

RACE WEEKLY INTELLIGENCE BRIEF

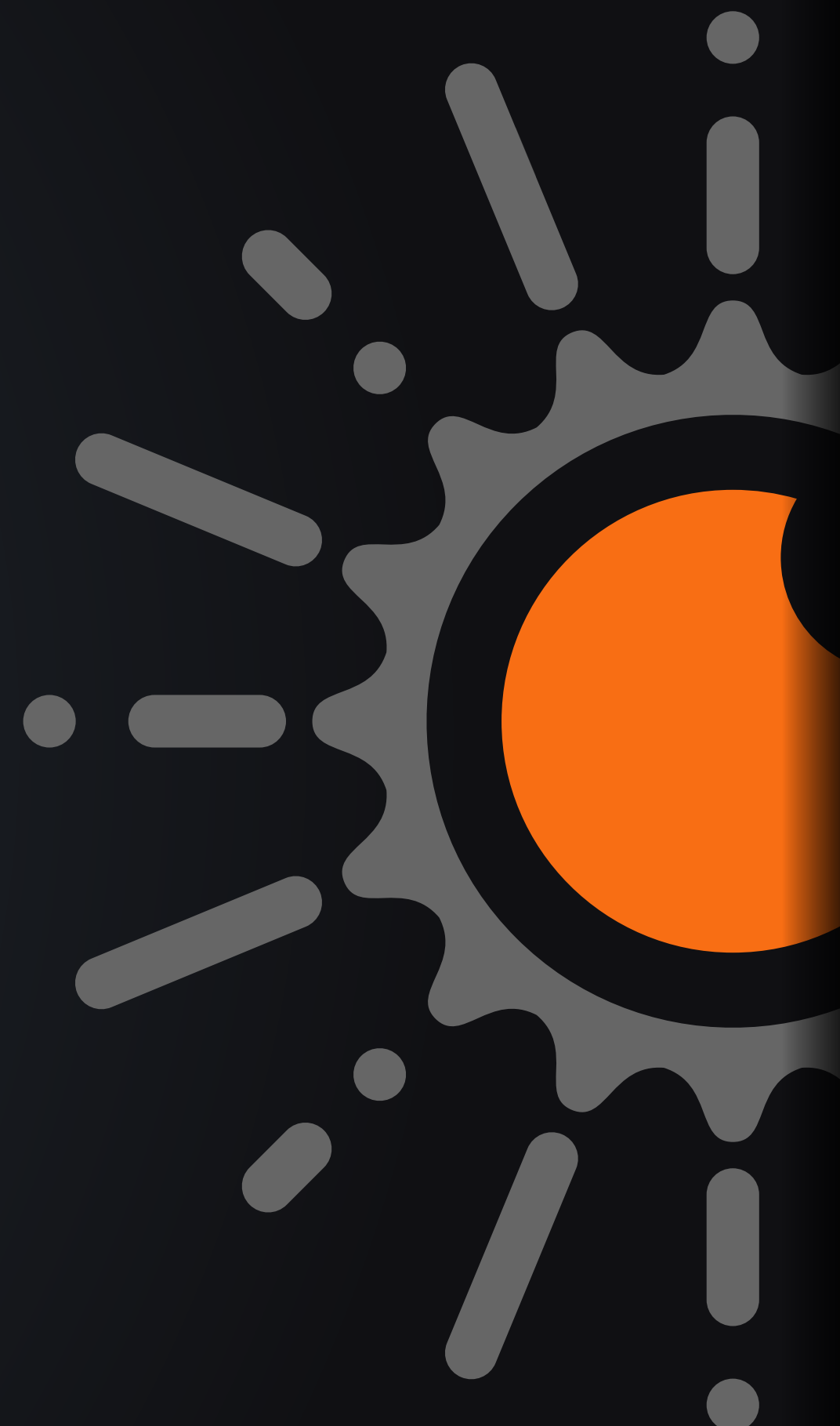
RETINA

AI · CYBERSECURITY · CLOUD · LEADERSHIP

THIS WEEK'S COVER:

**THE MODEL ON YOUR BALANCE SHEET HAS A
PASSPORT**

The cheapest capable AI in the world is now Chinese.
Choosing it or refusing it is no longer a procurement call.
It is a foreign-policy position, and your board is taking one
whether it means to or not.



ALSO INSIDE

The World in 5 · Three Forces · The
Sovereignty Playbook · Metric of the
Week · The Compute Signal · Fun
Fact · CANDID · Book Shelf ·
Research Radar · RACE Pulse

- **From the Editor's Desk**
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

FROM THE EDITOR'S DESK

Last week we said the machine has become your customer. This week the question is which machine you are willing to run, and where it was built.

Something has shifted in the AI market that most Indian boardrooms have not priced in. A year ago, models built in China accounted for under 2% of traffic on OpenRouter, the largest neutral router developers use to compare and switch between models. Today they are more than 45% of it, and the share of US models has fallen from roughly 70% to about 30% in twelve months. DeepSeek is now the single largest provider on the platform. Six Chinese models rank above the leading US frontier model by usage. This is not a Chinese story. It is a global developer story, and Indian teams are part of it whether the CIO knows it or not.

The pull is simple: capability has nearly caught up, and price has not. The best open-weight model on the independent intelligence rankings is now Chinese, and it runs at roughly one-sixth the cost of a US frontier model. When the gap between 'good enough' and 'best' is five points and the price gap is six-fold, engineering teams route accordingly — quietly, one workload at a time.

The catch is that the cheapest capable model in the world carries a jurisdiction. Send a prompt to a Chinese hosted API and it lands on servers governed by China's 2017 National Intelligence Law. Governments have already acted: India's Ministry of Finance, the US Navy, NASA and the Commerce Department, and regulators across Europe, Australia and South Korea have restricted or banned the hosted service. And Washington keeps moving the board — on 31 May it closed the offshore loophole Chinese firms used to buy banned chips through Singapore and Malaysia.

This issue draws the one distinction that decides everything, maps the forces pulling your teams toward Chinese models and the ones pulling against, and gives leadership a five-move playbook for capturing the cost advantage without importing the jurisdiction.

