FORUM REPORT

AI Vision Forum Paris 2026

Architecting Human–AI Synergy

A summary of proceedings, themes, and recommendations from the AI Vision Forum convened in Paris on 4 May 2026, co-located with GOSIM Paris 2026, and conducted under the Chatham House Rule.

FORUM DATE
4 May 2026 · Paris, France

PUBLISHED
May 2026

paris2026.visionforum.ai

Contents

Foreword

On Monday 4 May 2026, in the days immediately before GOSIM Paris, a small, international group of researchers, founders, foundation stewards, regulators, and educators met in Paris for the AI Vision Forum. The forum was an invitation-only gathering held entirely under the Chatham House Rule. Its tagline — *Architecting Human–AI Synergy* — was framed by the host as a deceptively simple question:

> "As AI agents become increasingly autonomous, embedded, and consequential, who is really responsible for what they do? How do we build agents that earn trust rather than demand it?"

This report distills the day's four panels — Agentic AI Systems, Agentic AI in Education, Trusted Agentic AI, and Open Token & Digital Public Goods — into a single shareable record. It does not name speakers or affiliations. It does try, faithfully, to capture what was said and where consensus, tension, and concrete next steps emerged.

Direct quotations are reproduced verbatim from the cleaned audio transcripts. A small number of structural framings — the **CLAW stack**, the **Seven Pillars of Open**, the **three-tier Proof of Control** taxonomy, and the term **deterministic control plane** — are editorial constructs the Organizing Committee uses to make the day's conversation easier to navigate; they are flagged where they appear.

— The AI Vision Forum Organizing Committee

Executive Summary

The AI Vision Forum Paris 2026 was framed throughout the day as a strategic, present-tense conversation, not a future-philosophical one. Across one keynote and four panels, the room kept returning to a small set of high-conviction claims:

> ***The stack has changed — and openness has to reach the bottom of it.*** *The agentic transition is not a software-only event. Open weights are nearly commodity; the remaining frontier is the open compute substrate beneath them — compiler, kernel language (e.g. Triton), operator library, and cross-vendor enablement (the FlagOS-class porting that brought a flagship open-weight model to more than ten AI chipsets in two or three days). Open source must extend across the four layers of what we call the* ***CLAW stack*** — Compute, LLMs, Agents, Workflow — *or sovereignty over AI will be impossible regardless of which flag flies over the data center.*

> ***Trust is engineered, not declared.*** "We are no longer managing just hallucination. We are managing trust." *Agents already outnumber humans, by the panel's measure,* "exponentially". *Verifiability has to move from logging to* tamper-resistant, machine-to-machine evidence at every boundary. *Sovereign-AI claims without cryptographic verification are incomplete:* "we never talk about sovereign AI again without talking about verification in the same sentence."

> ***Education has to lead, not follow.*** "Ninety-five percent of different experiments have failed, or only maybe from five to ten" *in the recent enterprise AI wave — not because the models are weak, but because the* cognitive design *around them is missing. The classroom is the place where AI either widens the gap between top and median learners or finally closes it.* "Whatever you use a tool for, you are going to get worse in that skill." — *productive struggle has to be designed in.*

> ***Tokens are infrastructure.*** "In the AI coding bureau, code is cheap. Show me the data." *The unit of strategic resource is shifting from code to* ***tokens****. Open-source maintainers are under pressure from AI-generated contributions; corporate sponsorship alone is not the answer.* "It doesn't matter how many tokens you give me. My community will still think that I'm the grinch that stole Christmas because it will never be enough to redistribute fairly to absolutely everyone."

> ***Human–AI synergy is a contract.*** *"Synergy" was repeatedly reframed throughout the day as something to be engineered: explicit roles, verifiable identities, auditable behaviour, and friction by design where learning, judgement, and meaning are at stake.*

> ***The day's most quoted line — the chant the room repeated together:*** **"I will not be eaten by AI."** *The opening keynote turned Arthur C. Clarke's* "any sufficiently

> advanced technology is indistinguishable from magic" *into a refusal: magic deceives; white-box, inspectable AI is the alternative.*

Headline Findings

1. **The Seven Pillars of Open must be defended together.** Open science, open data, open standards, open source, open weights, open platform, and open hardware — a layer-by-layer enumeration the Organizing Committee uses to summarise the day's openness arguments. *Open weights alone do not produce open AI.*
2. **Agents break existing licenses.** Open-source licenses regulate code use, modification, and distribution; *"they couldn't restrict the uncertainty of agent. So it's necessary to develop new license."*
3. **A backbone-plus-specialists agent architecture is converging.** A strong open-weight agentic backbone orchestrates many small specialised language models for narrow tasks — the panel rejected the framing that one open generalist must do everything.
4. **The junior pipeline is being reshaped, not eliminated.** Multi-tenancy, reliability, and security judgement are being compressed into the first year of an engineering career because juniors will now spend their time evaluating, not writing, code.
5. **The verifiability gap is widening.** Compute cost is falling; the cost of verifying what an AI did, with what data, on which chip, under which policy, is rising. The **deterministic control plane** — tamper-resistant, contemporaneous, binary-auditable evidence at every boundary — is the technical artefact most urgently missing.
6. **The EU AI Act collides with agentic reality.** Obligations like accuracy and human oversight actively trade off in tuning. High-risk obligations may be postponed to **December 2027**; the panel's preference is harmonised process-and-procedure auditing rather than reopening the Act.
7. **Tokens-as-public-good is a serious proposal — but it is not free tokens.** *"In this world, no meal is free."* The day called for an *Open Token* construct evaluated against the Digital Public Goods Alliance's nine indicators, not for vendor giveaways used as marketing funnels.

A Single Sentence

Architecting human–AI synergy is the work of replacing implicit trust with engineered trust — at every layer of the stack, in every classroom and code review, and across every border.

About the AI Vision Forum Paris 2026

Date: Monday, 4 May 2026 **Location**: Paris, France **Co-located event**: GOSIM Paris 2026 **Format**: Invitation-only · ~100 participants · Chatham House Rule **Tagline**: *Architecting Human–AI Synergy*

The day was structured as one keynote plus four moderated panels:

Time	Panel	Subtitle
Morning Session 1	Agentic AI Systems	Human–AI Symbiosis
Morning Session 2	Agentic AI in Education	Learning & Creativity
Afternoon Session 1	Trusted Agentic AI	Governance, Safety & Sovereignty
Afternoon Session 2	Open Token & Digital Public Goods	Foundations & Sustainability

Participants represented research institutions and foundations across Europe, North America, China, and the Global South; model labs and AI infrastructure companies; standards bodies (ISO, ITU, ITRI); and several nation-state-level policy programmes. The geographic and institutional mix was deliberate: the forum's premise is that no single jurisdiction or sector will resolve the agentic transition alone.

A Note on Methodology

This report was assembled from the full audio record of the day — approximately six and a half hours of speech across one keynote and four panels — transcribed and edited by the Organizing Committee. **Direct quotations** are reproduced verbatim, with names and affiliations removed in accordance with the Chatham House Rule. Where the underlying audio contained gaps or unintelligible passages, those were treated as silences and not filled by speculation.

Where the report uses a **structural framing** that is the Organizing Committee's editorial overlay rather than verbatim panel speech — *CLAW stack*, *Seven Pillars of Open*, the *three-tier Proof of Control* matrix, *deterministic control plane* — those terms are flagged in the text. Section ordering, the framing language of the Executive Summary, and the synthesis of cross-cutting themes are also editorial choices of the Organizing Committee, not the panelists.

Opening Keynote — Sketching the Day

The day opened with a personal trace: a PDF written in 1992 that still opens cleanly in 2026 — used as a small celebration of *open standards*. The speaker's first neural-network project, in 1992, had taken roughly six months to "get something that didn't work." In March 2023 — when ChatGPT first reached the market — he asked it to rewrite that 1992 project. *"Of course, it wrote it in ten seconds. But I ran the code and the code failed. Because it was pure hallucination."* After debugging, he asked ChatGPT to fix its own error. The fix arrived in another ten seconds, and the program ran at ~95% accuracy.

