Digital sovereignty is not a political debate. It is a condition for the capitalisation of performance.

Why the false dilemma between performance and sovereignty misses the regulatory reality of AI in regulated environments, and why compliance is a runtime-computable property, not an annual documentary declaration.

1. Introduction

The thesis of this article holds in one sentence: for an organisation deploying AI in the European regulated perimeter, the choice of cloud provider is not a performance choice arbitrated under compliance constraints; it is a compliance choice whose performance is a by-product, and whose juridical stability is the condition for capitalisation. Performance obtained in a juridically unstable space is not capitalised. It is recorded as a conditional operational debt, callable at the next regulatory event. And compliance itself then ceases to be an annual documentary declaration and becomes a *dual execution property: computed by the gate, proven by the ledger*. This is the architectural contribution proper to this doctrine, and the thread that runs through the entire text.

The current regulatory calendar transforms the nature of the question for those who wish to deploy in regulated environments. As of 16 May 2026, all active HDS certificates in France must conform to version 2 of the framework published by the order of 26 April 2024. On 2 August 2026, barring publication in the OJEU of the postponement adopted by the European Parliament on 26 March 2026, the obligations applicable to high-risk AI systems listed in Annex III of Regulation EU 2024/1689 enter into application. On 17 December 2025, ANSSI granted SecNumCloud 3.2 qualification to S3NS, validating for the first time a model of sovereignty-by-control over a partly non-European technology stack. Three deadlines, three normative bodies, one same substance: infrastructure compliance is no longer an arbitration option.

And yet, public debate continues to oscillate between two equally sterile framings, as if the matter were still in a phase of philosophical discussion. The first opposes protectionism and openness, as if the choice of a sovereign infrastructure stemmed from an ideological preference between Colbert and Smith. The second opposes hyperscaler performance and degraded sovereign cloud, as if compliance were merely an adjustment variable to be arbitrated against IO throughput. These two framings share the same vice: they treat sovereignty as a configuration parameter, whereas it is a structural property of

the system. And they invariably conflate three distinct notions that the law, for its part, separates with precision: the localisation of data, the qualification of the infrastructure, and the legal sovereignty of the operator.

The validity domain of this thesis must be made explicit before going further, on two distinct dimensions.

- First on the scope of use. It does not hold for all AI systems, nor for all contexts. It holds for systems deployed in the European regulated perimeter, processing sensitive data or exercising a function bearing on the security, the health, or the fundamental rights of persons. Outside this perimeter, classical cloud-computing arbitrations apply. Inside, they no longer apply in their classical form.
- Then on the layer addressed. Sovereignty is not monolithic: it is *stratified by layer*, and the sovereignty of one layer does not substitute for that of the others. At least four layers must be distinguished: the cloud infrastructure layer (datacentre, IaaS, PaaS, CaaS), the model layer (foundation models, fine-tunes, training data), the hardware layer (GPU accelerators, HBM memory, advanced lithography foundries), and the energy layer (powering of datacentres in the European electricity mix). The present article principally addresses the cloud infrastructure layer, because this is the layer whose regulatory obligations are dated and enforceable today (HDS v2, SecNumCloud 3.2, EUCC in progress), and because this is the layer that is architecturally instrumentable with the tools described herein. The other layers are treated in the discussion as authentic limits, and will be the subject of subsequent notes in this series. This restriction is explicit: it protects the argument against the legitimate objection that a cloud sovereignty without silicon sovereignty or energy sovereignty remains a logically bounded sovereignty.

2. Terminological clarification: three planes, three threats

The word sovereignty has become a semantic agglomerate that mixes three distinct questions, each having its own legal regime, its own assessment instruments, and its own admissibility criteria. To conflate these three planes is to believe one has resolved a compliance problem because one has ticked a geographical box. It is, in practice, the most frequent error in architecture committees.

But the separation of the planes, treated as a mere taxonomy, remains descriptive. To become doctrine, it must be stated as a *threat model*: each plane addresses a type of risk structurally irreducible to the other two. This orthogonality is not terminological, it is causal.

First plane: the localisation of data. Geographical question: where do the bits physically reside, and where are they processed. This is a direct requirement of several regimes: GDPR (Articles 45 and 46), HDS v2, SecNumCloud. *Threat addressed*: the unmastered physical and jurisdictional transfer of data to a territory whose protection regime is insufficient or non-equivalent. Localisation neutralises imposed flows. It neutralises nothing else. An Azure datacentre in Marseille certified HDS remains operated by a subsidiary whose parent company is subject to the U.S. CLOUD Act. The latitude of servers does not neutralise the nationality of the parent company's government.

