

Asia Pacific

Aesop or Zeno?

AI competition and the economics of staying ahead

October 2026

Two fabled tortoises offer different ways to think about AI. Aesop's eventually beats the hare. Zeno's remains tantalisingly ahead of Achilles: each time he reaches the place it occupied, it has moved a little farther on. One story ends in an overtaking; the other presents a gap that never quite closes.

Chinese AI laboratories have made this an unexpectedly practical distinction. Their pursuit of US frontrunners such as Anthropic and OpenAI has been ferociously inventive. Restricted access to advanced chips has encouraged designs and training methods that make scarce computing resources and capital go further. The question is whether they are catching up or simply reaching places from which the leaders have already moved on.

The gap is closing. SemiAnalysis's August study found that leading Chinese open models were matching capabilities demonstrated by OpenAI and Anthropic just five or six months earlier. In its 2022-24 comparisons, Chinese models had taken roughly eighteen months to catch up. The frontier continues to move, and the leaders' unpublished models may be further ahead. But competitors are reproducing their breakthroughs more quickly, leaving less time to earn a return before similar capabilities become available elsewhere.¹

For investors, there is a further question. How far ahead must a laboratory remain, and for how long, to justify the cost of getting there? The answer matters well beyond the companies selling AI models. It reaches the manufacturers of the chips, memory, circuit boards and power supplies that make those models possible, and the widening circle of investors and lenders financing the infrastructure. In

August, The Wall Street Journal put major technology companies' off-balance-sheet spending commitments, including future datacentre leases and purchases, at roughly US\$3 trillion. The duration of such commitments matters as much as their size: expenditure stretches over years, while a technological advantage may erode within months.²

Several leading Chinese laboratories continue to publish detailed technical accounts of their work. That openness gives us a chance to examine how their progress is being achieved, as well as the results.

Some of the most interesting progress in China has come from questioning what a model really needs. DeepSeek has found ways to compress information from earlier parts of a conversation so that it can be retrieved from only a small cue, cutting the cache-memory requirement of its V2 model by more than 90% relative to its predecessor. In V3, its mixture-of-experts architecture applies a similar principle to computation: although the model contains 670 billion parameters, only about 37 billion are activated for each token. It also uses lower-precision numbers selectively, reducing the resources needed for some calculations while retaining greater precision where errors would otherwise accumulate. Across each of these innovations, the underlying idea is the same: what looks like a hard computational requirement may instead be an expensive consequence of how the system was originally designed.^{3,4}

Other advances relate to how models learn. Distillation allows a student model to learn from a teacher's reasoning. The important economic distinction is straightforward:

discovering a capability and learning to reproduce it need not have the same cost. The follower may benefit from the transfer of knowledge without paying for every unsuccessful experiment along the way.

The launch of Anthropic's Fable 5.1 made the value of that teaching material unusually visible. Karo Zieminski's comparison of English- and Chinese-language coverage of the new release found different priorities: English reports foregrounded pricing and safety, while Chinese results gave greater prominence to anti-distillation, access to reasoning and relay services. Anthropic had strengthened protections against extracting its reasoning traces, the intermediate working from which another model can learn. News reports of the same release looked different depending on whether the reader wanted to use the teacher or learn to reproduce its abilities.^{5,6}

Recent work on multi-teacher on-policy distillation (MOPD) extends this idea. Specialist teachers are trained in different tasks, and the student learns from the teachers' feedback and guidance on its attempts to do the task itself. The aim is to combine their strengths without learning one subject at the expense of another. The June paper describing this approach reports its use in Xiaomi's MiMo-V2-Flash. There is considerable ingenuity in the teaching process as well as in the machinery being taught.⁷

These teachers can be a laboratory's own models; distillation need not involve copying a competitor. Nor does learning from another laboratory mean remaining permanently behind it. Researchers can use distillation to acquire existing capabilities, then develop new ones through further training and research of their own. Today's follower could make tomorrow's breakthrough.