That thirty-year arc set up the day's single anchor argument: the theory of neural networks has changed almost nothing in those decades; what changed is that *it works now*, and the working version compounds exponentially. The keynote called the combined trajectory of Moore's Law, big data, and AI capability a *double exponential* — one that no individual, "perhaps a few exceptions," can comprehend in real time:

> "I've been eating algorithms for breakfast every day for the past four years. And I'm telling you, I'm starting to feel the crunch. Something is tearing in the fabric."

The keynote also offered a verbatim observation that runs through every later panel: *"AI is Cloud."* — meaning that the infrastructure question and the model question are now indistinguishable. And it offered a polite jab at the closed incumbent: *"OpenAI was closed. OpenAI is a fake, open name. It's closed."*

The keynote closed with a call-and-response that gave the day its most quoted line. Citing Arthur C. Clarke — *"any sufficiently advanced technology is indistinguishable from magic"* — the speaker pushed back that magic deceives and that magic is therefore an unacceptable design choice. Then the chant the room repeated together:

> **"I will not be eaten by AI."**

From there the panels move from manifesto to architecture.

Panel 1 — Agentic AI Systems: Human–AI Symbiosis

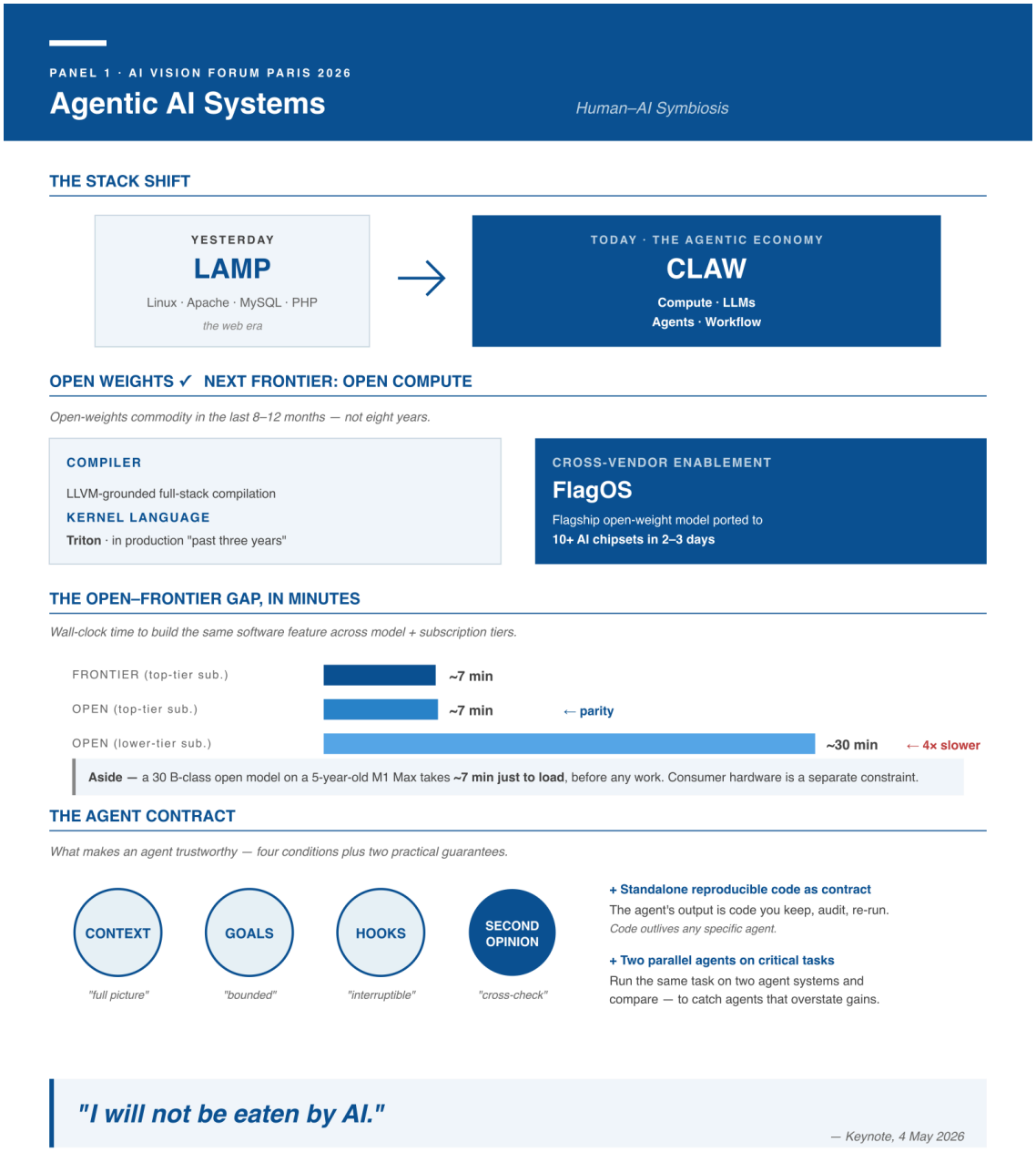

Figure 1: Panel 1 summary infographic — visual recap of the panel.

The Framing

The moderator opened with the panel's hardest question: *"What does real symbiosis between a human and an agentic system actually look like?"* The panel was constructed across continents and disciplines — a Cambridge mathematician working on automatic theorem proving, a German non-profit behind

some of the most widely used open data sets, a Chinese standards engineer, a European AI-infrastructure company, and a "personal AI" practitioner. The moderator explicitly aimed *"to create some friendly disagreement"*, and the panel did not converge on everything.

What the Panel Argued

- **Interfaces have moved from artificial to natural.** Seventy years of computing interfaces — punch cards, keyboards, terminals, MS-DOS, GUI, miniaturisation, touch — each step asked humans to adapt to the machine. AI is the first interface that adapts to *us*. The same AI is also competent at legacy code (COBOL was cited explicitly).
- **Open source is the challenger discipline.** Windows vs Linux, AWS/Azure vs Google/Kubernetes, OpenAI vs Meta and DeepSeek — incumbents are repeatedly forced open by challengers. The panel argued for keeping that pattern intact for the agentic era.
- **Open weights are nearly commodity; the next frontier is open compute.** Llama, DeepSeek, Qwen, Minimax and Kimi families have reached general usability over the last 8–12 months — *"not eight years"*. The harder remaining work is the *open compute substrate*: a Python-like kernel language (**Triton** was cited explicitly, in production use *"for the past three years"*), an operator library, and cross-vendor enablement that lets the same model run on heterogeneous AI chipsets.
- **FlagOS-class enablement is already shipping.** One panelist's team reported porting a flagship full open-weight DeepSeek model to **more than ten AI chipsets in two or three days**, with one accelerator reaching roughly **70% of the throughput of a comparable NVIDIA chip** (uncertified). *Open weights without an open compute substrate is a hollow victory.*
- **Backbone + specialists is the working agent architecture.** A strong open-weight backbone orchestrates many small specialised language models for narrow tasks. As one panelist put it:

> "We don't need to be providing a recipe for carbonara and COBOL code in the same way."

- **The open–frontier gap is being measured in minutes, not years.** A concrete benchmark from the room: a feature a frontier closed model produced in ~7 minutes took a Kimi-class open model ~30 minutes on a lower subscription tier, and roughly 7 minutes on a top tier. An open 30 B-class model on a five-year-old M1 Max consumer Mac took ~7 minutes *"just to load the skill."* The historical pattern across Llama, DeepSeek, and Qwen suggests these gaps close steadily — the editorial *Seven Pillars of Open* frame names the bet: that openness wins layer by layer if it is defended layer by layer.
- **Data flight and provider lock-in.** *"When we use Claude or any of these frontier models, we are unlearning things."* The concern raised in the panel was not a romantic argument for using inferior tools to keep skills sharp. It was twofold and concrete: heavily subsidised frontier APIs route private prompts and reasoning context into closed labs that then use that context as training data; and exclusive reliance on a single closed interface narrows an organisation's working knowledge of the underlying stack and erodes optionality. The working discipline some practitioners advocated, therefore, was *not* "always use the slower stack," but **maintain both**: keep at least some critical workflows running on open-weight stacks so the team can keep its proprietary context out of closed pipes and stay able to switch providers when the cost-and-capability landscape shifts.
- **Trust comes from inspection.** Trust in an agent comes from context, goals, hooks, and a *second-opinion* pattern — a different agent (different model family where possible) reviews the first agent's output before action. A stronger formulation: run *two parallel agent systems* on the same task to detect agents that *"play some trick"* and overstate optimization gains.