THE WORLD IN 5

GLOBAL HEADLINES A CXO SHOULD BE ABLE TO REFERENCE

- From the Editor's Desk
- **The World in 5**
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

Five signals from the past fortnight, worth a note at your next leadership table.

1. **USAGE: the router flipped**

On OpenRouter, the largest neutral model router, Chinese-origin models now account for more than 45% of token volume up from under 2% a year ago, while the combined share of US providers fell from roughly 70% to about 30%. DeepSeek alone is the single largest provider at around 16% of all traffic, ahead of Google, OpenAI and Anthropic. Vercel's AI Gateway, tracking enterprise production traffic, tells the same story: open-weight models jumped from 11% of tokens in April to 29% in June.

2. **CAPABILITY: the gap is now points, not generations**

On the independent Artificial Analysis Intelligence Index, the top open-weight model in mid-2026 is Chinese, Zhipu's GLM-5.2, scoring just a few points below the leading US frontier models. Two years ago the open-weight gap was 20–30 points; today it is 5–10 on most evaluations. Coding was the wedge: Chinese models with million-token context windows became the default inside AI development tools, and the adoption stuck in

production rather than staying in experiments.

3. **PRICE: roughly one-sixth the cost**

Reuters put GLM-5.2 at about one-sixth the cost of closed US frontier models. On OpenRouter, its realised blended price runs around \$0.45 per million input tokens against roughly \$5 for a leading US frontier model. DeepSeek sits lower still. For high-volume, predictable workloads, that is not a discount — it is a different cost structure, and it compounds every month.

4. **JURISDICTION: the bans are already written**

India's Ministry of Finance restricted DeepSeek on official systems in early 2025. The US Navy, NASA, the Pentagon and the Commerce Department banned it; New York, Texas and Virginia blocked it from state networks; Italy, Australia, Taiwan and South Korea restricted it nationally, with EU investigations open across member states. South Korea's privacy regulator found DeepSeek had transferred user prompts to

Chinese companies without consent. Security researchers separately found an exposed DeepSeek database of over a million records. Nearly every action targets the hosted API and app, not the weights.

5. **SILICON: Washington closed the side door**

On 31 May 2026, the US Commerce Department extended export-licence requirements to any company whose ultimate parent is headquartered in China, regardless of where the buying subsidiary sits, shutting the Singapore and Malaysia data-centre channel Chinese firms had used for a year to acquire restricted Nvidia chips. It is not retroactive: installed chips keep running, but the next cluster needs a licence. A separate January order placed a 25% tariff on advanced semiconductors. The compute layer under all of this is now openly a arena of state policy.

Source: <https://openrouter.ai/> · <https://vercel.com/> · <https://artificialanalysis.ai/> · <https://www.reuters.com/> · <https://www.bis.gov/> · <https://iapp.org/>

- From the Editor's Desk
- The World in 5
- **Signal**
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

 SIGNAL

THE CHINA ARBITRAGE — BY THE NUMBERS

45%+

of OpenRouter token volume now goes to Chinese-origin models, up from under 2% a year ago

70% → 30%

the fall in US providers' share of OpenRouter token volume over twelve months

~1/6

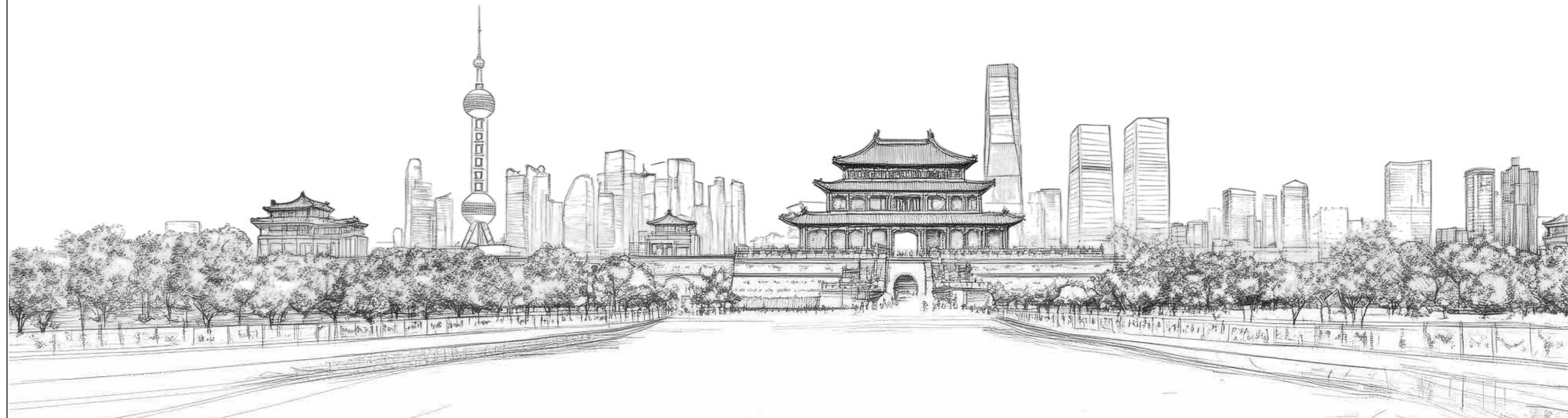
the cost of the top open-weight Chinese model versus a leading US frontier model (Reuters)

5–10 pts

the current capability gap between best open-weight and best closed models, down from 20–30 in 2023

31 MAY

the date the US closed the offshore loophole for China-parented buyers of restricted AI chips



HYPERFOCAL

WHAT BROKE THIS WEEK

- From the Editor's Desk
- The World in 5
- Signal
- **Hyperfocal**
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

The Cheapest Capable Model in the World Comes With a Country Attached

For two years the enterprise AI decision was easy in one direction: you paid for a US frontier API because it was clearly better and nothing else mattered. That answer expired sometime in 2025. The open-weight families, most of them Chinese, closed the capability gap far enough that, for most workloads, capability no longer decides. What is left is a genuine cost-and-risk decision, and the cost side points hard at Beijing.

