Tencent's Yuanbao each spent roughly Rmb 1bn per month on advertising in mid-2025. The logic is defensive. If a rival builds the dominant Do-Engine, it can disintermediate every other ecosystem. A user who asks an AI agent to find the cheapest latte across three delivery platforms turns those platforms into interchangeable commodity suppliers. Baidu has integrated over 18,000 third-party services into its AI framework. The spending is about seizing the conversational operating system for daily commerce, not building a better chatbot. (The Do-Engine war)

The second front is the contest over who mediates between AI agents and the services they need to access. In December 2025, ByteDance launched Doubao Phone Assistant on ZTE's Nubia M153. The product used a GUI Agent approach: rather than integrating with apps through official APIs, it operated at the system level, reading screens and simulating finger taps. Within days, WeChat forced users to log out. Taobao triggered CAPTCHA verifications. ByteDance retreated, disabling WeChat functionality. Simple tasks took roughly 30 seconds, compared with roughly 10 seconds for a human. The first batch comprised only 30,000 units at a retail price of Rmb 3,499. Secondhand prices ran Rmb 700 to 1,500 above retail. ByteDance was not trying to compete in smartphones. It was testing whether a forbidden path could work at all, because the permitted path, API cooperation, requires permission from the platforms the agent aims to disrupt.

By June 2026, Tencent had reversed course selectively. WeChat opened a controlled A2A channel, a Tencent-defined Agent-to-Agent access framework distinct from the open A2A protocol promoted by Google, to phone manufacturers including Huawei, Honor, Xiaomi, OPPO, and Vivo. Users could ask their phone assistant to send a WeChat message or make a call. Honor completed integration across three device series, with 50 per cent of online active devices connected. The mechanism keeps execution inside WeChat. The phone assistant initiates the request; WeChat processes it within its own environment. Tencent defined the scope, granted access to communication functions, and retained the ability to expand or withdraw permissions. The door opened, but Tencent designed the door. (WeChat as AI gatekeeper)

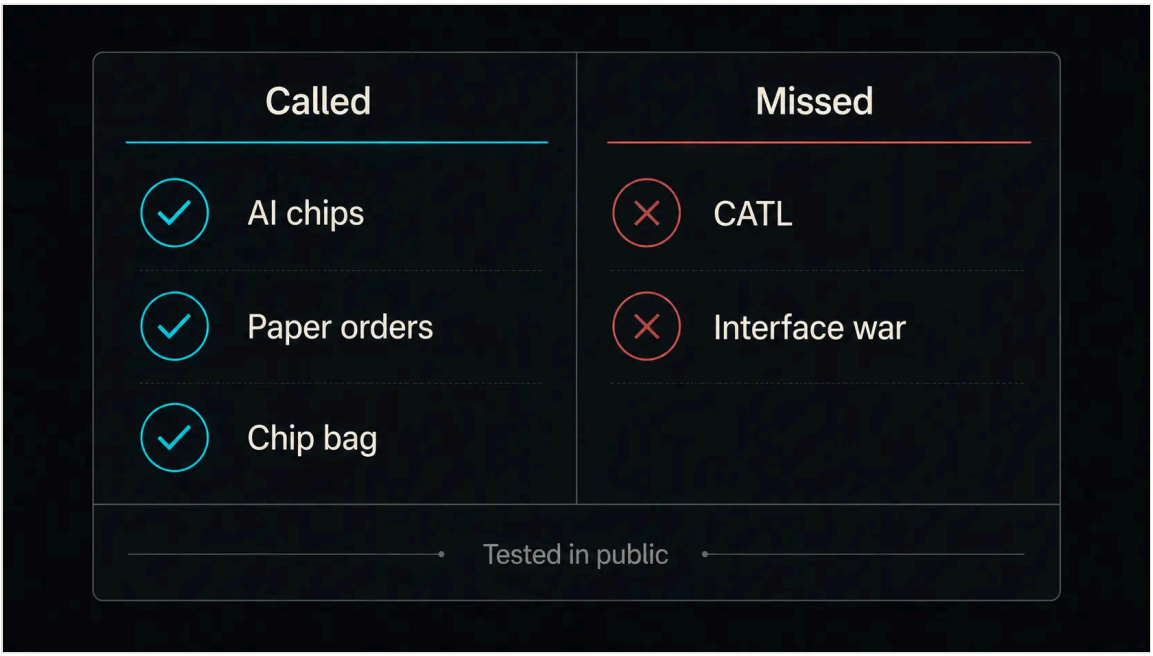


The door opened, but Tencent designed the door.