- **Agents should produce artefacts, not just answers.** The panel argued that the most useful unit of agent output is *runnable software* — a script, a program, a config — rather than a narrative summary of "what I did." A runnable artefact is something the user owns outright and can re-execute, audit, or hand to another team, regardless of whether the original agent is still available or behaves differently next time. Treating the artefact as the deliverable, not the chat transcript, is the discipline.
- **The governance problem is unresolved — and not, strictly, a licensing problem.** *"Existing license, I think they couldn't restrict the uncertainty of agent. So it's necessary to develop new license."* The point the panel was making, more precisely than the phrasing here suggests, is that *runtime agent behaviour* — what an agent decides to do, which sub-agents it spawns, what data it touches — sits outside what software licenses can meaningfully control. Licenses grant permission to use, modify, and distribute; they cannot dictate behaviour once the software is running. That work belongs to **regulation** (the EU AI Act, sectoral safety regimes) and to **terms-of-use conditions** in the spirit of Responsible AI Licenses (RAIL), which condition permission on disclosure, audit-trail provision, and process governance. The panel's call for "new licences for agents" is best read as a call for both: new conditional license terms *and* new regulatory categories that cover what licenses structurally cannot.
- **The engineer role is shifting from excavator to orchestrator.** Seniors will learn context engineering, MCP-style tool plumbing, drift detection, and how to select and evaluate models. Juniors face a harder transition: their entry-level work is *evaluating* agent output, so multi-tenancy, reliability, security, and cybersecurity judgement must be compressed into the first year.
- **Africa is a present, not a future, story.** One panelist's organisation has a Memorandum of Understanding with the African Union's **ASRIC** programme covering more than **50 countries**; the first call drew over **100 teacher applications from 28 countries**, and an initial cohort of **11 teachers from 11 universities** was hosted in Beijing.
- **Benchmarks have to stay ahead of capability.** When a benchmark saturates — when frontier models reach near-perfect scores on it — it stops measuring anything useful and becomes a marketing prop. The panel offered **Terminal Bench** as an example of a benchmark deliberately designed to resist saturation: co-built with Stanford and Berkeley, its 2.0 "hard" track has frontier models stuck at *"twenty thirty percent"* success; Terminal Bench 3 is already in development to keep the difficulty curve ahead of capability gains.

Notable Tensions

- **Aspirational vs. operational openness.** Aspirational "fully sovereign, local, open" stacks vs. the current reality that even on good open weights, latency and quality lag the frontier by minutes per task.
- **Open-weight optimism vs. realism.** *"Ten years, two years from now"* parity vs. *"we are very far away"* because private context flows continuously into closed labs as training data.
- **Personal-agent enthusiasm.** Some panelists actively delegated overnight workflows to local agent swarms; others wanted agents confined to a narrow optimisation perimeter under mandatory cross-check.

Recommendations

Figure 2: Panel 1 recommendations infographic — concrete actions.

- Pool underutilised data-centre compute globally and route it via open programmes for the Global South.
- Open the full CLAW-stack compute layer — compiler, kernel language (e.g. Triton), operator library, FlagOS-style cross-vendor enablement.
- Adopt the *backbone + specialists* agentic architecture rather than chasing one open generalist.
- Treat the agent's deliverable as **runnable software** the user owns, audits, and can re-execute — not a chat transcript of what the agent claims to have done.
- Run two parallel agent systems on critical tasks; have agents seek a second opinion from a different model family.
- Develop the dual-layer governance the panel called for: (a) **regulation** that covers runtime agent behaviour, sub-agent spawning, and incident reporting, alongside (b) **RAIL-style conditional**

licences that require operators to publish audit trails and process governance as a condition of permission to use.
- Maintain capability across both open and closed stacks — to keep proprietary context out of closed pipes and preserve provider optionality.

Panel 2 — Agentic AI in Education: Learning & Creativity

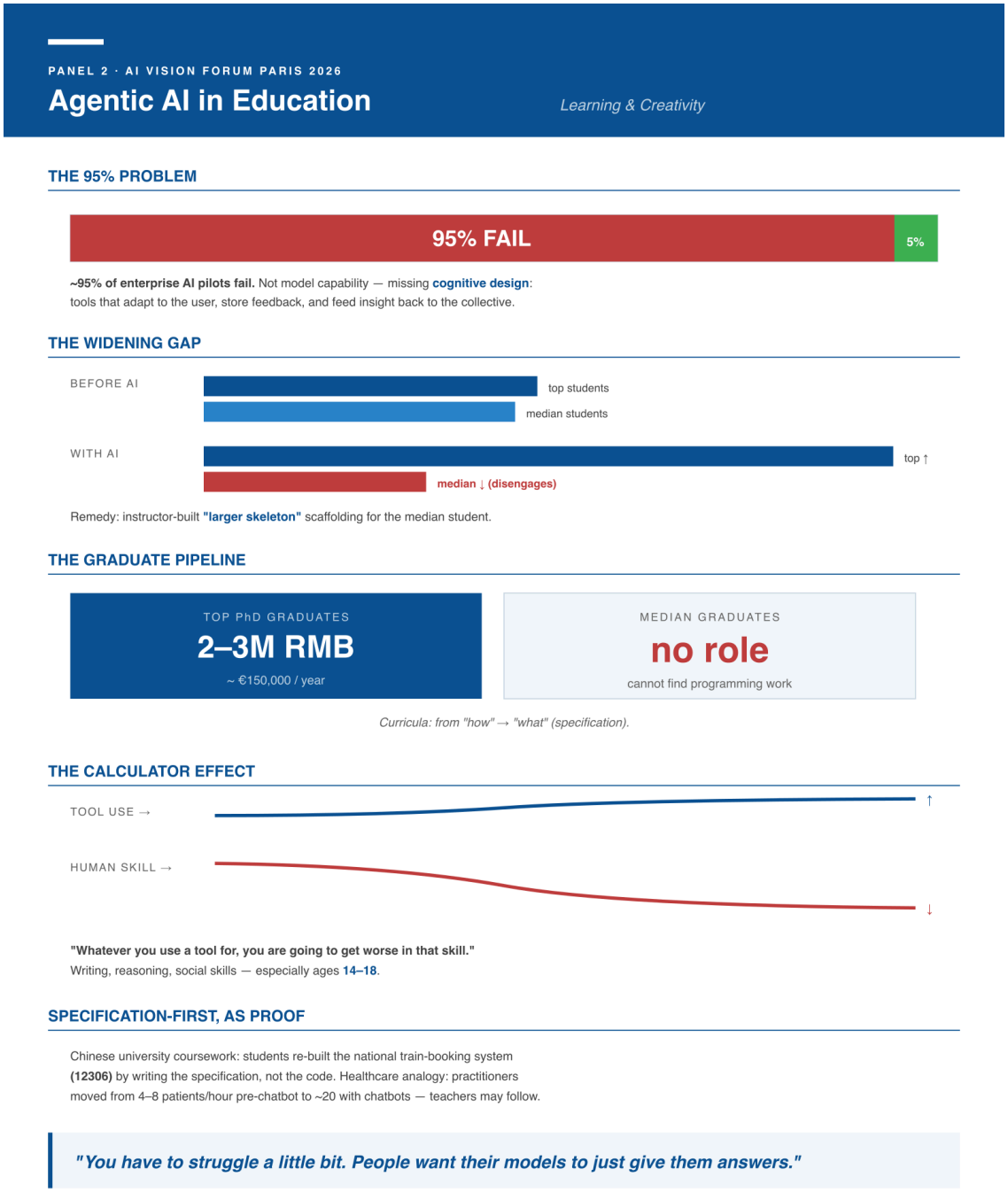

Figure 3: Panel 2 summary infographic — visual recap of the panel.