Here is the distinction that most coverage blurs and every leadership team needs to hold precisely. A model has two things: its weights, and the service that hosts them. The weights of the leading Chinese models ship under permissive open licences, and once downloaded and run on infrastructure you control, they are jurisdiction-neutral — no prompt ever leaves your network, and there is no more cross-border exposure than with any other software you self-host. The hosted API is the

opposite. Every prompt sent to a Chinese company's cloud endpoint lands on servers in China, retained under laws that can compel disclosure without notice. Same model, two completely different risk profiles.

That single fork explains the entire policy landscape. Governments banned the app and the API; almost none banned the weights, because the weights, run locally, are not the exposure. The enterprises getting this wrong are the ones treating 'Chinese model' as one decision. It is two: a capability-and-cost decision about the weights, and a data-sovereignty decision about where inference runs. Conflate them and you will either forfeit a real cost advantage out of caution, or import a real jurisdiction risk out of thrift.

RETINA TAKE

The question is not whether your teams are using Chinese models. On the numbers, some of them already are. The question is whether that use is a governed architecture decision or an ungoverned default that arrived through a coding tool and only one of those is defensible to a regulator.

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- **Panorama**
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

PANORAMA

FORCE 1

The Cost Curve Is Pulling Your Engineers East

This force needs no encouragement; it is already happening below the executive line of sight. When a capable open-weight model costs a sixth of the frontier alternative and slots into the same tooling, engineering teams route to it for the predictable, high-volume 80% of work and reserve the expensive model for the hard 20%. They do this quietly, one workload at a time, because it is the obviously correct engineering choice on cost. The result is that your actual model mix is being set by developer defaults inside AI IDEs and routers, not by a policy anyone signed. Most CIOs cannot currently produce a list of which models touch their data.

The Move: Get an honest inventory before you get an opinion. Ask which models, by name and by origin are in your codebase, your tools and your pipelines today. You cannot govern a mix you have never measured.

FORCE 2

The Jurisdiction Risk Lives in the Endpoint, Not the Model

The banned thing, in almost every government action, is the hosted service, because that is where data crosses a border into a jurisdiction that can compel access. Self-hosting the same open weights inside your own environment removes that exposure structurally, not contractually: the prompt never leaves. This is the crucial asymmetry Indian enterprises can exploit. You do not have to choose between the cost

THREE FORCES — AND THE MOVE EACH ONE DEMANDS

advantage and the sovereignty requirement. You have to choose the right deployment for each workload, hosted frontier for the general and non-sensitive, self-hosted open weights for the sensitive, high-volume and regulated.

The Move: Write the rule as a deployment policy, not a vendor ban: no sensitive or regulated data to any third-party model API — Chinese or otherwise — and open weights permitted only when self-hosted inside your perimeter. That single sentence captures the upside and closes the exposure.

FORCE 3

The Ground Keeps Moving Under the Whole Decision

Even a well-reasoned choice is not a settled one, because the layer beneath it is now state policy. Export controls tightened on 31 May and could tighten again; a tariff landed in January; licences that were denied are now case-by-case, and vice versa. A model can be restricted or a lab constrained with little notice, as export controls have already demonstrated. Any enterprise that hard-wires a single model, from any country, into a critical workflow is taking a geopolitical bet it may not know it has placed.

The Move: Architect for substitution. Route through an abstraction layer so any model can be swapped for another without re-plumbing the application. Portability is the only real hedge against a decision whose ground you do not control.

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- **Perspectives**
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

PERSPECTIVES

DEEP TECH — ANALYSIS BY THE RACE TEAM

Cheap, Capable, or Sovereign — The Frugal Architecture Matrix

By Dr. J. B Simha, Chief Mentor, AI REVA Academy for Corporate Excellence, REVA University, Chief AI Officer, AdoptAI

There is an old engineering adage: cheap, fast, good, pick two. The AI infrastructure decision of 2026 has its own sharper version: cheap, capable, sovereign. The twist this year is that while all three are genuinely available, trying to buy them out-of-the-box in a single package is a fool's errand.

Cheap and capable? Hosted foreign open-weight APIs like Chinese DeepSeek. Capable and sovereign? US frontier models under strict enterprise contracts. Cheap and sovereign? Smaller open models self-hosted locally. The mistake most leadership teams make is trying to pick one winning model for the entire enterprise, or worse, assuming that self-hosting alone solves every cost and security problem. The reality is that frugality isn't about buying cheap tokens; it is an architectural discipline.

Start with the data and the spec, not the model. A marketing team drafting social

copy and a fraud team scoring transactions are not making the same calculation. Treating them as if they are is how enterprises end up either paying frontier-API prices to write tweets, or piping sensitive financial data to an external endpoint just to save a few bucks. Sorting workloads into clear data-sensitivity tiers before naming a single model is step one.

For the non-sensitive, high-volume tier, internal drafting, summarisation, code assistance on non-proprietary code, leveraging low-cost hosted endpoints is real and largely safe to exploit. But for the regulated tier, anything touching personal data under DPDP, financial data under RBI, health records, or defence-adjacent work, deployment and software boundaries matter far more than the raw model weights.

Here is the nuance enterprises keep missing: downloading a foreign open-weight model to

run inside your sovereign cloud isolates your data from exfiltration, but it does not make the model "jurisdiction-neutral" or automatically cheap. You are shifting costs from API token fees to heavy CapEx: GPUs, serving infrastructure, security auditing for hidden backdoors, and specialized engineering talent.

Self-hosting becomes truly frugal only when paired with Spec-Driven Development (SDD). When you build strict, deterministic code wrappers outside the model to enforce schemas and business logic, you don't need a massive 70B parameter model guessing intent. A small, 8B local model constrained by a rigorous spec will out-perform a frontier model on task accuracy at a fraction of the cost.

