Sovereignty Is a Stack, Not a Label

Seven layers, three verdicts, an arbitration doctrine for AI in the regulated European environment.

1. Introduction

The previous article argued that digital sovereignty is not a political debate but an architectural constraint, and that it distributes itself across three planes: data localization, infrastructure qualification, and the operator's legal immunity. That doctrine remains valid for the IaaS-PaaS layer qualified under SecNumCloud. But it does not descend into the material stack that makes such infrastructure operable.

The SecNumCloud 3.2 qualification delivered to S3NS's PREMI3NS on 17 December 2025[S1] (and which Bleu, after validating its J0 milestone in April 2025, still aims for without yet having obtained it [S2]) qualifies the cloud service operator. It does not qualify the silicon that operator runs, nor the high-bandwidth memory that feeds its AI accelerators, nor the firmware that controls its servers, nor the transcontinental infrastructure across which its data transit. The doctrinal silence on these deeper layers produces a misleading effect of simplification: it suggests that a SecNumCloud qualification resolves the entire sovereignty problem, when in fact it addresses one specific layer, the operator's extraterritorial legal immunity.

The European political agreement of 7 May 2026 on the Omnibus VII package does not waive this work; it only displaces it in time. Parliament and Council validated the postponement of the high-risk obligations of Annex III from 2 August 2026 to 2 December 2027, and of Annex I to 2 August 2028[S3]. Most embedded industrial uses were exempted. Healthcare was not. For AI systems in healthcare under HDS, NIS2, and AI Act high-risk regimes, the preparation window has lengthened, but the perimeter of structural constraints remains intact. A critical supply-chain dependency remains an architectural property of the system, not a regulatory entry-into-force date.

The thesis of this note is fourfold. *First*, sovereignty is not a label but a stack composed of distinct layers, each calling for its own autonomous verdict. *Second*, the material supply chain now constitutes the deep layer of effective sovereignty, because it conditions operational continuity, maintenance, availability, and predictability of the system. *Third*, no commercial offering, European or otherwise, today closes the stack in pure sovereignty; technological autarky is an industrial fantasy, and only a composite, audited sovereignty is practicable. *Fourth*, the emergence of hardware attestation mechanisms

and usage telemetry on AI accelerators is progressively turning certain material components into objects governed at a distance, and changes the very nature of the problem.

Domain of validity. The thesis applies to AI systems deployed in regulated European environments (GDPR, HDS, NIS2, AI Act high-risk) whose operational continuity is part of the functional requirements. For such systems, the brutal interruption of the supply chain of a critical component is not a contractual nuisance: it is a service rupture with regulatory and clinical consequences. The doctrine is written under that constraint.

The doctrinal question is therefore no longer only *where is my data?* It becomes: *which components of my stack can be interrupted, altered, conditioned, or rendered indefensible by a third-party decision external to my governance perimeter?*

2. Securing Is Not Insourcing

Debates on digital sovereignty often conflate three notions that are neither equivalent nor substitutable: *insourcing*, *diversifying*, *securing*.

Insourcing means producing oneself, or under direct capital control, a critical component of the chain. This is the maximalist ambition of pure technological sovereignty. In the field of advanced semiconductors, it runs into a severe industrial reality. Producing a leading-edge wafer requires not only a fab but a complete upstream ecosystem: EUV lithography, photoresists, masks, special gases, advanced packaging, hybrid bonding, high-bandwidth memory, EDA software toolchains. No European actor today masters the entire chain. ASML produces the EUV machines, but Carl Zeiss SMT produces the optics, JSR and Tokyo Ohka dominate photoresists, Synopsys and Cadence dominate the EDA tools, TSMC concentrates leading-edge fabrication, Samsung and SK Hynix dominate HBM. The fragmentation is structural.

Diversifying means distributing dependencies so that no single rupture causes the system to stop. It is a resilience strategy, not a sovereignty strategy. Successful diversification can rest entirely on non-European suppliers, provided they are not all simultaneously exposed to the same geopolitical, climatic, or logistical risk.