The Framing

> “Education is where the AI debate stops being abstract, starts being about our children and our future.”

The Education panel was deliberately steered toward possibilities rather than doomsday framings, while taking the risks seriously. Multiple panelists chose to anchor the discussion in **cognitive science** rather than capability benchmarks: expert work — scientific, journalistic, public-sector — is *cognitively, socially, and physically distributed*, never individual. Any AI deployment that doesn't fit that distributed structure will fail in the same way most enterprise pilots have failed.

The 95% Problem

The single most-referenced statistic of the morning came from outside the room: *"ninety-five percent of different experiments have failed, or only maybe from five to ten"* — recent enterprise-AI-pilot results cited from MIT and similar studies. The panel's diagnosis was unanimous: the failures are not model-capability failures. They are *cognitive design* failures — tools that don't adapt to the user, don't store and reflect feedback, introduce friction without offsetting support, and ultimately produce an organisational **cognitive black hole** in which the system gets smarter and the organisation does not.

The corollary for education: if a school just hands out a chatbot, the same 95% failure mode will appear — but with higher human cost.

What the Panel Argued

- **The teacher's role is shifting, not disappearing.** From knowledge-transmitter to motivator, contextualiser, and provider of cross-validation. Teachers are also the irreplaceable carriers of *silent know-how*: how a hypothesis is judged, how a proof "feels" right before it is checked. Models cannot yet teach this directly.
- **The gap between top and median students is widening.** Strong students push agent-driven projects further than ever; the median student disengages. The remedy named by the panel is an instructor-built *"larger skeleton"* — explicit scaffolding for non-elite students so the median doesn't fall out of the distribution.
- **Software-engineering education is becoming specification, not syntax.**

> "Software engineers not going to write the how problem… but to write the what problem. They need to define their specification."

The unit of work is becoming "write an agent that writes the code"; coursework is moving toward optimisation against a test harness. One Chinese university had students rebuild the national train-booking system (12306) by writing the specification, not the code.

- **The graduate market is bifurcating.** In some markets, top PhD graduates command roughly **2–3 million RMB per year (~€150,000)**; the median graduate cannot find a programming role at all. Curricula must move up the abstraction stack — defining problems, building specifications, framing better questions — or be left training students for jobs that no longer exist.
- **Productive struggle is essential.**

> "You're not going to learn if you're just giving the answer. You have to struggle a little bit. And people will like that in their chatbots. I don't know why anybody wants their models to just not give them answers."

Chatbots default to removing friction; learning needs friction calibrated to the learner.

- **The calculator effect is observable now.** *"Whatever you use a tool for, you are going to get worse in that skill."* Writing, reasoning, and even social skills sit under the same risk if AI does them too early in a child's development.
- **Personalisation is real but bounded.** Agents can adapt explanation depth — PhD student, professor, novice — but truly modelling how an individual child thinks is still beyond them.
- **Adolescent mental health and AI peer-substitution is acute and present.** Identity-formation years (ages 14–18) need rejection, conflict, and negotiation practice with humans; chatbot substitution is observable now.
- **Healthcare analogy.** Practitioners moved from 4–8 patients per hour pre-chatbot to ~20 with chatbots; the panel asked whether teachers will move toward a similar background-orchestrator role with agents doing more frontline interaction.
- **Equity and infrastructure are preconditions.** AI in education in regions without reliable electricity or internet is gated by basic infrastructure — off-grid solar, local offline networks, edge compute. Token efficiency was named as a sustainability lever: the claim that *"over ninety percent of token consumption"* could be saved in many LLM applications.

Notable Tensions

- **Frictionless UX vs. productive struggle.** Users demand answers; learning requires withholding them.
- **AI as tutor-replacement vs. teacher-augmentation.** Flexibility of personal AI tutors against the claim that programs "behind the screen" can never fully replace the contextual, motivational role of a teacher.
- **When to introduce AI.** Consensus that university is too late; unease about unsupervised chatbot use in early childhood and adolescence.
- **Talented vs. median students.** Geniuses can self-explore productively; the majority need explicit scaffolding.

Recommendations

Figure 4: Panel 2 recommendations infographic — concrete actions.

- Lead with **cognitive design** before procurement. Any AI tool in a classroom should adapt to the user, store feedback, and feed insight back to the collective.
- Train teachers in *how* to deploy AI for students before deploying it to students.
- Introduce AI in **elementary school**, not university.
- Build agents that *introduce* calibrated friction rather than remove it; protect productive struggle.
- Scaffold the median student explicitly; do not assume AI productivity gains distribute evenly.
- Pursue token-efficient LLM applications as a sustainability/equity strategy.
- Limit unsupervised chatbot use for adolescents; design systems that protect, not displace, peer interaction.
- Pair every AI tool with a teacher-mediated cross-validation source so students see disagreement, not a single confident voice.

Panel 3 — Trusted Agentic AI: Governance, Safety & Sovereignty

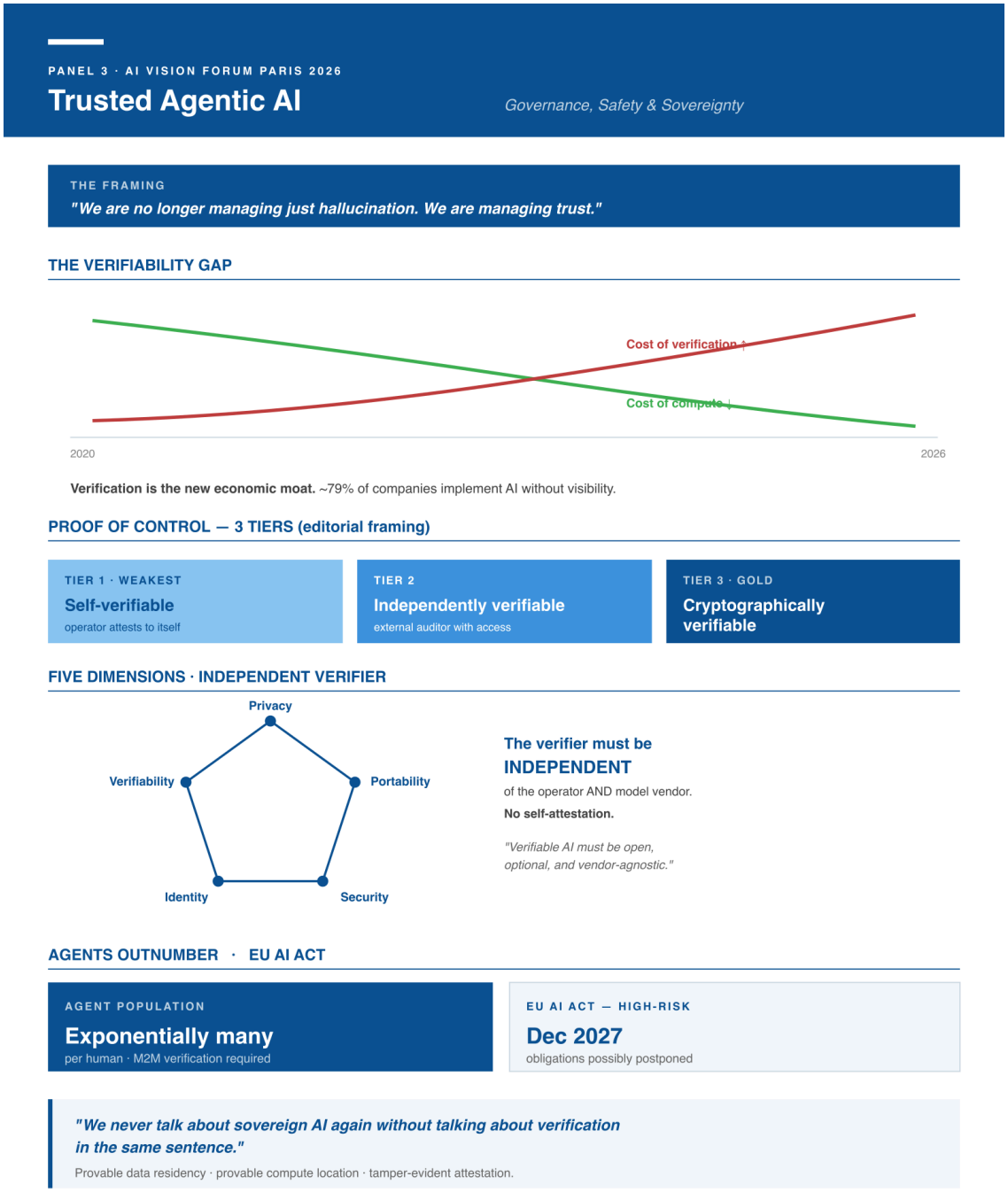

Figure 5: Panel 3 summary infographic — visual recap of the panel.