The case for a persistent lead rests instead on the difficulty of discovering the next capability. A laboratory that repeatedly makes new discoveries could maintain that valuable gap. Better models could also help its researchers conduct the work that produces the next generation. Access to computing capacity, useful customer feedback and strong research teams may sustain that process. Each advantage needs examining; none should be assumed permanent.

There is, however, a third possibility between overtaking and endless pursuit. The follower may never catch the leader but still weaken the economics of leading.

A cheaper model might perform most of a customer's work adequately while the best model retains an advantage on the hardest tasks. Technical leadership would survive, but the market willing to pay a substantial premium could be smaller than expected. A laboratory could keep winning the model wars but struggle to earn enough from the victory.

The opposite is also possible. A modest improvement in reliability might make the difference between completing a valuable task and leaving it unfinished. If a more capable model saves hours of expensive human supervision, a higher price per token may still produce a lower cost per successful job. Distribution, integration into customers' work and dependable service also matter. Benchmark scores alone cannot tell us what the lead is worth.

The frontier laboratories are also competing on cost. Released as we were drafting this piece, Anthropic's Claude Opus 5.5 and OpenAI's GPT-6 Sol and Luna brought lower prices alongside reported improvements in capability and efficiency. Cheaper models and adjustable levels of reasoning allow customers to match expenditure to the difficulty of the task while staying with the same provider. This may help the leaders retain routine work that seemed most vulnerable to challenge from the open weight models, as well as the hardest assignments. The financial consequences depend on whether prices fall because delivery costs have declined, margins are being sacrificed, or both. Access to capital also affects how long laboratories can sustain aggressive pricing while funding further research. Capability, pricing and financing need to be understood together.^{8,9}

Falling prices need not prevent attractive returns if costs fall faster or the market expands sufficiently. More efficient models could make intelligence cheap enough to stimulate greater overall demand for computing: the possibility captured by Jevons Paradox. Yet growing usage alone does not establish that every provider can honour its commitments. If returns disappoint, financing can become harder and spending plans can change. Enormous long-term growth in compute demand can coexist with an uncomfortable investment cycle along the way.

Our Asia portfolios participate in several parts of this system. Across them, holdings include memory manufacturers, TSMC, Silergy, MediaTek, Lite-On, Tripod and Realtek. Their businesses differ substantially, as do the opportunities and risks we are taking through them.

TSMC manufactures chips for a range of designers; our memory holdings supply another essential component of computing. Their manufacturing expertise involves difficult physical processes and the ability to deliver reliably at scale. Reproducing those capabilities is a different problem from learning a model's behaviour. But technological importance does not settle investment value. We still need to understand the price paid for scarce capacity, the supply being built and the potential returns when conditions become less favourable or competitors become more capable.

China is pursuing a parallel catch-up in hardware. In September, CXMT announced mass production of its fifth-generation DRAM platform, alongside new memory products for smartphones and portable electronics. This is a different achievement from producing the most demanding high-bandwidth memory at competitive yields. But stronger competitors can affect returns through additional supply in established products well before they match every leading-edge technology. It would be inconsistent to admire Chinese laboratories' ingenuity while assuming Chinese semiconductor manufacturers cannot find ways around their constraints.¹⁰

Elsewhere, Silergy makes analogue and power-management chips within electronic systems, while Lite-On supplies power solutions, including high-voltage power racks for datacentres. MediaTek's activities include custom chip design, offering customers ways to tailor computing to their needs. Realtek designs communications and connectivity chips, and Tripod manufactures printed circuit boards. Several of these businesses serve substantial markets beyond AI. Their relevance to our portfolios extends beyond the outcome of the race between model laboratories.^{11,12,13,14}

This spread of exposure is deliberate, but it does not eliminate common risks. Suppliers at different points in the chain can depend on the same customers' spending. Efficiency gains can expand demand while reducing the amount of a particular component required for each task. Our job is to examine those effects company by company and then consider what they mean for the portfolio as a whole.