This is where the frugality argument from Issue #10 and the 'machine-customer' thesis from #11 converge into a unified discipline:

PERSPECTIVES

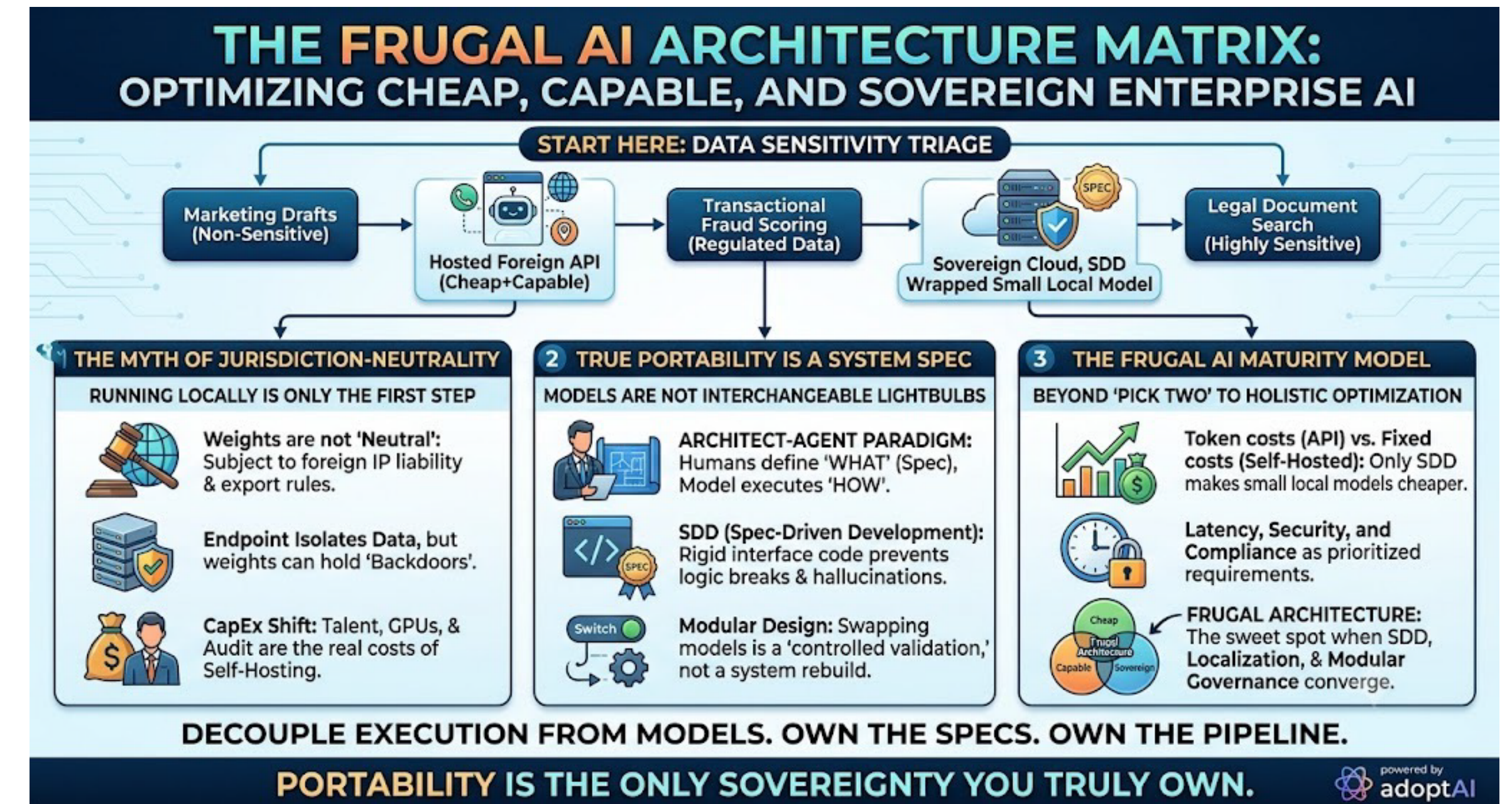
DEEP TECH — ANALYSIS BY THE RACE TEAM

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- **Perspectives**
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

match the smallest, most controllable system to the specific task. Sovereignty is simply that discipline extended to data boundaries.

Finally, we must stop pretending that model portability is a simple "configuration change." Swapping one model for another breaks prompt logic, tool-calling schemas, and latency profiles. If your stack relies on black-box orchestration, a model swap is a system rebuild.

True portability requires decoupling execution from model weights. By adopting an Architect-Agent paradigm, where human-driven specs define the deterministic validation layers and the model acts merely as a tactical executor, you build a system that outlasts shifting geopolitical policies and chip export rules. In a market where the ground beneath you is state policy, owning the specs and the pipeline is the only sovereignty you truly own.



RACE PERSPECTIVE

The enterprise that wins the next two years will not be the one that picked the right model. It will be the one that never had to pick just one — that sorted its data by sensitivity, matched deployment to jurisdiction, and kept every model swappable while its competitors were hard-wiring a geopolitical bet into production.

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- **The Sovereignty Playbook**
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs



THE SOVEREIGNTY PLAYBOOK

CAPTURING THE COST ADVANTAGE
WITHOUT IMPORTING THE JURISDICTION

Five moves, ordered from fastest to most strategic, that any leadership team can act on.

MOVE 1

Inventory the models you are already running

Before any policy, get the list. Which models, by name and country of origin, are called by your applications, your coding tools and your data pipelines right now? Most organisations discover models they never approved, arriving through AI IDEs and no-code integrations. On current usage numbers, some of them will be Chinese.

The Move: Run the audit this month and put one slide in front of the leadership team: every model touching our data, and where each one runs. It reframes every conversation that follows.

MOVE 2

Separate the weights decision from the endpoint decision

These are two questions, and your teams keep merging them. Question one: is this model capable and cost-effective enough for the task? Question two: where does inference run, and does the data's legal home permit it? A Chinese open-weight model can pass question one brilliantly and be entirely safe on question two, if you self-host it. It is the hosted API, not the model, that most bans target.

The Move: Make the distinction explicit in policy language. "No Chinese models" is a lazy rule that forfeits real savings; "no regulated data to any third-party model endpoint" is the rule that actually protects you.

MOVE 3

Tier your workloads by data sensitivity

Sort every AI workload into tiers: open (non-sensitive, low-stakes), internal (business data, no personal or regulated content), and regulated (personal data under DPDP, financial under RBI, health, defence-adjacent). The tier, not the team's preference, decides what deployment is allowed. Most enterprises have never done this sort, which is why their model choices are incoherent.