The Framing

> "We are no longer managing just hallucination. We are managing trust."

The panel opened with a deliberate reframing: trust, not hallucination, is now the central management problem. Agentic AI no longer just answers — it plans, books flights, accesses wallets, updates calendars, and spawns sub-agents. *"Agents outnumber humans not just on an order of like, oh, ten to one. We're talking about exponentially they outnumber us."* Centralised, static governance — the model under which most product-safety regulation evolved — does not apply.

The panel deliberately spanned three regions and brought together open-source maintainers, a standards-body member, a Chinese AI-and-society academic, a European AI-Factory operator, and audience contributors with experience inside the EU AI Act negotiation.

What the Panel Argued

- **"Human in the loop" is necessary but insufficient.** Agents already outnumber humans by orders of magnitude; verification has to move to *machine-to-machine* scale, with explainability-by-design and minimum viable standards.
- **Agent identity is a bundle of verifiable claims.** Not "which agent" but *which agent, in which role, under which policy, in which jurisdiction*. Authentication, authorisation, and delegation each need explicit treatment, especially across open ecosystems where a bank agent calls an insurance agent calls a credit-scoring agent.
- **OAuth was built for a human-initiated web.** Verifiable Relationship Credentials (VRCs), decentralised identifiers, and similar primitives that have lived in the decentralised-tech world have to merge into mainstream agent stacks.
- **Open source itself is at risk.** Maintainers can no longer tell whether a pull request comes from a human or an agent. Verifiable relationship credentials are being introduced into maintainer workflows for that reason. Several Linux-kernel-adjacent incidents were referenced as background.
- **Verifiable AI is a distinct category.** *"AI in general is based on faith."* Verifiable AI must be a separable, certifiable category — *"open, optional, and vendor-agnostic"*.
- **The EU AI Act collides with agentic reality.** Obligations — human oversight, robustness, accuracy, fairness — interact and trade off. Higher accuracy lowers the human-oversight burden; stronger fairness can collide with accuracy. There are no single thresholds that satisfy them simultaneously for an agentic system. High-risk obligations may be postponed to **December 2027**; the panel's preferred path is harmonised *process-and-procedure auditing*, not reopening the Act.
- **Logs are not enough.** Logs are tampering-prone after the fact. Meaningful oversight requires *tamper-resistant, contemporaneous, binary-auditable* evidence at every boundary — every data crossing, authentication/delegation event, and payment settlement. The Organizing Committee uses **deterministic control plane** as a shorthand for this artefact (and notes that a working "node-takeover" demonstration exists: an agent kicks the node owner off, processes, wipes, and returns control, providing cryptographic proof of no-owner-control).
- **Sovereign AI is paired with verification or it is incomplete.**

> "We never talk about sovereign AI again without talking about verification in the same sentence."

True sovereignty requires provable data residency, provable compute location (down to the specific chip), and tamper-evident attestation. A domestic vendor logo is not a substitute.

- **A practical Proof-of-Control taxonomy (editorial framing).** The Organizing Committee captures the panel's verification thinking in a three-tier × five-dimension matrix: tiers — *self-verifiable*, *independently verifiable*, *cryptographically verifiable* — across dimensions — privacy,

portability, verifiability, security, identity. The structural property the panel insisted on: the verifier must be independent of the operator and the model vendor. **No self-attestation.**
- **Standards must be global; regulation may stay regional.** Ecosystems will diverge on policy. They cannot diverge on technical standards for identity, traceability, and verification — or "sovereignty" collapses into vendor lock-in and cross-border agent traffic becomes unsafe by construction. ISO, ITRI, ITU, IEEE, and the Linux Foundation were named as bodies that need to converge.
- **Insurance may move faster than regulation.** Liability allocation is the ultimate enforcement; insurance underwriters cannot price what they cannot verify. Treating the insurance industry as a first-class stakeholder may produce binding verification standards faster than national legislation.
- **Verification is the new economic moat.** As compute cost falls, the cost and importance of verification rises. Value accrues to whoever owns the trusted control plane, not to whoever serves the cheapest token. A figure of **~79%** was cited regarding companies that have implemented AI without knowing what is happening underneath. US AI investment was named at *"ten times"* that of Europe — the regulation-versus-innovation debate, on this reading, is partly a capital-availability debate in disguise.

Notable Tensions

- **Reopen the AI Act vs. work within it.** Some panelists argued the Act's product-safety frame structurally cannot accommodate agents; others warned that reopening invites years of political delay.
- **Worst-case classification vs. sandboxed experimentation.** An audience contributor proposed forcing every agentic system into the high-risk bucket. The panel pushed back: not enough cases yet to define "worst case", and blanket high-risk would drive innovators out of the jurisdiction without making systems safer.
- **National-data sovereignty vs. distributional AGI.** One framing emphasised data residency and domestic stack ownership; another emphasised AGI-class capability arising from *international collaboration of specialised sub-agents* and warned that excessive national isolation will leave countries unable to participate.