Securing combines partial insourcing, diversification, strategic contracts, buffer stocks, volume guarantees, continuity clauses, and substitution trajectories. To secure a dependency is to turn it into an audited property of the system rather than a vulnerability one merely endures.

The decisive distinction: insourcing is an industrial strategy, diversifying is a resilience strategy, securing is a sovereignty strategy. The three can coexist. None replaces the other

two, and the composite sovereignty defended here presupposes all three activated simultaneously, in different proportions across layers.

3. The "Sovereignty by Control" Model: What It Solves, What It Does Not

The model validated by ANSSI with S3NS (and still pursued by Bleu) rests on one thesis: sovereignty depends less on the origin of the technology than on the effective control exercised over its operation. The distinction is serious. It has been institutionally validated. But its scope must be precisely delimited.

What the model does solve is real: the legal immunity of the cloud operator from direct extraterritorial injunctions. S3NS operates under French law, with French governance, French personnel, French data centers, and legal separation from Google Cloud through the capital control of Thales and a technological quarantine mechanism [1]. Bleu pursues a comparable architecture around Microsoft technologies, 100 % owned by Capgemini and Orange [2]. The risk of direct injunction under the CLOUD Act is thereby significantly reduced and controlled, without disappearing entirely from the analysis as long as the technological stack remains licensed from American actors.

What the model does not solve, and does not claim to solve, is the structural technological dependency on American suppliers for the layers it assembles. S3NS operates Google Cloud under license; Bleu will operate Azure and Microsoft 365 building blocks under license. Product roadmaps, deep cryptographic mechanisms, build chains, CPU microcodes, AI accelerators, and firmware dependencies remain under non-European technological control. Legal control has been acquired; control over technological production has not changed hands.

The doctrinal formula that follows is sober. *A SecNumCloud certification strongly reduces the extraterritorial legal risk at the operator level. It transfers neither the intellectual property of the technological stack, nor the mastery of product roadmaps, nor the control of the deep material components. It qualifies an operating layer under European law; it does not mechanically make sovereign the entire production chain.*

The decisive distinction: control and production are not the same thing. The model resolves a **sovereignty of use**; it does not resolve a **sovereignty of fabrication**. This nuance avoids two symmetrical errors: regarding these models as fictively sovereign, or regarding them as useless because not perfectly sovereign. Both stances are intellectually weak. The reality is more demanding: these models resolve one critical layer of sovereignty without resolving the others, and the doctrine must precisely name which layers remain open.

4. Seven Layers, Examined One by One

The sovereignty of an AI architecture can no longer be evaluated globally. It must be examined layer by layer, according to a uniform grid: *what sovereignty question does it pose? what is the current European verdict? what residual dependency persists? what CTO decision follows?* Seven layers appear today as structuring.

Layer 1 : Logic silicon (CPU, GPU, AI accelerators). The question is who designs, who manufactures, who can condition the operation. The European verdict for 2026 is one of partial design (with the SiPearl Rhea1 expected to ship during 2026[S4]) but absent leading-edge fabrication. The critical residual dependencies bear on TSMC for etching, ARM for the architecture, and cryptographic attestation under non-EU governance. The typical CTO decision is *compensable* provided it is covered by long-term contracts and a dated migration trajectory toward European accelerators as they mature.

Layer 2 : Advanced memory (HBM, DRAM, NAND). Who produces the memory that feeds the GPUs and servers? The European verdict is null: no actor, no announced industrial program. The market remains concentrated around SK Hynix, Samsung, and Micron, with 2026 capacity largely pre-empted by hyperscalers and major American AI actors[S5]. The CTO decision is *compensable* on a case-by-case basis through pluri-annual volume commitments, but it remains non-substitutable in the short term and it is the hardest dependency in the stack.

Layer 3 : Firmware and roots of trust. Who signs the microcode, the BMC, the secure boot? Coreboot and LinuxBoot offer open alternatives, but their adoption in production server environments remains marginal. Residual dependencies bear on Intel ME, AMD PSP, and proprietary BMCs (iLO, iDRAC, IPMI), which are not fully auditable. The CTO decision is *compensable* by independent audit, reproducible-signature requirements, and a quarantine zone for updates.