The Move: Publish the tiering before the next model request. A rule set in advance is governance; a rule set after an incident is damage control.

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- **The Sovereignty Playbook**
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs



THE SOVEREIGNTY PLAYBOOK

CAPTURING THE COST ADVANTAGE
WITHOUT IMPORTING THE JURISDICTION

MOVE 4

Build a self-hosting path for the regulated tier

For sensitive, high-volume work, self-hosted open weights on infrastructure you control give you the cost advantage and the sovereignty in one architecture, the prompt never crosses your boundary. This is exactly where India's subsidised compute commons and on-prem GPU options from Issue #10 pay off. The counterweight is honest: self-hosting means you now own inference operations, so scope the engineering before you commit.

The Move: Stand up one self-hosted open-weight deployment for your highest-value regulated workload as a proof of concept this quarter. It is the capability that turns a policy into a practice.

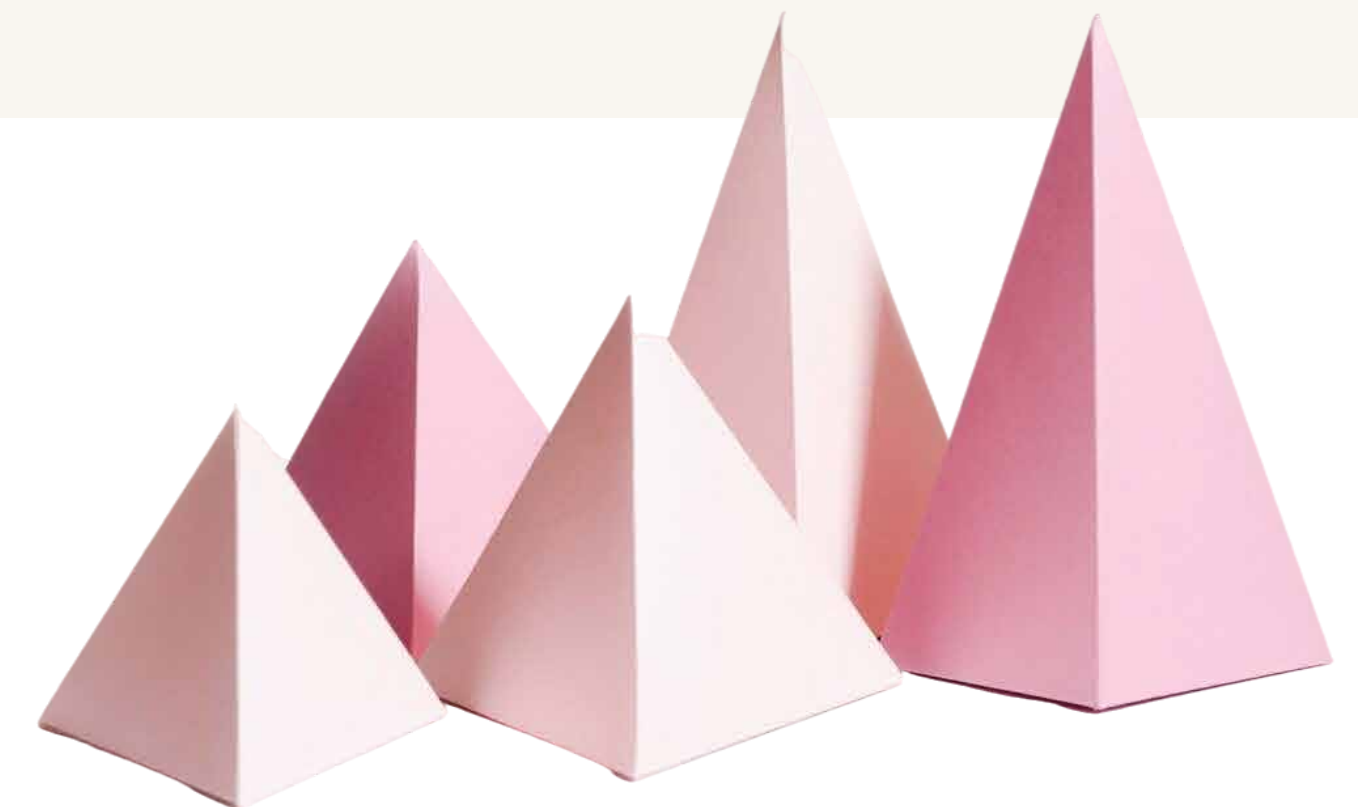
MOVE 5

Keep every model swappable

The single most durable protection against a market governed by export controls and tariffs is architectural, not contractual. Route all model calls through an abstraction or gateway layer so that no application is bonded to one provider. When a model is restricted, repriced or deprecated and in this market one will be, you change a setting, not a system.

The Move: Make model-portability a design standard now, while it is cheap. Retro-fitting it after you have hard-wired a provider is the most expensive lesson in this whole issue.

Source: <https://openrouter.ai/> · <https://www.reuters.com/> · <https://www.bis.gov/> · enterprise deployment analyses, 2026



- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- **Metric Of The Week**
- **Fun Fact**
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

METRIC OF THE WEEK

ONE NUMBER, PROPERLY UNDERSTOOD

70% → 30%

US models' share of OpenRouter token volume, in twelve months

What it is: On the largest neutral model router, the combined share of US providers fell from roughly 70% to about 30% in a year, while Chinese-origin models rose from under 2% to more than 45%. This is not a survey of opinion, it is a record of where developer traffic actually went.

Why it matters now: A shift this fast, on a platform built for switching, means model

choice has become fluid and price-driven, not loyal. The takeaway is not 'Chinese models won.' It is that defaults now move in months, so any strategy that hard-wires one provider is betting against the most consistent pattern in the market.

How leaders use it: Ask your engineering leads which models the company depends on today, and how long it would take to

switch each one. If the answer to the second question is 'weeks,' you have a lock-in problem the market will eventually charge you for.