The protocol power at stake is structural. WeChat has 1.4 billion monthly users. Tencent's standalone AI app, Yuanbao, had roughly 57.35 million monthly active users in March 2026, compared with 345 million for Doubao and 166 million for Qwen. Tencent's capex in 2025, approximately Rmb 79bn, trailed ByteDance's reported Rmb 150bn and Alibaba's Rmb 380bn three-year commitment. WeChat's ecosystem remains Tencent's strongest card. Any phone agent that cannot interact with WeChat appears strategically incomplete in China. That gives Tencent protocol-level leverage over the interface war even while its standalone AI product lags. Whether that leverage holds depends on whether WeChat's forthcoming AI agent delivers enough value to keep intent flowing through its ecosystem, or whether users begin routing tasks around it.

Part II: The Evidence So Far

The working hypotheses behind these principles took shape during the first months of Hello China Tech's coverage, starting in July 2025. An internal working document from October 2025 codified them as five principles. Twelve months of published analysis now provide enough evidence to score those hypotheses against outcomes. The cases below are divided into what the analysis anticipated and what it did not. The wins and misses use the same evidentiary standard: specific claims matched against specific facts.



What the framework called

The good enough threshold in AI chips

The working hypotheses codified in October 2025 argued that finding a good enough substitution path mattered more than closing the gap with global leaders. They argued that parallel substitution paths would emerge, not the most advanced, but sufficient to support

the majority of application scenarios.

That argument has been tested commercially. Cambricon's Q1 2025 results, detailed in Principle 1, showed the financial side of good enough in action. The technology side was equally telling. Its Siyuan 590 chips deliver approximately 80 per cent of Nvidia A100 performance at 30 per cent lower cost than Nvidia's China-specific H20 in certain scenarios. Revenue recognition shifted from project-based contracts concentrated in Q4 to consistent quarterly sales, suggesting a move from one-off government procurement toward recurring commercial demand. Gross margin stabilised around 70 per cent, comparable to Nvidia's data centre business.

Metax, operating at a different scale, confirmed the structural picture from the other side. Its 300-page regulatory filing for the Shanghai Stock Exchange confirmed the generation gap detailed in Principle 5. But Metax secured a $3bn valuation on what amounts to a security premium logic, where customers value supply chain independence over marginal performance advantages. The framework's claim was that good enough would find a paying market. The evidence suggests the market arrived faster than the October 2025 version anticipated.

The paper-order problem

The framework identified three risks in the robotics sector: technology and manufacturing cost barriers, limited deployment scenarios, and excessive short-term hype. It placed the real opportunity roughly 5 years out, based on a technology maturity assessment.

The order wave that followed appeared to challenge the timeline. UBTECH announced a Rmb 250m contract, a record for the sector. Galbot secured contracts exceeding Rmb 700m. China Mobile awarded a Rmb 124m joint tender to AgiBot and Unitree Robotics. The headline numbers suggested the market was arriving faster than expected.

Closer examination supported the framework's caution rather than refuting it. Tiantai Jiqiren announced a 10,000-unit order worth billions of renminbi. The purchasing company had been established 12 days before the deal announcement, with overlapping personnel between buyer and seller. Unitree Robotics, in its IPO filing, disclosed that 73.6 per cent of its humanoid robot revenue in the first 9 months of 2025 came from research institutions and universities. Its adjusted net profit fell 52.55 per cent year-over-year in Q1 2026 as spending surged. The order rush was real. So was the paper-order problem. The framework correctly identified the hype pattern. Its 5-year timeline proved slightly conservative on order announcements but appears closer to accurate on actual commercial deployment readiness.