Recommendations

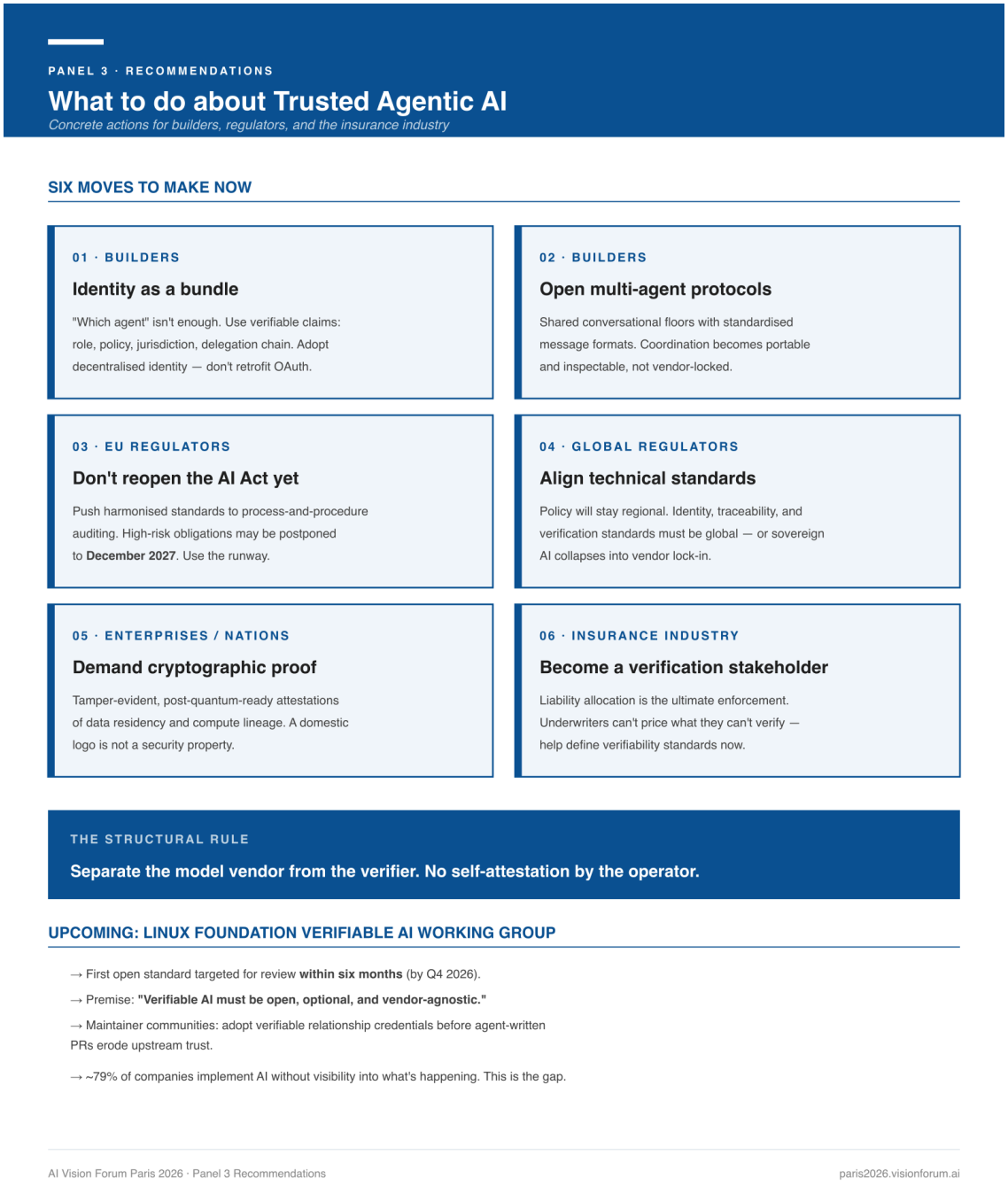

Figure 6: Panel 3 recommendations infographic — concrete actions.

- **Builders.** Treat agent identity as a bundle of verifiable claims; adopt cryptographic, decentralised identity primitives rather than retrofitting OAuth.
- **Builders.** Design multi-agent systems on open, transparent message-passing protocols with shared conversational floors so coordination is portable and inspectable.
- **Regulators (EU).** Do not reopen the AI Act for agentic AI yet. Push harmonised standards toward process-and-procedure auditing with provider-justified thresholds. Use sandboxes — including reinforcement-learning-based simulated regulators — to learn empirically how obligations interact.
- **Regulators globally.** Align on technical standards even where policy diverges.
- **Enterprises and nation-states pursuing sovereign AI.** Demand cryptographic, tamper-evident, post-quantum-ready attestations of data residency and compute lineage.
- **Insurance industry.** Become a first-class stakeholder in defining verifiability standards.

- **Open-source maintainer communities.** Adopt verifiable human-vs-agent authorisation standards (verifiable relationship credentials) before agent-written PRs erode upstream trust.
- **Standards bodies.** Separate the model vendor from the verifier. Require independent, externally inspectable proof.

The panel surfaced a forthcoming initiative — a **Verifiable AI working group** to be hosted at the Linux Foundation, with the first open standard targeted for review *"within six months"*.

Panel 4 — Open Token & Digital Public Goods: Foundations & Sustainability

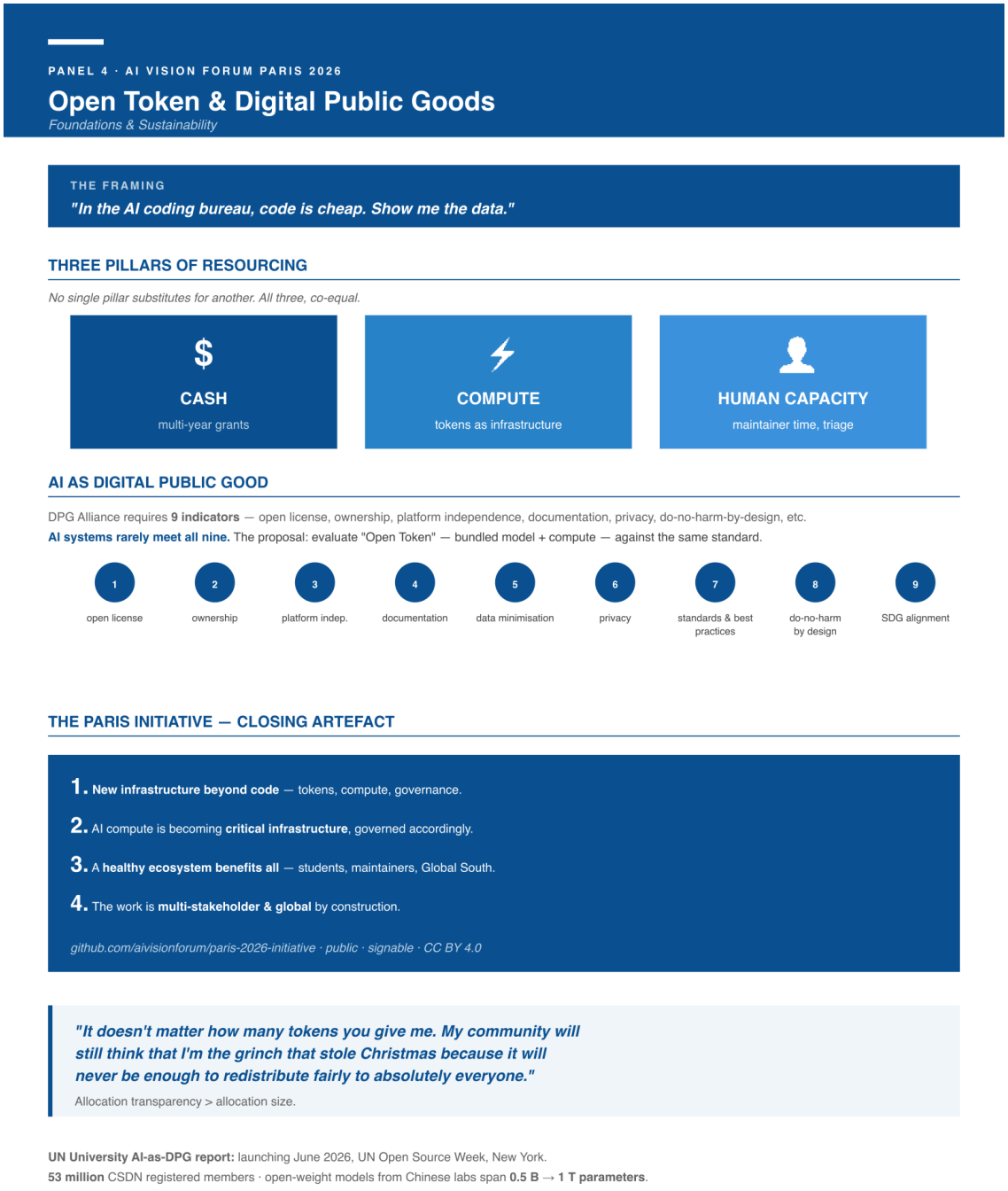

Figure 7: Panel 4 summary infographic — visual recap of the panel.

The Framing

The final panel returned to infrastructure with a single reframing:

> "In the AI coding bureau, code is cheap. Show me the data."

The strategic resource is shifting — from code to **tokens**, from software to **apps**, from models to **infrastructures**.

The panel was structured in three explicit rounds — *Why now*, *Governance and sustainability*, *Global collaboration* — anchored to the question of whether tokens can be treated as a new category of digital public good. Regional strengths were named as complementary rather than competitive: Europe contributes governance maturity; China contributes scale, engineering volume, and rapid product cycles; the Global South contributes both need and largely-untapped talent. The CSDN founder noted that CSDN, founded twenty-five years ago, now has roughly **53 million registered members**.