Layer 4 : Physical network (backbone, fiber, submarine cables). Who produces, who owns, who operates the routes of traffic? The European verdict is strong on cable manufacturing (Alcatel Submarine Networks, nationalized at 80 % by the French State in November 2024[S6]) and on fiber (Prysmian, Nexans), moderate on the backbone (Nokia, Ericsson). Residual dependencies bear on Cisco, Juniper, and Arista for data-center equipment, and on the growing hyperscaler ownership of transatlantic cables. The CTO decision is *acceptable* as long as regulated traffic remains under verified European control.

Layer 5 : Operating system and hypervisor. Who controls the kernel and the virtualization layer? The verdict is technically strong, with Linux, KVM, Xen, but weak in governance, with American foundations dominating (Linux Foundation, CNCF). Residual

dependencies bear on global kernel maintenance and on the governance of projects outside the EU. The CTO decision is *acceptable* on condition of build auditability and a documented defensive fork.

Layer 6 : Orchestration and platform. Who governs Kubernetes, OpenShift, the steering tools? The verdict is technically high, weak in governance (the CNCF remains dominated by American actors). The residual risk is that of a hostile fork or a release blockage. The CTO decision is *acceptable* if an internal fork is maintained at reasonable cost by the teams or by a European consortium.

Layer 7 : Qualified sovereign cloud. Who operates, under what law, with what extraterritorial immunity? The European verdict is growing, with a landscape of two coexisting philosophies: pure technological sovereignty (OVHcloud, Outscale, Scaleway, NumSpot) and sovereignty by control (S3NS qualified, Bleu pending). Residual dependencies bear on the licensed technological stack for hybrid models, or on the software ecosystem for indigenous actors. The CTO decision becomes *disqualifying* in sensitive healthcare contexts if the operator is not qualified SecNumCloud 3.2 or on a verifiable trajectory to qualification.

The structural verdict is clear: European sovereignty is strong at both ends of the stack (the operator can be SecNumCloud-qualified, the server can be Eviden, the cable can be ASN, the model can be Mistral) but fragile in the middle (leading-edge silicon, HBM, firmware, governance of open-source projects). This hourglass structure is a historical feature: Europe has retained its integration capabilities and applicative competencies while the production of critical components has shifted to Asia and the United States.

5. Silicon Changes Political Nature

The most significant change of the past two years lies not only in the industrial concentration of AI GPUs. It lies in the gradual emergence of a hardware-attestation infrastructure. The matter must be formulated with precision.

To date, no publicly documented mechanism constitutes an operational kill switch enabling the remote deactivation of NVIDIA GPUs. NVIDIA's position is explicit and publicly repeated by its leadership: no backdoor, no kill switch, and the telemetry used for the location-verification pilot remains read-only[S7]. The mechanism currently in pilot on the Blackwell architecture is an optional software agent, customer-installed on the customer's own servers, which leverages GPU telemetry, cryptographic attestation, and network-latency measurement to estimate the country in which a chip is operating (a precision comparable to IP geolocation, not embedded GPS[S7] [S8]). Extension to Hopper and Ampere is only at the exploration stage; nothing has been announced on that perimeter.

Several building blocks nonetheless converge, and it is their composition rather than each block in isolation that deserves doctrinal attention. On the export-policy side, the new licensing policy of the U.S. BIS, published on 13 January 2026, governs the export of the H200 to China through case-by-case review, mandatory third-party laboratory verification in the United States before shipment, a 25 % import duty, volume caps, and buyer-side compliance requirements including no-remote-access clauses[S9] [S10]. On the hardware side, the Blackwell architecture embeds reinforced cryptographic attestation capabilities relative to previous generations[S7]. On the software side, the location-verification pilot is being deployed as an optional service.