The chip bag test for deployment readiness

Galaxy General, China's best-funded embodied AI startup, raised nearly Rmb 7bn and reached a valuation exceeding Rmb 20bn on a thesis that the robot's brain, its AI model, matters more than its body. A FamilyMart convenience store on the ground floor of its Beijing headquarters became the test case.

The framework identified the right failure mode. The chip bag gave it a concrete name.

The store carries roughly 1,000 SKUs. Galbot handles fewer than 10 of them. A bag of Lay's potato chips, pressurised with food-grade nitrogen, defeats the robot's grippers roughly 4 times in 10. The robot is priced at approximately Rmb 700,000. Four full-time employees staff the store, the same headcount as a comparable FamilyMart without a robot. Galaxy General's purchases from Unitree, a competing hardware company, surged from Rmb 364,700 in 2024 to Rmb 18.19m in 2025, a roughly 50-fold increase. Dexterous hands demonstrated at the Spring Festival Gala came from Sharpa, a specialist manufacturer, at a cost of several hundred thousand renminbi per pair.

The framework argued that technology maturity, manufacturing costs, and deployment limitations would constrain the sector. Galaxy General's convenience store deployment provided a measurable version of those constraints. The brain-first thesis encounters a trilemma: sensing, execution, and cost constrain each other simultaneously. The AI model does not yet compensate for simple grippers on deformable packaging, does not yet eliminate latency from cloud-based voice processing, and does not yet bring dexterous-hand costs to a level retail can absorb. At Rmb 700,000 per unit, a convenience store operator could hire several full-time workers for years. The framework identified the right failure mode. The chip bag gave it a concrete name.

What it missed

CATL's data centre playbook

The early analysis discussed EV battery overcapacity and correctly identified export as the primary exit route for manufacturing surplus. It did not anticipate that a dominant battery company would begin assembling a vertical energy chain for AI data centres.

The wrong assumption was specific: the framework treated each industry's ripen-then-glut exit as relatively contained. EVs would export. Models would deflate into domestic enterprise. Chips would consolidate under policy. The framework did not consider that a company dominant in one cycle could use overcapacity as a launch point into an adjacent value chain that the framework had not yet mapped.

CATL invested Rmb 4.1bn for a 49 per cent stake in the controlling shareholder of Zhongheng Electric, a provider of high-voltage direct current power systems deployed in data centres. It agreed to pay up to $942m for as much as 38.1 per cent of VNET Group, a data centre operator running more than 50 facilities with 889 megawatts of wholesale capacity. Reports put its commitment to DeepSeek's fundraise at roughly $740m. Total exposure potentially exceeds Rmb 15bn. Robin Zeng, CATL's founder, estimated the data centre energy opportunity at "ten times" the size of EV battery supply. The company expects energy storage to account for half of global sales by 2030.

The framework's ripen-then-glut cycle can explain the motivation retrospectively: battery penetration in China passed 60 per cent, margins face compression from second-tier manufacturers, and a company locked into hardware sales needs a new growth vector. The pattern of securing critical links across a value chain, from upstream materials through midstream power equipment to downstream channels, is structurally identical to how CATL captured the lithium battery chain. But this is framework extension after the fact, not prediction. The original working hypotheses did not foresee data centres as the destination, did not consider cross-industry value-chain migration as a category of exit from overcapacity, and did not anticipate that a battery company would invest alongside a frontier AI lab. The miss reveals that the exit paths from state-driven overcapacity are more diverse than the framework's initial categories assumed. CATL is replaying the battery playbook in a domain the framework left blank. (CATL's data centre playbook)

The speed of the interface war

The early analysis discussed common misreadings of Chinese corporate behaviour and cautioned against over-interpreting commercial decisions through a geopolitical lens. It said nothing about super apps versus AI agents, nothing about the contest for the conversational interface layer, and nothing about what would become the most structurally consequential competitive dynamic of the period that followed.