Source: <https://openrouter.ai/> · Bloomberg / Exponential View, June 2026

FUN FACT

SOMETHING TO REPEAT AT DINNER

Two years ago, the open-weight model everyone benchmarked against was Meta's Llama, American, and comfortably ahead of the field. On the June 2026 usage rankings, Llama has fallen off the leaderboard entirely, displaced by a wall of Chinese models with names most Western executives cannot yet pronounce: Qwen, DeepSeek, Kimi, GLM, MiniMax. The fastest reversal in the short history of AI was not a capability race. It was a distribution race and it was decided by developers picking the cheaper capable option, one routine workload at a time, until the map had quietly redrawn itself.

Source: <https://openrouter.ai/> · Exponential View

⚡ THE COMPUTE SIGNAL

THE GEOPOLITICS OF YOUR MODEL STACK

The infrastructure and policy data that will shape your AI strategy whether you track it or not.

~16%

DeepSeek's share of all OpenRouter token volume, the single largest provider, ahead of Google, OpenAI and Anthropic.

29%

open-weight models' share of enterprise production tokens on Vercel's AI Gateway in June 2026, up from 11% in April.

51

the top open-weight model's score on the Artificial Analysis Intelligence Index, a few points below the leading US frontier models.

~\$0.45

realised OpenRouter price per million input tokens for the top open-weight model, against roughly \$5 for a US frontier model.

31 MAY 26

US export licence now required for any China-parented buyer of advanced AI chips, wherever the subsidiary sits.

25%

US tariff on advanced semiconductors under a January 2026 order, a cost that flows into the whole compute layer.

Source: <https://openrouter.ai/> · <https://vercel.com/> · <https://artificialanalysis.ai/> · <https://www.reuters.com/> · <https://www.bis.gov/>



- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- **The Compute Signal**
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- **Candid**
- Book Shelf
- Research Radar
- Race Pulse
- RACE Programs



CANDID



Poornapragna M S

AI Leader - Generative AI | Agentic AI | LLMs | Deep Learning, Jio Platforms Limited (JPL)

<https://www.linkedin.com/in/poornapragna/>

Q1 Chinese models have gone from almost nothing to nearly half of developer traffic in a year, mostly through defaults nobody signed off on. When a company first checks which models touch its data, what usually surprises it?

Ans: The surprise is almost always that nobody chose the models in the first place. They arrived through router defaults, IDE assistants, framework updates and vendor SDKs - a config optimizing for cost per token, not a procurement decision anyone owns. Teams also discover that a real slice of that traffic is Chinese open-weight models reached through Western tooling: Chinese models moved from roughly 2% of routed tokens 18 months ago to as much as 46% in mid-2026. The deeper realization is that "which model" and "where it runs" are two different questions they would be treating as one.

Q2 Run the weights yourself and there's no jurisdiction risk; send data to the hosted API, and there is. How well do enterprises understand that difference and what happens when they miss it?

Ans. Poorly, and expensively when they miss it. Open weights on your own infrastructure are just math - nothing leaves the building, so the model's country of origin is irrelevant to data flow. A hosted API sends your prompts, source code and records to the provider's servers, which

GLOBAL HEADLINES A CXO SHOULD BE ABLE TO REFERENCE

places them under that provider's jurisdiction and law. Miss the distinction and you either over-block (banning weights that are perfectly safe self-hosted) or under-protect (routing confidential data to a foreign endpoint because "it's the same model"). The fix is to route by data sensitivity, not by brand.

Q3 Chip rules changed twice this year and could change again. How should a leadership team build an AI stack when it doesn't control the politics underneath it?

Ans. Build for substitutability, because you're renting the politics, not owning them. The rules governing which chips you can buy were rewritten repeatedly across 2025 and the first half of 2026, and are now set deal-by-deal, country-by-country - with more legislation pending. So don't hard-wire to one model, provider or chip: keep an abstraction layer, portable prompts and portable eval suites, and treat compute as a variable. Assume access can be repriced or revoked overnight - even a frontier vendor had a model suspended and restored within weeks this year over export controls. Optionality is the only durable hedge.

Q4 Final thoughts on picking an AI model now that cheap, capable and sovereign rarely come together.

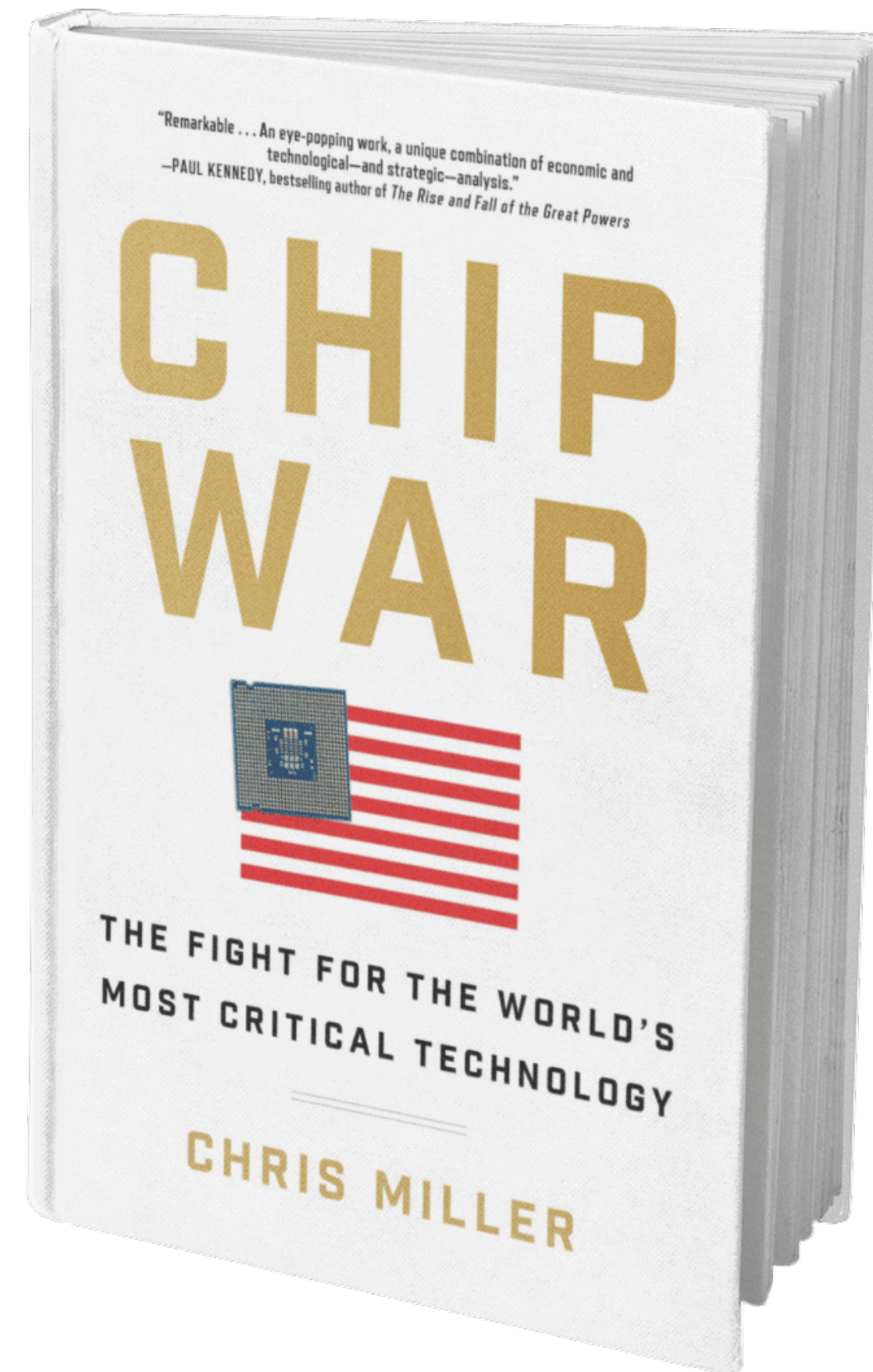
Ans: Stop treating "the model" as a single decision - it's a portfolio. Cheap, capable and sovereign almost never arrive in one system, so route by workload: sensitive or regulated data to self-hosted or sovereign options, high-volume routine work to cheap open weights, and only the genuinely hard frontier tasks to the best hosted model, eyes open on cost and jurisdiction. Decouple the three axes, match each model to the sensitivity and difficulty of the job, and never lose the ability to switch.