Second plane: the qualification of the infrastructure. Normative question: has the host obtained the appropriate certification for the type of data concerned. HDS for healthcare, SecNumCloud for sensitive State, OIV (Operators of Vital Importance) and OSE (Operators of Essential Services) data, ISO 27001 for the generic foundation of information security management, future EUCC at European level. *Threat addressed*: the lack of operational and security mastery of the execution environment, regardless of the trust placed in the operator. Qualification neutralises technical and organisational failures. It neutralises nothing else. A host can be HDS-qualified and yet exposed to extra-European injunctions. For a healthcare establishment in France, hosting patient data on a non-HDS-certified cloud exposes the entity to risks of non-compliance, of administrative, contractual and possibly criminal sanctions depending on the responsibilities engaged, independently of the GDPR. Nor is qualification a rebadged ISO 27001: the HDS v2 and SecNumCloud 3.2 frameworks impose technical and organisational requirements that go far beyond the generic foundation, and that are audited by bodies accredited by COFRAC (the French national accreditation body) under prescriptive modalities.

Third plane: legal sovereignty. Question of extraterritoriality: is the operator subject to extra-European disclosure obligations, typically the U.S. CLOUD Act of 2018 (Stored Communications Act as amended), Section 702 of FISA, or the Executive Orders covering SIGINT collection. *Threat addressed*: coercive extraterritorial capture, that is, the capacity of a foreign authority to compel the operator to hand over data or to modify its operations regardless of the law applicable at the place of storage. This third layer is precisely what SecNumCloud 3.2 frontally addresses through its criteria of immunity to extra-Community laws, and what the EUCC label currently being finalised at European level will structure for its "high" level, on the basis of French proposals. A French HDS certification does not neutralise an American head office.

The CLOUD Act does not read datacentres, it reads org charts.

The orthogonality of the three planes rests on the orthogonality of the three threats: none reduces to the other two, none is annulled by the resolution of the other two. Localisation is geographical, qualification is normative, legal sovereignty is capitalistic and statutory. The three planes are jointly necessary for certain classes of data, and each produces its

own admissibility verdict independently of the others. An operator can be located in France, HDS-certified, and yet exposed to the CLOUD Act. An operator can be immune to the CLOUD Act but lack HDS qualification. An operator can be HDS-qualified and CLOUD-Act-immune but fail to guarantee the localisation of support operations outside the EU. Each of these combinations produces a partially compliant system, that is to say, in the regulated perimeter, non-compliant.

To remember it, the triad fits in one sentence. *Localisation says where the bits are. Qualification says what one has the right to do with them. Legal sovereignty says who can, in last resort, compel the operator to surrender them*. This ternary structure is the formulation to which an architecture committee must be able to return without opening a legal text.

3. Diagnosis: genealogy of a false debate

If the conflation of these three planes persists at the high level one observes in COMEX and CTO architecture committees, it is not for lack of information. It is by effect of framing. Three forces converge to maintain the false dilemma.

1. The first force is commercial. The American hyperscalers have built a narrative according to which European compliance is a configuration option, accessible through "sovereign regions" localised on the territory concerned. This narrative is technically true for the first plane (localisation) and for part of the second plane (sectoral HDS). It is silent on the third plane (legal immunity). The doctrinally exact diagnosis is not that the narrative is selective: selectivity is a journalistic critique, not an architectural one. The exact diagnosis is that *the sovereign region answers a question different from the one it leads its audience to believe it resolves*. The sovereign region answers a geographical constraint. The CLOUD Act acts on a relation of capitalistic control. The two do not live in the same logical space. The region cannot, by definition, neutralise a threat that is not exercised on geography. This is not a communication defect; it is an *incommensurability of layers*. To fault the sovereign region for lying about the CLOUD Act is to fault an altimeter for failing to measure latitude.
2. The second force is cultural. Cloud marketing has spent fifteen years associating sovereignty with degraded performance. The typical argument: sovereign cloud is more expensive, less tooled, less agile. This argument conflates a precise historical period (2018-2022, when the tooling gap was real and when the French offering lacked AI scalability) with a timeless truth. The SecNumCloud 3.2 qualification of S3NS in December 2025, which simultaneously covers IaaS, PaaS and CaaS and integrates Google Cloud's data and AI services under French control, breaks the argument. The qualification of Bleu (Capgemini-Orange-Microsoft) in progress, and the rapid evolution of the OVHcloud, Outscale,

Scaleway and NumSpot offerings, confirm it: the functional gap is shrinking. The cultural argument persists even as its empirical basis erodes.

3. The third force, deeper, is epistemic, and rests on a structural organisational failing: *each body of regulation has its master; no master has the coherence*. The CISO masters the GDPR, the CIO masters ISO 27001, the CTO masters HDS, the legal director masters the CLOUD Act. No one, in the standard decision-making chain, is held responsible for the articulation of the four corpora as simultaneous constraints on a single architecture. The system is validated by addition of partial validations, and the coherence of the whole is validated by no one. This is exactly what the doctrine of *governance-as-architecture*, developed earlier in this series, designated as the structural failure of organisations confusing documentary governance and architectural governance.