The wrong assumption was one of omission. The framework treated the application layer as a venue for iteration and competition, shaped by the layered market and the ripen-then-glut cycle. It did not identify a specific mechanism by which platform incumbents would use protocol power to determine which AI agents could access which services, or by which new entrants would attempt to bypass that control.

In December 2025, ByteDance launched Doubao Phone Assistant, the GUI Agent product described in Principle 6. WeChat shut it down within days. By June 2026, Tencent had reversed course selectively, opening a controlled A2A channel to phone manufacturers while keeping execution inside WeChat's environment. Tencent's stock rose 10.5 per cent on reports that it was testing a WeChat-native AI agent. The interface war arrived hotter and faster than anything the framework's original framing anticipated.

The original working hypotheses contained no analysis of platform gatekeeping as a structural force. It had no discussion of GUI Agent versus API cooperation as competing access models, no assessment of how WeChat's 1.4 billion users would function as a chokepoint in the agent economy, and no anticipation that the question of who controls the conversational interface would become the single most capital-intensive competitive contest in Chinese technology during 2026. This gap is the primary reason the 2026 Edition adds a sixth principle. The interface war was not a minor update to an existing category. It was a structural force the framework failed to see.

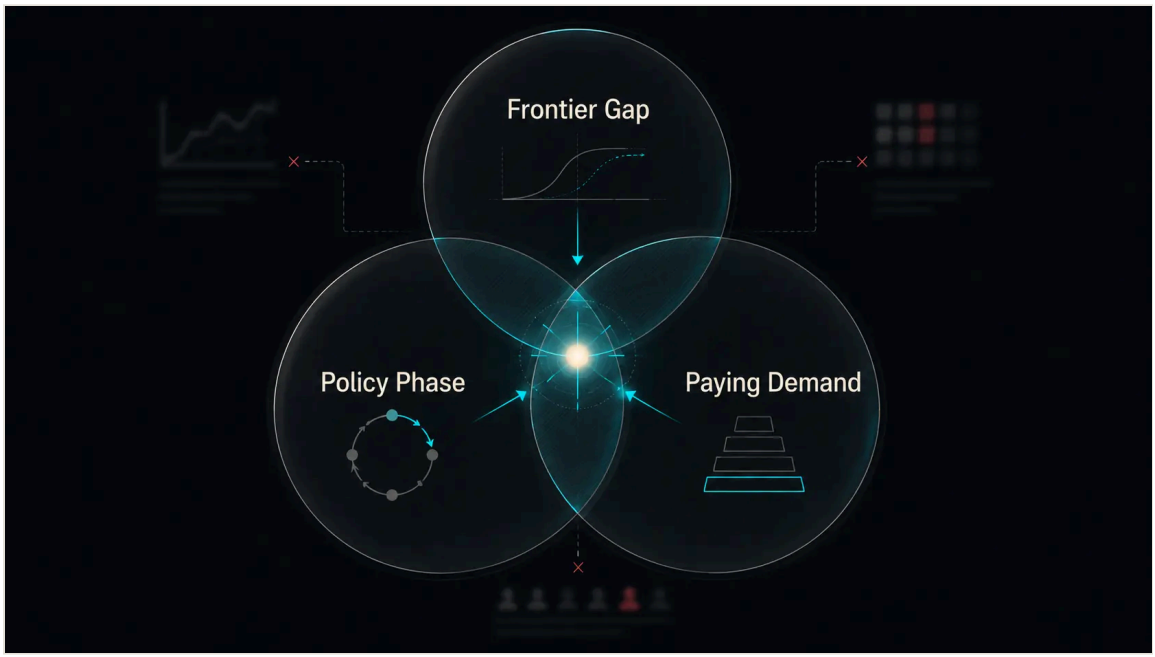
Scoring the ledger

The original working hypotheses identified the right structural forces in AI chip substitution, robotics hype, and the gap between announcements and deployment readiness. It missed the speed and shape of two developments that proved consequential: CATL's cross-industry expansion into data centre energy and the emergence of the interface war as a distinct competitive arena. The cases above are illustrative, not exhaustive. They were chosen because each maps directly to a principle's core claim. The full record is public: every essay behind this framework remains timestamped in the archive, open to the same scoring applied here.