Tripod offers a useful example of the judgement we value. It declined to supply NVIDIA because demanding specifications and design changes at short notice

diluted the returns on offer. The opportunity to work for a celebrated customer was weighed against the economics of the business. Tripod was prepared to remain a step behind the leading edge where it judged the returns more attractive.¹⁵

That willingness to say no matters. The capital-allocation decisions made by our companies are part of our own investment risk. During a period of exceptional excitement, we want management teams to examine what growth requires of them: the investment, the commitments to customers and the consequences if expectations prove too optimistic. The quality of the opportunity depends on what remains for shareholders after those demands have been met.

We apply the same discipline to position sizes. Our holdings in several of the larger companies are often modest and below their respective benchmark weights. Owning a business does not oblige us to match the importance the market assigns to it. We seek to limit our exposure to concentration risks, including those embedded in the benchmark itself.

This can mean accepting relative underperformance when a small group of large companies continues to lead the market. That is a real trade-off. Our absolute-return approach places the risk of permanent capital loss ahead of the discomfort of differing from an index. An index weight cannot tell us whether an investment offers an adequate margin of safety.

Aesop and Zeno's fables are useful because they expose how much remains unsettled. We can recognise the ingenuity of Chinese laboratories without assuming they will make frontier research uneconomic. We can recognise the strengths of today's leaders without assuming their lead will always command an attractive return. Humility should influence the prices we pay, the size of our positions and the demands we place on management.

We will keep studying the technology and revising our judgements as the evidence changes. Our portfolios should allow us to participate in its development while limiting the damage if an important assumption proves to be wrong. Whichever story the race comes to resemble, our first responsibility remains the preservation of our clients' capital.

- 1 SemiAnalysis, Are Open Models Catching Up, 21 August 2026. The historical comparison covers open models more broadly, including US models; recent examples are Chinese. Benchmarks differ across the scaling, reasoning and agentic eras. [Read source](#)
 - 2 Peter Rudegeair and Peter Santilli, Why Big Tech's AI Spending Is \$3 Trillion Higher Than It Seems, The Wall Street Journal, August 2026. The estimate concerns commitments across major technology companies, not US\$3 trillion of debt owed by model laboratories alone. [Read source](#)
 - 3 DeepSeek AI, DeepSeek V2 technical report, May 2024. The reported 93.3% reduction concerns the KV cache relative to DeepSeek 67B, not total model memory. [Read source](#)
 - 4 DeepSeek AI, DeepSeek V3 Technical Report, December 2024. Mixture of experts and selective FP8 training. [Read source](#)
 - 5 Karo Zieminski, Claude Fable 5.1 What Chinese Coverage Says and English Doesn't, 3 September 2026. A comparison of coverage surfaced by 136 Chinese-language queries and an English-language search, not a representative survey of national opinion. [Read source](#)
 - 6 Anthropic, Preserved thinking changing how the Messages API handles thinking blocks to protect against distillation, September 2026. Scope and purpose of the reasoning-extraction protections. [Read source](#)
 - 7 Ma et al, MOPD Multi Teacher On Policy Distillation for Capability Integration in LLM Post Training, first submitted 29 June 2026. [Read source](#)
 - 8 Anthropic, Introducing Claude Opus 5.5, 22 September 2026. Company-reported capability, efficiency and pricing changes. [Read source](#)
 - 9 OpenAI, Introducing GPT-6 Sol and Luna, 22 September 2026. Company-reported capability improvements and API price reductions. [Read source](#)
 - 10 CXMT, CXMT Announces Mass Production of 5th Generation DRAM Technology Platform, 20 September 2026. Company announcement concerning DRAM and LPDDR5X; it does not establish commercial competitiveness in leading HBM. [Read source](#)
 - 11 Silergy, Company Profile. Analogue, power and signal-chain products and applications. [Read source](#)
 - 12 Lite-On, NVIDIA GTC 2026 announcement, 17 March 2026. Data-centre power solutions. [Read source](#)
 - 13 MediaTek, Data Center Solutions. Custom chip design and related engineering services. [Read source](#)
 - 14 Realtek, Communications Network ICs. Product portfolio. [Read source](#)
 - 15 FSSA research and presentation material on Tripod, September 2026. The NVIDIA example reflects the investment team's account of the company's decision.
- * Reference to specific securities (if any) is included for the purpose of illustration only and should not be construed as a recommendation to buy or sell the same. All securities mentioned herein may or may not form part of the holdings of First Sentier Group's portfolios at a certain point in time, and the holdings may change over time.