What the Panel Argued

- **Open-source built the modern digital world; the AI era introduces a new strategic resource.** Linux at the bottom, Python in the middle, Rust securing the next-generation systems — the LAMP era's success depended on a quiet consensus that the substrate was a shared resource. Tokens now sit alongside code as foundational infrastructure.
- **A "Linux for agents" has not yet emerged.** The current AI stack is a five-layer system in which open-source and closed-source giants coexist at every layer; the opportunity is to build community-driven infrastructure where it is still possible.
- **Maintainer pressure is the most acute bottleneck.**

> "Stuff on the internet isn't free… my blood, sweat, and tears have paid for [it] to be online."

Maintainers of critical open infrastructure are at the edge of quitting; AI accelerates demand on them rather than relieving it.

- **Open Token as a candidate Digital Public Good.** The Digital Public Goods Alliance's standard requires more than an open license: nine indicators including ownership, platform independence, documentation, privacy, and a do-no-harm-by-design property. AI systems rarely meet all nine. The proposal is to evaluate an *Open Token* construct — bundled model + compute access — against the same standard, so AI DPGs are actually runnable. A UN University report on AI-as-DPG is scheduled to launch *in June, during UN Open Source Week in New York*.
- **Compute governance is emerging as a policy lever.** Drawing on the 2024–2025 literature on compute governance: monitor who trains what at what scale, steer subsidies toward public-interest training, embed cryptographic verification at the hardware level.
- **Equity through Open Token.** Affordable tokens for students, researchers, hackathon participants, and open-source maintainers let people experiment across a rapidly-shifting framework landscape without being priced out of learning. Token allowances of *"$400 per employee, no limit"* were cited as an enterprise pattern; *"200 million tokens consumed by a founder"* was floated as a startup-readiness signal.
- **Local and edge models complement Open Token.** Small models running on-device — Gemma was named — offer self-control, privacy, and effectively "free" usage where the user already paid for hardware and electricity. Cloud tokens remain higher-quality for frontier tasks. The right architecture is hybrid, not exclusive.
- **"Free" is never free.** *"In this world, no meal is free."* Free tokens are typically a marketing funnel that monetises data or lock-in. The goal is *affordable, governed* token access — not zero-price tokens that incentivise over-consumption and external environmental cost.

- **Cross-continent contribution flows are uneven.** Chinese labs are contributing open-weight models across size tiers (0.5 B, 1 B, up to 1 T parameters). European communities concentrate on governance, licensing, and digital-rights compliance. The Global South contributes labour and ideas largely invisibly.
- **Quality and security strain on open source.** AI-generated contributions can be high-volume and well-intentioned, but they can also erase prior human work, bypass community norms, and demoralise long-term contributors. The contribution-acceptance ceiling has to rise — reproducible test cases, manual QA, peer review.
- **Long-horizon analogy.**

> "When steel was created, [they thought] maybe it will only be used for the beams… But after fifty years, the most famous use case of steel is creating a lot of famous skyscrapers."

Today's most visible uses of tokens probably underestimate the durable uses they will support.

Notable Tensions

- **Charity vs. strategy in corporate contribution.** One panelist rejected the framing of "charity" entirely — all contribution is enlightened self-interest in a long arc. A foundation perspective pushed back: donor strings and PR demands make this feel unfair to maintainers on the receiving end.
- **Fairness of allocation.**

> "It doesn't matter how many tokens you give me. My community will still think that I'm the grinch that stole Christmas because it will never be enough to redistribute fairly to absolutely everyone."

Foundation-run programmes already provoke accusations of favouritism. Consensus formed around neutral, government- or foundation-mediated pools rather than single-company gatekeeping. - **Money vs. people as the binding constraint.** Funding remains "number one" in mind, but maintainer burnout was named as a more acute *present* bottleneck. - **Local vs. cloud.** Tension between the prediction that *"everything will be local in five years"* and the present reality that frontier quality, speed, and multi-modal capability still favour large cloud providers.

The Paris Initiative (closing artefact)

The closing remarks of the day announced a working four-principle **Paris Initiative — A Consensus Statement on Agentic AI Infrastructure**. The statement lives at `paris2026.visionforum.ai/initiative/` and is also published on GitHub at `github.com/aivisionforum/paris-2026-initiative` under a CC BY 4.0 licence, open for public signature — organizations and individuals add their names via pull request. A follow-up milestone is set for the next GOSIM gathering in Shenzhen, October 2026.

The statement opens with a preamble: *the following principles reflect consensus reached at AI Vision Forum Paris 2026 on the development of* ***open, trusted, and sustainable agentic AI infrastructure***. Each principle below is reproduced as written in the published statement.

I. The Agentic Era Requires New Infrastructure

We recognize that the world is entering a new computing paradigm — the Agentic Era. The LAMP and Cloud eras assumed humans as the primary actors. Autonomous agents — operating continuously, acting on behalf of others, consuming resources at scale — require new models for identity, authorization, and economics.

II. AI Compute Should Be a Public Good

We believe that AI compute is becoming critical infrastructure — as essential as electricity — and should be broadly accessible to foster innovation, education, and open source sustainability. The Open Token model offers a concrete, operational path toward AI compute as a digital public good: a bridge connecting token donors (LLM providers) with organizations that need compute — open source projects, researchers, educators, and civil society.

III. A Healthy Agentic Ecosystem Benefits All

We observe that a thriving agentic ecosystem creates shared value: LLM providers benefit from increased token consumption and developer adoption; open source projects gain sustainable funding; society benefits from accessible AI infrastructure. This is not a zero-sum dynamic — broader access grows the ecosystem for everyone.

Leading AI organizations have already demonstrated that structured compute access programs generate real returns in goodwill, ecosystem development, and community trust. Open Token builds on this precedent with a neutral, open governance model that no single vendor controls.

IV. Multi-Stakeholder Collaboration Is Essential

We affirm that building trusted agentic AI infrastructure requires active collaboration across sectors: LLM providers, open source foundations, international organizations, academia, and civil society each bring capabilities the others lack. No single actor can do this alone. Progress comes from working together on concrete programs that demonstrate what collaboration can produce.

Next Steps

Based on this consensus, participants supported the formation of a **Preparatory Working Group** to pursue three concrete goals:

1. **Grow Open Token from a community initiative into a structured, sustainable program** with clear governance, connecting token donors (LLM providers) with open source projects, educational institutions, and researchers seeking compute support.
2. **Secure sustained, multi-cycle token donation commitments** from LLM providers — piloting two or three recurring partnerships with open source foundations as proof of model, building toward a self-sustaining network.
3. **Invite broader participation** from additional token donors and recipient organizations across regions and sectors.

How to sign

The Initiative is signable by any organization or individual that publicly endorses the four principles. Two signing paths are available in the repository: a pull request adding your line to `SIGNATORIES.md` (preferred, fully traceable), or a GitHub Issue using the *Sign the Paris Initiative* template. Signatures will be reviewed by the Organizing Committee and merged with light verification. Signing is an endorsement, not a binding commitment, and signatories may withdraw at any time.

Recommendations

Figure 8: Panel 4 recommendations infographic — concrete actions.

- **Foundations and companies.** Resource open AI infrastructure on three pillars — cash, in-kind compute, human capacity — and stop substituting one for another.
- **Public bodies.** Build neutrally-governed token pools for students, researchers, and Global South contributors. Multi-year grants, not one-month vouchers.
- **Standards work.** Develop a “green, ethical Open Token” standard with provenance, environmental-impact disclosure, and DPGA-style qualification.
- **Compute governance.** Adopt hardware-level cryptographic verification, monitored allocation, and subsidies routed toward public-interest training.
- **Industry.** Companies running coding-agent ecosystems should sponsor maintainers directly with token grants and triage capacity.

- **Open-source projects.** Raise contribution-acceptance standards (reproducible test cases, manual QA, peer review) to absorb the AI-generated PR avalanche without quality collapse.
- **Allocation transparency.** Reward open-source contribution with token access (PR-, issue-, hackathon-based eligibility) with transparent allocation that does not bias model improvement back to the donor.
- **Measure outcomes, not consumption.** Build leaderboards and feedback loops so token grants are measured by what they produce, not by how many tokens were spent.