Part III: How to Use This

The three-lens test

Each principle above describes one structural force. Applying them requires reading any sector, company, or policy development through three lenses simultaneously. Omitting any one of the three produces a distorted picture.



The first lens is global technology state. Where does a given technology sit on its maturity curve, and what is the gap between China's capability and the current global frontier? This lens prevents overestimating a domestic advance that remains several generations behind state of the art. Applied to AI chips: a domestic GPU achieving 80 per cent of a 2020 Nvidia architecture occupies a different competitive position from one matching current-generation performance, even if both qualify as good enough for their respective market segments.

The second lens is the state's policy position within the ripen-then-glut cycle. Is the government still pumping capital into the sector, or has overcapacity already triggered the exit phase? This lens determines whether a company's growth reflects genuine demand or

policy-driven acceleration that will eventually reverse. Applied to humanoid robotics: the current wave of record-setting orders arrives during the ripening phase, when state capital flows freely and procurement announcements carry political as well as commercial incentives. The same order book would read differently if it appeared after consolidation had begun.

The third lens is layered market demand. Which tier of the market is the product actually serving, and does demand persist without subsidy? This lens separates commercial traction from policy-supported demonstration projects. Applied to AI toys: products selling hundreds of thousands of units through consumer e-commerce channels at unsubsidised prices represent a different quality of demand signal from government-procured pilot deployments.

All three lenses are required because each corrects for the others' blind spots. Technology state without policy context misses the role of state capital in compressing timelines. Policy position without demand analysis misses whether the market the state is building will sustain itself. Demand analysis without technology state misses whether the product can improve fast enough to retain the customers it acquires.

Real innovation versus policy arbitrage

Within any sector, distinguishing between companies that have built durable capability and those riding temporary policy tailwinds requires three tests.

First, does demand survive without subsidy? Strip away government procurement preferences, state-fund investment, and policy-mandated adoption targets. If the remaining demand base cannot sustain the company's unit economics, the business is a policy arbitrage position, not a technology company.

Second, is the technology path dependent on external access that could be revoked? A company whose core product relies on components, manufacturing processes, or software ecosystems controlled by parties outside China's policy reach faces a constraint that no amount of domestic capital can eliminate on a short timeline. This does not disqualify the company, but it changes the risk profile and the appropriate valuation framework.

Third, do unit economics hold at scale? A product that works at pilot volume but requires subsidised inference, below-cost hardware, or loss-making pricing to reach commercial adoption has not yet crossed the threshold from demonstration to business. The test is whether the company can deliver its product profitably at 10 times its current volume without external support.

What would prove this wrong

> A framework that cannot be wrong is not a framework. It is a slogan.

A framework earns credibility by specifying the conditions under which it would require revision. Each principle below includes one observable test. If the stated condition materialises, the principle needs to be downgraded, rewritten, or abandoned.

Principle 1, capital structure beats technology

If a Chinese AI chip or model company reaches sustainable profitability at scale without state capital backing, relying entirely on private venture funding and commercial revenue, the principle's claim that capital structure is the primary survival determinant would need qualification.

Principle 2, the state is the market's architect

If a major technology sector completes its development cycle with private capital as the dominant funding source and state funds playing a minority role, the principle overstates the state's architectural function. The sector to watch is frontier AI models, where DeepSeek's trajectory suggests a path less dependent on direct state investment than the principle assumes.

Principle 3, the layered market is the durable advantage

If a single company captures more than 60 per cent market share in any major Chinese technology category and sustains it for two consecutive years, the winner-take-all dynamic the principle argues against would have a live example. The structural persistence of layered demand would be in question.

Principle 4, pipes beat apps

If consumer-facing AI application companies achieve stable profitability with margins consistently exceeding those of the infrastructure providers underneath them, the value-accrual thesis reverses. The companies to track are ByteDance's Doubao and Alibaba's Quark: if either builds a sustainably profitable consumer AI business that outearns its cloud division, the principle requires revision.