Taken in isolation, none of these elements turns a GPU into a remotely controlled terminal. Taken together, they establish a doctrinal precedent: *the possibility that the authorization to operate an AI accelerator may progressively become an attested, verified, and conditioned property of an authority external to the physical owner of the hardware*. The precedent bears on the infrastructure of control, not on a present capacity to interrupt. The distinction is essential to the rigor of the argument.

The decisive distinction: owning an AI accelerator no longer necessarily means controlling, in full, the attestation that this accelerator produces about its own conformity. The data center may be French, the operator may be European, the cluster may be physically located under national jurisdiction; the attestation reference, the cryptographic chain, and the conformity mechanisms can remain external. Local legal sovereignty then ceases to be sufficient for the most regulated classes of systems.

6. HBM: The Real Breaking Point

High-bandwidth memory is probably the most critical and least compensable dependency in the current European AI stack. The world market is concentrated around SK Hynix, Samsung, and Micron, and 2026 capacity is largely pre-empted by hyperscalers and major American AI actors[S5]. HBM3E is in structural shortage; HBM4 enters production in 2026 for the NVIDIA Rubin platform.

No European actor produces HBM. No public European HBM industrialization project has been announced to date. The geographic concentration is extreme: production essentially in South Korea and a secondary American Micron capacity; the HBM base die is manufactured by TSMC. The entire chain, for the world's AI GPUs, rests on three Asian suppliers and a geographic triad (South Korea, Taiwan, the United States via Micron) that does not include Europe.

This absence is not accidental. HBM requires competencies in vertical stacking (TSV, *through-silicon vias*), hybrid bonding, and precision thermals that Europe has not developed at industrial scale. The Chips Act 2 currently in discussion for 2026 could

identify HBM as a priority, but no industrialization project is today at the operational planning stage. The flagship Intel Magdeburg project, announced in 2022, was cancelled in July 2025[S11]. ESMC Dresden, slated to start production in 2027, targets 28/22 nm CMOS and 16/12 nm FinFET nodes for the automotive and industrial markets, and is not a leading-edge AI fab[S12].

The constraint touches directly on the principal limiting factor of modern AI GPUs: memory bandwidth. Without HBM, modern AI accelerators lose an essential part of their operational value. Even a hypothetical European accelerator would today still depend on a non-European HBM chain (Rhea1 itself embeds four stacks of Samsung HBM2E[S4 S13]).

HBM is the pebble in the shoe of European AI sovereignty. As long as it is there, the full stack cannot be qualified as sovereign without a usage caveat. This finding does not call for fatalism; it calls for prioritization.

7. Three Verdicts: Acceptable, Compensable, Disqualifying

A useful doctrine cannot remain descriptive; it must produce operational arbitration rules. A dependency becomes problematic when it cumulates four properties: it is critical, it is non-substitutable, it is not contractualized, and it is exposed to an uncontrollable third-party decision.

A dependency is not a sovereignty failure. It becomes one when it is simultaneously critical, non-substitutable, non-contractualized, and exposed to an uncontrollable third-party decision. The doctrine should therefore not ask: *is this layer European?* It should ask: *can this layer interrupt, alter, or render indefensible the regulated service, and what verifiable countermeasures exist?*

Three verdicts follow.

1. **Acceptable.** The dependency exists but remains technically substitutable at reasonable cost and delay, without asymmetric capacity for interruption by the third-party actor. Typical example: a Kubernetes orchestrator maintained by the CNCF, for which a documented defensive fork is technically maintainable by the internal teams or by a European consortium. The dependency exists, but it is not asymmetric: rupture would be costly to both parties, and the substitution effort is bounded.
2. **Compensable.** The dependency is not substitutable in the short term, but it can be covered by a verifiable contractual or operational mechanism: strategic stock, long-term supply contract with isolation clauses against political decisions, multi-

supplier redundancy, or a dated and financed migration trajectory. The Samsung HBM embedded in a Rhea1 falls into this category as long as a volume-commitment contract covers the lifetime of the cluster and a migration trajectory exists for the next generation. Compensation is the lucid acknowledgment of a dependency, not its denial.