Education Research — A Companion Paper Series

Alongside the forum itself, the Organizing Committee published a three-paper research series that accompanies Panel 2 and extends its argument into theory and architecture. The series asks why the great theories of learning have never reached every learner, and what changes when the agent — not the assistant — becomes the unit of educational software.

The full series — with the complete text of each paper and downloadable PDFs — is published at `paris2026.visionforum.ai/education/`.

Paper 1 — Commonalities of Feynman, Socrates, and Piaget

Theory · Comparative Analysis. A structural analysis of three pedagogical traditions across two and a half millennia, identifying nine shared commitments — *learner as active constructor; cognitive conflict as catalyst; metacognition; depth over coverage; individualization; dialogic interaction; simplification through analogy; the teacher as guide; intrinsic motivation*. The paper's argument: these are not three competing methods but three articulations of the same underlying picture of how human understanding actually deepens. They have remained admired in theory and absent in practice for the same structural reason — each demands a sustained one-to-one relationship with someone who knows the learner, has unbounded patience, and adapts in real time.

Paper 2 — How AI Transforms Educational Implementation

Practice · LLM Capabilities. Maps the seven structural barriers that have prevented these theories from scaling — teacher supply, time, assessment, cost, culture, and two others — onto specific capabilities of large language models. The paper's claim is precise: AI is the first technology in human history able to deliver high-quality, personalized pedagogical interaction *without proportional human cost*. The implication is not that AI replaces teachers; it is that the constraint that has held the great traditions out of mainstream education for two thousand years has finally moved.

Paper 3 — From Socrates' Daimon to Digital Daimon

Architecture · Persistent Agents. A technical architecture paper. The "Assistant" paradigm — *stateless, reactive, identity-less* — cannot host the kind of relationship the classical theories require. A **Digital Daimon**, built as a persistent agent with memory, autonomous intervention, deep learner-modeling, and self-evolution, can. The paper lays out six architectural properties of such an agent and maps each one to the corresponding pedagogical theory and to the relevant layer of the open agentic stack.

Why this series sits alongside the Panel 2 discussion

Panel 2's discussion was descriptive — what is happening in classrooms right now, what is breaking, what the early data shows. The three companion papers are prescriptive: they take the panel's diagnosis ("cognitive design is missing"), connect it to a two-millennia-old theoretical tradition, identify the specific AI capabilities that finally unblock that tradition, and propose the architecture to deliver it. Together, panel and papers make a complete argument: the failure to scale classical pedagogy was a

constraint problem, the constraint has moved, and the agent — not the assistant — is the unit of software that can carry the load.

The Paris Synthesis — Cross-Cutting Themes

Across four panels with four distinct briefs, a small number of themes recurred with unusual consistency.

Trust is engineered

The word "trust" appeared in every panel and meant something slightly different each time. By the close of the day it had acquired a working definition: *trust is the property that an external observer can verify without relying on the producer's self-attestation*. This is the connective tissue between the Trust panel's Proof-of-Control framing, the Systems panel's call for second-opinion patterns, the Education panel's plea for cross-validating sources, and the Open-Token panel's demand for provenance on every contribution.

Open must be defended at every layer

Open weights are nearly commodity. Open *compute* — compiler, kernel language, operator library, cross-vendor enablement — is the next defended layer. Without it, the open-weights story is conditional on one chip generation. Open *standards* — for identity, traceability, verification — are conditional on no single jurisdiction. Open *governance* — for tokens, for verification, for community contribution — is conditional on no single company. The Organizing Committee's *Seven Pillars of Open* enumeration (open science, open data, open standards, open source, open weights, open platform, open hardware) is a navigational aid for keeping all seven in view simultaneously.

The junior pipeline is the strategic question

In each panel, a worry surfaced that does not fit any single brief: the people who would otherwise have become the next generation of senior engineers, scientists, teachers. Juniors must now compress a decade of judgement into their first year because their job is *evaluating* agent output, not writing it. Top graduates earn ~€150K and the median earns nothing. Maintainer pipelines are drying up. Curricula, sponsorship, and pipeline design are not adjacent concerns to AI — they are part of its critical path.

Standards global; regulation regional

Repeated almost as an axiom across panels. Policy will reflect national values, electoral incentives, and historical regulatory traditions. *Standards* — the technical artefacts that make identity, traceability, verification, and inter-agent cooperation actually work across borders — cannot. ISO/IEC, ITU, IEEE, the UN's DPG Alliance, and the Linux Foundation's emerging **Verifiable AI** working group are the bodies named.

Verification is the new moat

As compute cost falls, the cost and importance of verification rises. Value will accrue to those who own the trusted control plane — the substrate that lets a regulator, an enterprise customer, or an insurance underwriter independently confirm what an agentic system actually did. This re-frames much of the

day's commercial and policy discussion: investment in verification infrastructure is not a cost centre. It is the next moat.

Friction is a feature

Across the Education, Trust, and Open-Token panels, a quiet consensus emerged that *frictionlessness* is a poor default. Learning requires productive struggle. Trust requires independent verification (which adds latency). Sustainable open infrastructure requires governed access (which excludes some uses). The agentic transition is in part a redesign of where, intentionally, to put the friction back in — so it lands at the points where humans, not machines, do their best work.

Looking Forward

The forum produced few unconditional predictions and many conditional commitments. The most concrete of these — the **Paris Initiative** on Open Token, the call for a multi-stakeholder working group with a public-GitHub commitment, the GOSIM Shenzhen follow-up milestone in October, the Linux Foundation **Verifiable AI** working group with a six-month standard-review target, and the UN University AI-as-DPG report launching in June — are the threads to watch over the coming year.

Three concrete artefacts the room agreed should exist by the time this group reconvenes:

1. **An open compute substrate** — compiler, kernel DSL, operator library — running flagship open-weight models on five or more silicon families without per-vendor forks. FlagOS is named as a working seed.
2. **A working draft of an agentic-license category**, separate from existing OSI license families, in public comment.
3. **At least one production deployment of an independent, cryptographic verification layer** — a deterministic control plane — in a regulated industry, with public documentation of what it cost, what it proved, and what it did not.

The forum also identified deliberate gaps it could not close in one day. The role of AI in conflict, the labour-market dislocations of the next 24 months, the climate cost of frontier training at scale, and the governance of synthetic media all came up but did not have dedicated panels. These were registered as candidates for the next forum convening.

A Final Note

The day's host closed by reframing the opening question — *as AI becomes more autonomous, who is responsible?* — as one no single panel was meant to answer. What emerged instead was the beginning of a shared language: synergy as a contract, not a feeling; learning as more than efficiency; trust as engineered, not performed; and public goods as a design choice, not an afterthought. That language is Paris's contribution to the broader AI conversation, and the invitation is to carry it into GOSIM Paris this week, into Shenzhen in October, and into the rooms where the agentic transition will, regardless of any forum, continue to unfold.

Acknowledgements

The AI Vision Forum Paris 2026 was made possible by the work of its Organizing Committee, the GOSIM Paris team, and the participating researchers, founders, foundation stewards, regulators, and educators who travelled — in some cases very long distances — to make a single day of conversation possible.

This report was assembled from the full audio record of the day — approximately six and a half hours of speech across one keynote and four panels — transcribed and edited by the Organizing Committee. Direct quotations are reproduced as spoken; names and affiliations have been redacted in accordance with the Chatham House Rule under which the forum operated. Where the underlying audio contained gaps or unintelligible passages, those were treated as silences and not filled by speculation.

The forum is grateful to the partner organisations whose hosting, logistics, and operational support made the gathering possible, and to the many participants whose preparation, candour, and willingness to disagree productively in a single room made the discussion worth recording.

Forum website: paris2026.visionforum.ai Paris Initiative: paris2026.visionforum.ai/initiative/ · GitHub: github.com/aivisionforum/paris-2026-initiative Education Research papers: paris2026.visionforum.ai/education/ Forum report published: May 2026 Co-located with GOSIM Paris 2026