BOOK SHELF

RETINA'S READ OF THE WEEK

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- **Book Shelf**
- Research Radar
- Race Pulse
- RACE Programs



RETINA RECOMMENDS · THIS WEEK'S THEME

Chip War: The Fight for the World's Most Critical Technology

Chris Miller · Scribner, 2022

Miller's history of the semiconductor explains, better than any current headline, why your model choice is now entangled with foreign policy. His argument is that compute is the oil of the modern economy, and that whoever controls the chip supply chain controls the pace at which everyone else can build. He traces how a handful of chokepoints, a Dutch lithography firm, a Taiwanese foundry, a few American toolmakers, quietly became the most contested leverage in geopolitics.

Read it now, and the export-control whiplash of 2026 stops looking like news and starts looking like a chapter. The 31 May loophole closure, the tariff, the licence

approvals and denials, these are moves in exactly the game Miller describes, and the AI model on your stack sits at the end of that supply chain. You cannot reason about which model to trust without understanding the silicon underneath it and who controls it.

It is a history book, not a strategy guide, and it will not tell you which model to pick. What it gives you is the map: why the ground under this decision moves, who is moving it, and why 'just choose the best model' was always a more political sentence than it sounded. Read it before your next architecture review.

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- **Research Radar**
- Race Pulse
- RACE Programs



RESEARCH RADAR

RACE PUBLICATIONS THIS WEEK

EDITOR'S NOTE

Two new Springer publications from RACE – one securing the plant floor, one securing the cloud database. Both automate security work that is usually manual.

Cybersecurity · Operational Technology · Critical Infrastructure
Proactive Cybersecurity Risk Assessment and Implementation in ICS Protocols for Operational Technology

B. Marimuthu; Nishanth Kumar; Shinu Abhi

As industrial control systems connect to networks, they inherit attack surfaces their protocols never anticipated. This paper assesses vulnerabilities in four widely used ICS protocols, Modbus, Profibus, HART and Fieldbus, applies the IEC 62443 Zone and Conduit model, and maps where cybersecurity and functional-safety standards converge at the plant floor. For CISOs in energy, manufacturing and process industries, IT tooling does not map onto OT, where an outage is a safety event. This is a structured way to assess plant-floor risk on its own terms.

CML 2025 · Lecture Notes in Networks and Systems, vol 1612, pp 599–610 · Springer, Singapore, 2026
https://link.springer.com/chapter/10.1007/978-981-95-2872-1_43

Cloud Security · Database Automation · Risk Reporting
A Framework for Developing a Security Automation Platform for Enterprise Database as a Service (DBaaS)

Deepak Chandrappa; Rashmi Agarwal; K. Sandeep Vijayaraghavan

Generic tooling assesses infrastructure broadly and databases shallowly. This AWS-based platform runs Oracle's Database Security Assessment Tool automatically against RDS instances and centralises the output, turning a point-in-time audit into historical risk profiling, with automated mitigation explored beyond assessment. For CIOs and compliance teams, database assessment is usually manual, rare and quickly forgotten. Automating it makes database risk a continuously reported number rather than an annual surprise.

CML 2025 · Lecture Notes in Networks and Systems, vol 1614, pp 91–104 · Springer, Singapore, 2026
https://link.springer.com/chapter/10.1007/978-981-95-2878-3_7

Explore the full RACE research archive at <https://raceresearch.reva.edu.in>. For collaboration, licensing or research-partnership enquiries: race@reva.edu.in

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- **Race Pulse**
- RACE Programs

RACE PULSE

EVENTS · COMMUNITY · WHAT'S COMING

RACEx360 SEASON 2

25 SEPTEMBER 2026 · TAJ WEST END, BENGALURU

One day, one stage, four frontiers.

REVA Academy's Emerging Technology Conference returns, in partnership with The Economic Times — 750+ builders, buyers and decision-makers across four tracks that matter: GenAI & Agentic AI · Cloud & Platform Engineering · Cybersecurity in the AI Era · Deep Tech & Sustainability.