3. **Disqualifying.** The dependency can interrupt or alter a critical system without realistic recourse. Typical example: an external attestation mechanism conditioning the use of an accelerator deployed in a vital clinical system, with no isolation clause against the regulatory evolution of the issuing jurisdiction and no documented alternative trajectory. If the dependency is disqualifying, the architecture must be revised, not softened by discourse.

This taxonomy is not a scoring grid. It is an arbitration scheme that transforms doctrine from *finding* into *governance*. Without this passage, composite sovereignty remains a descriptive exercise. With it, it becomes an instrument of decision.

8. Why the "Made in Europe" Paradigm Falls Short

Public debate is gradually sliding toward a binary logic: European or non-European. This grid is insufficient for two symmetrical reasons.

A system assembled in Europe with NVIDIA GPUs, Korean HBM, Intel ME firmware, American EDA dependencies, and TSMC etching does not become sovereign because its final integration is European. The nationality of the last assembler does not summarize the sovereignty of the chain. Buying *Made in Europe* does not manufacture *Made in Europe*, and this confusion produces counter-productive trade-offs in which the organization pays a premium for local assembly without substantially reducing its dependencies.

Conversely, a composite stack using certain non-European building blocks but exhibiting local governance, strong auditability, a migration trajectory, legal isolation, and documented supply-chain resilience can be far more defensible than an opaque "European" assembly whose dependencies have not been examined.

The decisive distinction: Made in Europe is a commercial category, audited composite sovereignty is an engineering category. They are not the same objects, and conflating them degrades the quality of the architectural decisions made on that basis.

9. Conclusion

The contemporary problem of digital sovereignty is no longer only a cloud problem. It is a problem of operational continuity governed under multiple dependencies. Models such as S3NS and Bleu resolve a real part of the problem, the operator's legal immunity, without resolving the sovereignty of the silicon they operate, of the HBM that feeds their AI accelerators, or of the deep firmware and attestation mechanisms. This does not invalidate them; it fixes their actual perimeter.

Sovereignty is not a pure state. It is an arbitration discipline that consists in identifying, layer by layer, which dependencies are tolerable, which are compensable, and which become disqualifying given the threat model of the system under consideration.

A SecNumCloud certification immunizes the operator, not the silicon it operates; the CLOUD Act reads organizational charts, export regimes read bills of materials, and AI accelerators are now beginning to learn how to attest to their own conditions of operation. In this context, composite sovereignty is not a pragmatic concession to the imperfection of the real : It is the only technically defensible doctrine for building regulated AI systems without succumbing to either the illusion of autarky or strategic resignation.

Sources and References

Footnotes

1. S3NS, *S3NS Announces SecNumCloud Qualification for PREMI3NS, its Trusted Cloud Offering*, press release of 19 December 2025 (ANSSI security visa delivered on 17 December 2025). https://www.s3ns.io/actualite/s3ns-annonce-qualification-sec-num-cloud

2. Bleu, *Bleu validates the J0 milestone of SecNumCloud 3.2 qualification*, press release of 17 April 2025. The qualification, targeted for the first half of 2026, has not been obtained to date. https://www.bleucloud.fr/bleu-valide-le-j0-de-la-qualification-secnumcloud-3-2/

3. Council of the European Union, *Artificial Intelligence: Council and Parliament agree to simplify and streamline rules*, official press release of 7 May 2026; European Commission, press release IP/26/1024 of 7 May 2026. Provisional political agreement under the Omnibus VII package: postponement of Annex III high-risk obligations to 2 December 2027 and Annex I to 2 August 2028; exemption of mots embedded industrial uses; medical devices remain in scope. Formal adoption expected before 2 August 2026. https://www.consilium.europa.eu/en/press/press-releases/2026/05/07/artificial-intelligence-council-and-parliament-agree-to-simplify-and-streamline-rules/ ; https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1024