Principle 5, parallel substitution, not catch-up

If domestic AI chips cannot narrow the training efficiency gap to within 2 times Nvidia's cost-performance ratio within the next product cycle, the parallel stack remains confined to inference and lower-value workloads. Parallel substitution would then describe a partial strategy rather than a comprehensive alternative.

Principle 6, the interface war decides the next decade

If operating-system-level AI agents from device manufacturers successfully bypass super-app gatekeeping and provide end-to-end service access without platform permission, the protocol power thesis weakens. The test case is whether Huawei's HarmonyOS or Xiaomi's HyperOS can route transactions, payments, and social interactions through system-level agents without requiring WeChat's A2A authorisation.

Closing

This is the 2026 Edition of The China Tech Field Guide. It will be re-scored and re-issued annually, each July. The six principles, the evidence record, and the falsification conditions will be updated against whatever the following 12 months provide. Principles that hold up will be sharpened. Principles that break will be replaced. The 2027 Edition will carry a longer track record, a more demanding audience, and the accumulated weight of being wrong in public.

The analysis was conducted in public. This framework consolidates it for the first time. Every principle links to the published essay from which it was derived. The scored ledger in Part II applies the same standard to hits and misses.

This Field Guide is re-scored and re-issued every July. The analysis between editions, company-level breakdowns, earnings deep dives, and the running record of hits and misses, appears in Hello China Tech: **hellochinatech.com**

Further reading

Essays referenced in this Field Guide, grouped by principle.

Principle 1: Capital structure beats technology

- A Spike, Not a Pyramid: China's State-Engineered AI Boom — How state capital reshaped AI venture funding distribution
- Good Enough to Matter: How Cambricon Redefined China's AI Chip Game — Cambricon's revenue breakout and the economics of forced innovation
- China's AI Startups Enter Their Darwin Moment — Three survival paths for startups squeezed by big tech and capital scarcity

Principle 2: The state is the market's architect

- The $21B Bet: Inside China's Most Successful Government VC Story — The Hefei Model from BOE to NIO to CXMT

Principle 3: The layered market is the durable advantage

- China's Practical AI Play: The Case for Right-Sized Intelligence — Why smaller models find paying markets in China's layered demand structure
- China Shipped Millions of AI Toys While the West Was Still Debating — How supply chains, cheap inference, and layered demand created a $3.5bn category

Principle 4: Pipes beat apps

- 140 Trillion a Day: The Token Economy China Is Building — Token consumption as economic indicator and the cloud infrastructure race
- The Token Reckoning — Why AI coding plan economics point toward token efficiency over token price
- Samsung Lost China's Living Rooms, Memory Became the Prize — Value migration from the consumer layer to upstream memory fabrication
- China Builds the Plumbing, Not the Agent — Baidu's DAA metric and the infrastructure layer for China's agent economy

Principle 5: Parallel substitution, not catch-up

- The Metax Paradox: Transparency, Deficits, and the Investment Case for China's AI Chips — What a 300-page regulatory filing reveals about the chip gap and the security premium
- DeepSeek Is Turning China's Chip Problem Into a Compute-Shaping Bet — How model architecture reshapes hardware requirements
- China's AI Chips Enter the Training Stack — Domestic chips move from inference to frontier-scale training

Principle 6: The interface war decides the next decade

- The Billion-Dollar Bonfire: China's AI War and the Creation of the 'Do-Engine' — The defensive spending logic behind the conversational OS race
- The New Interface War: Why ByteDance Had to Launch an Agent No Platform Would Approve — ByteDance's GUI Agent gambit and the structural constraints behind it
- WeChat Becomes the Gatekeeper for China's AI Agents — How Tencent's A2A framework defines AI agent access in China

Part II: Evidence

- Billion-Dollar Bets or Paper Orders: China's Humanoid Robot Rush — Order quality hierarchy in China's humanoid robotics sector
- Galaxy General and the Chip Bag Test — What a convenience store reveals about the brain-first thesis in embodied AI
- CATL Replays the Battery Playbook — CATL's vertical energy chain for AI data centres