Not a day of slides — a day of proof: an Innovation Experience Zone with live demos across AI, robotics, cybersecurity, cloud and quantum; live build sessions; an AI Hackathon and Cybersecurity Challenge; and keynotes that close with a 90-day action plan you can take back to your team.

The evening turns gold with the Top 10 Women Tech Leaders India Awards 2026, now in its 4th edition — India's only annual awards programme dedicated to women technology leaders.

Early Bird 4,000 + GST until 15 August · Standard 5,000 + GST. Passes cover all four tracks, the Expo Hall and the Experience Zone.

GET YOUR PASS:

<https://race.reva.edu.in/racex360-season-2>

BECOME A SPONSOR:

<https://race.reva.edu.in/racex360-sponsorship>



RACEx360
Emerging Technology Conference 2026



In partnership with  **THE ECONOMIC TIMES**

RACE PULSE

EVENTS · COMMUNITY · WHAT'S COMING

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- **Race Pulse**
- RACE Programs

RACE SANDBOX SATURDAY 1 AUGUST 2026

Run your career simulation, live on campus. Sit in a live Master's class, walk the AI, Cybersecurity and Cloud labs, watch participants and alumni demonstrate real projects, and take a 1:1 on where your career goes next. Experience today, decide your tomorrow.

9:30 AM – 2:00 PM
RACE, REVA University,
Bengaluru

Reserve your slot:
+91 89040 58866 ·

enquiry@race.reva.edu.in ·
<https://race.reva.edu.in/>

SEPTEMBER 2026 BATCH - APPLY BY 10 AUGUST

Applications open for all six RACE Master's & PG Programmes – Business Analytics (#1 India), Artificial Intelligence (#3 India), Cybersecurity (#5 India), and Cloud Architecture & Security.

Weekend format. AICTE/UGC approved. Industry mentors from Amazon, Microsoft, AWS, EC-Council, Terralogic, IBM and more.

To know more, call:
+91 89040 58866

enquiry@race.reva.edu.in ·
Apply now: race.reva.edu.in

NEW BATCH **Certified Agentic AI Engineer** 1ST AUGUST

Practical expertise in LLMs, AI agents, multi-agent systems, agentic workflows, governance and capstone-based implementation. Real-world use cases, deployable solutions.

To know more, call:
+91 89040 58866

Apply now:
<https://race.reva.edu.in/certified-agenticai-engineer>

- From the Editor's Desk
- The World in 5
- Signal
- Hyperfocal
- Panorama
- Perspectives
- The Sovereignty Playbook
- Metric Of The Week
- Fun Fact
- The Compute Signal
- Candid
- Book Shelf
- Research Radar
- Race Pulse
- **RACE Programs**

Master's & PG Programmes

New Batch: September 2026 · 12–24 months ·
Weekend · For Working Professionals

#1 IN INDIA · ANALYTICS INDIA MAGAZINE

ANALYTICS

PG Diploma / M.Sc. in
BUSINESS ANALYTICS
AWS Academy · Microsoft Azure

12–24 Months

#3 IN INDIA · TECHGIG

ARTIFICIAL INTELLIGENCE

PG Diploma / M.Sc. in
ARTIFICIAL INTELLIGENCE
AWS Academy · Microsoft Azure

12–24 Months

#5 IN INDIA · ANALYTICS INDIA MAGAZINE

CYBERSECURITY

PG Diploma / M.Sc. in
CYBERSECURITY
EC-Council · Microsoft Azure · Terralogic.

12–24 Months

#3 IN INDIA · TECHGIG

ARTIFICIAL INTELLIGENCE

M.Tech. in
ARTIFICIAL INTELLIGENCE
AWS Academy · Microsoft Azure

24 Months

#5 IN INDIA · ANALYTICS INDIA MAGAZINE

CYBERSECURITY

M.Tech. in
CYBERSECURITY
EC-Council · Microsoft Azure · Terralogic

24 Months

CLOUD

PG Diploma / M.Sc. in
**CLOUD ARCHITECTURE &
SECURITY**
AWS Academy · Microsoft Azure

12–24 Months

Short-Term Certification Programmes

100 Hours · Project-Driven · Customisable for Organisations

AGENTIC AI | 4–6 weeks

AGENTIC AI SYSTEMS DESIGN

Multi-agent systems using LangChain, LangGraph, CrewAI. Capstone:
deployed autonomous agent.

AI ENGINEER | 8–12 WEEKS

AI ENGINEERING CERTIFICATION

Data pipelines, model training, LLMs, deployment. Real production
projects every batch.

CLOUD | 8–12 WEEKS

CLOUD ARCHITECTURE & DEVSECOPS

Multi-cloud, IaC, secure CI/CD on real AWS/Azure/GCP environments. Not
demos.

CYBERSECURITY | 8–12 WEEKS

CYBERSECURITY & SOC OPERATIONS

Threat detection, incident response, SIEM/SOAR. Hands-on in RACE's live
Security Operations Centre.

DATA & AI | 12 WEEKS

DATA SCIENCE & BUSINESS ANALYTICS

Statistics to ML to BI. Python, SQL, Power BI, Tableau on real business
datasets.

LEADERSHIP | 8–12 WEEKS

TECHNO-FUNCTIONAL LEADERSHIP

Executive presence, decision-making, leading tech-driven teams. Built for
mid-to-senior managers.

Enroll with us · race@reva.edu · race.reva.edu.in



REVA Academy for
Corporate Excellence

RETINA

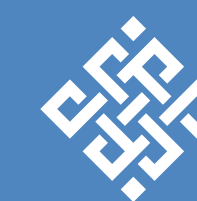
AI · CYBERSECURITY · CLOUD · LEADERSHIP

THIS WEEK, ASK YOUR LEADERSHIP TEAM

"Which AI models touch our data today, where does each one run, and if one were restricted tomorrow, how fast could we switch?"

REVA Academy for Corporate Excellence,
REVA University
Rukmini Educational Charitable Trust
Kattigenahalli, Yelahanka, Bangalore, Karnataka, India, 560 064
Phone: +91 89040 58866

RETINA PUBLISHES EVERY SUNDAY · FORWARD TO A CXO YOU RESPECT



REVA
UNIVERSITY
Bengaluru, India



Visit our website
race.reva.edu.in