4. SiPearl, product page *Rhea1 for HPC and AI Inference*: 80 ARM Neoverse V1 cores, fabrication by TSMC at N6 (7 nm), four embedded HBM2E stacks, first commercial samples available in early 2026, deployment in the Cluster module of the JUPITER exascale supercomputer (Forschungszentrum Jülich) planned for mid-2026. Tape-out confirmed in July 2025. https://sipearl.com/en/rhea1

5. On the HBM concentration around SK Hynix, Samsung, and Micron and the pre-emption of 2026 capacity by hyperscalers and American AI actors, see TrendForce and Astute Group analyses 2025-2026; Bird & Bird, *Digital Omnibus on AI Provisional Agreement Reached at the May Trilogue*, 8 May 2026. https://www.twobirds.com/en/insights/2026/digital-omnibus-on-ai-provisional-agreement-reached-at-the-may-trilogue

6. On the nationalization of Alcatel Submarine Networks by the French State in November 2024 (80 % equity stake for approximately 100 million euros), see communications of the Agence des participations de l'État and ASN's hearing before the French Senate, February 2026. ASN remains one of the three major global submarine-cable manufacturers, alongside TE SubCom (United States) and NEC (Japan).

7. Reuters, *Exclusive: Nvidia Builds Location Verification Tech That Could Help Fight Chip Smuggling*, 9-10 December 2025 (full republication via U.S. News). NVIDIA's repeated public position: no backdoor, no kill switch, no remote-disable capability (statement by David Reber Jr., Chief Security Officer at NVIDIA, August 2025: "No Backdoors. No Kill Switches. No Spyware."). Pilot initially deployed on Blackwell; extension to Hopper and Ampere at the exploration stage. https://www.usnews.com/news/top-news/articles/2025-12-09/exclusive-nvidia-builds-location-verification-tech-that-could-help-fight-chip-smuggling

8. On the underlying technical mechanism (Hardware-Enabled Mechanisms, cryptographic attestation, latency-calibrated landmark servers, precision equivalence with IP geolocation), see Tom's Hardware, *Nvidia develops software-based tracking for AI GPUs*, 10 December 2025. https://www.tomshardware.com/pc-components/gpus/nvidia-develops-software-based-tracking-for-ai-gpus-to-quash-smuggling-concerns-solution-devised-to-prevent-shipments-to-nations-with-export-controls-in-place

9. Bureau of Industry and Security (BIS, U.S. Department of Commerce), *Department of Commerce Revises License Review Policy on Semiconductors Exported to China*, press release of 13 January 2026. Case-by-case review for NVIDIA H200, AMD MI325X, and equivalent chips; mandatory third-party laboratory verification in the United States before shipment; buyer-side compliance requirements; no-remote-access clauses. https://www.bis.gov/press-release/department-commerce-revises-license-review-policy-semiconductors-exported-china

10. On the economic terms of H200 exports to China (25 % import duty in the United States as a constitutional workaround for the prohibition on export taxes, volume caps, Jensen Huang's statements of 17 March 2026), see Bloomberg, *Nvidia Gets US License for Small Amount of H200 Exports to China*, 26 February 2026. https://www.bloomberg.com/news/articles/2026-02-26/nvidia-gets-us-license-for-small-amount-of-h200-exports-to-china

11. On the cancellation of the Intel Magdeburg project in July 2025, originally announced in 2022 with €30 billion of investment and €9.9 billion of German subsidies committed under the European Chips Act, see Intel and German Ministry of Economy communications, July 2025.

12. On ESMC Dresden (joint venture TSMC + Bosch + Infineon + NXP): production scheduled to start in 2027, 28/22 nm CMOS and 16/12 nm FinFET nodes, automotive and industrial markets, target capacity of 480 000 wafers per year by 2029. See TSMC and partners' press release, 2025 updates.

13. On the memory details of Rhea1 (four Samsung HBM2E stacks in package), see *The Register*, *SiPearl finally tapes out Rhea1 supercomputer chip*, 9 July 2025. https://www.theregister.com/2025/07/09/sipearl_rhea1_tape_out/