Important Information

The information contained within this material is generic in nature and does not contain or constitute investment or investment product advice. The information has been obtained from sources that First Sentier Group believes to be reliable and accurate at the time of issue but no representation or warranty, expressed or implied, is made as to the fairness, accuracy, completeness or correctness of the information. To the extent permitted by law, neither First Sentier Group, nor any of its associates, nor any director, officer or employee accepts any liability whatsoever for any loss arising directly or indirectly from any use of this material. The information herein is for information purposes only; it does not constitute investment advice and/or recommendation, and should not be used as the basis of any investment decision. Some of the funds mentioned herein are not authorised for offer/sale to the public in certain jurisdiction.

The value of investments and the income from them may go down as well as up and you may not get back your original investment. Past performance is not necessarily a guide to future performance. Please refer to the offering documents for details, including the risk factors. Reference to specific securities (if any) is included for the purpose of illustration only and should not be construed as a recommendation to buy or sell the same. All securities mentioned herein may or may not form part of the holdings of First Sentier Group's portfolios at a certain point in time, and the holdings may change over time.

This material has been prepared for general information purpose. It does not purport to be comprehensive or to render special advice. The views expressed herein are the views of the writer at the time of issue and not necessarily views of First Sentier Group. Such views may change over time. This is not an offer document, and does not constitute an investment recommendation. No person should rely on the content and/or act on the basis of any matter contained in this material without obtaining specific professional advice. The information in this material may not be reproduced in whole or in part or circulated without the prior consent of First Sentier Group. This material shall only be used and/or received in accordance with the applicable laws in the relevant jurisdiction.

In Hong Kong, this material is issued by First Sentier Investors (Hong Kong) Limited and has not been reviewed by the Securities & Futures Commission in Hong Kong. In Singapore, this material is issued by First Sentier Investors (Singapore) whose company registration number is 196900420D. This advertisement or material has not been reviewed by the Monetary Authority of Singapore. First Sentier Group, First Sentier Investors, FSSA Investment Managers, Stewart Investors, RQI Investors and Igneo Infrastructure Partners are the business names of First Sentier Investors (Hong Kong) Limited. First Sentier Group (registration number 53507290B), First Sentier Investors (registration number 53236800B), FSSA Investment Managers (registration number 53314080C), Stewart Investors (registration number 53310114W), RQI Investors (registration number 53472532E) and Igneo Infrastructure Partners (registration number 53447928J) are the business names of First Sentier Investors (Singapore).

First Sentier Investors (Hong Kong) Limited and First Sentier Investors (Singapore) are part of the investment management business of First Sentier Group, which is ultimately owned by Mitsubishi UFJ Financial Group, Inc. ("MUFG"), a global financial group. First Sentier Group includes a number of entities in different jurisdictions.

To the extent permitted by law, MUFG and its subsidiaries are not responsible for any statement or information contained in this material. Neither MUFG nor any of its subsidiaries guarantee the performance of any investment or entity referred to in this material or the repayment of capital. Any investments referred to are not deposits or other liabilities of MUFG or its subsidiaries, and are subject to investment risk, including loss of income and capital invested.