4. Structural limits of the "performance versus sovereignty" paradigm

The limits of the dominant paradigm are not of degree, they are of nature. Five structural limits make it inadequate to describe the current regulatory reality.

First limit: causal inversion. The paradigm assumes that compliance is a cost that reduces performance. This assumption is false for systems in regulated environments. Without qualification, there is no system to compare. Compliance does not reduce performance: it defines the space within which performance can be measured.

Performance is defined over the set of admissible architectures.

Outside this set, the magnitude "performance" has no defined value, because the system is not deployable. A comparison between a non-SecNumCloud-qualified hyperscaler cloud and a SecNumCloud-qualified cloud on sensitive State or OIV data is not a performance comparison: it is a comparison between a system and the absence of a system.

Second limit, prolonging the first: the non-capitalisability of non-admissible performance. Even in zones where temporary deployment remains possible (by regulatory tolerance, by lack of guidance, by conjunctural political decision), performance obtained in a juridically unstable space is not an asset. It is not recognised on the operational balance sheet as a durable capacity. It is recorded, on the economic plane, as a conditional debt callable at a future regulatory event whose probability of occurrence is not nil, indeed whose recent chronicle (Schrems I, Schrems II, Latombe appeal, instability of PCLOB-FTC-DPRC after change of administration) suggests is tendentially high. The technical language of architecture meets here the language of management control: the unguaranteed duration of use of a performance forbids its

accounting treatment as a fixed asset. An architecture that does not validate, by construction, its admissibility conditions does not authorise the amortisation of the performance it produces.

Third limit: the conflation of compliance and commercial certification. Holding an ISO 27001, ISO 27701 or ISO 27018 certification is not equivalent to a SecNumCloud or HDS qualification, even though the three corpora share part of their requirements. ISO 27001 is a generalist management system, self-declarative within its scope, audited against a neutral international standard. SecNumCloud is a prescriptive cloud-specific framework, which imposes capitalistic and legal requirements without equivalent in the ISO standard. HDS v2 is a sectoral healthcare framework with precisely delimited activities. To conflate the three is to conflate a generic driving licence with a hazardous-materials heavy-goods qualification: both attest to a competence, but what they attest to is not commensurable.

Fourth limit: the elision of the temporal factor. The dominant paradigm reasons in static equilibrium: at a given moment, is the architecture compliant. But obligations are dated and evolutive. HDS v2 transitions in May 2026. AI Act Annex III has its application date oscillating between August 2026 and December 2027 depending on the outcome of the Digital Omnibus. SecNumCloud 3.2 is in deployment, with twelve candidacies under instruction. The future European EUCC could replace SecNumCloud in the long run.

Compliance is not a state, it is a flow.

An architecture compliant today may no longer be so in eighteen months if it does not embed an integrated regulatory tracking mechanism. This temporal property disqualifies architectural choices that would have been dimensioned on a snapshot.

Fifth limit: the occlusion of the algorithmic stack. The dominant paradigm reasons on the infrastructure layer (IaaS, PaaS, CaaS) and largely ignores the software layer and the model layer. Yet the AI Act, for high-risk systems, imposes direct obligations on model and system providers (Article 16), obligations on deployers (Article 26), an obligation of CE marking and registration in the European database. These obligations do not bear on the datacentre, they bear on the software object and its life cycle. The sovereignty of the infrastructure does not exempt one from the sovereignty of the development process, the control of the foundation-model supply chain, and the traceability of training data. An AI system can be compliant when hosted on a sovereign cloud and yet rest on a foundation model whose training data, documented biases, and version governance are opaque. The compliance of the lower layer does not suffice.

To this fifth limit one must add, for healthcare AI systems specifically, that the "high-risk" qualification stemming from Annex III of the AI Regulation does not by itself summarise the applicable regime. Healthcare AI systems remain simultaneously subject to the sectoral regimes of MDR (Regulation 2017/745) and IVDR (Regulation 2017/746), with

their mechanisms for clinical evaluation, post-market surveillance, and management of substantial modifications. The articulation of the two regimes is partly settled by Recital 64 of the AI Regulation and by the first doctrinal documents of the Commission, but it remains an object of technical instruction for notified bodies. An architecture that lays claim to AI Act compliance without explicit articulation with MDR/IVDR is not compliant: it is partially assessed.

5. Architecture: sovereignty as a computable structural property

If sovereignty is not a configuration parameter, how does it translate into the architecture of an AI system deployed in a regulated environment. Four principles structure the answer, organised around two distinct functions: a gate function and a ledger function.

First principle: qualification is a port of the system, not an attribute of the environment. This formulation reprises the hexagonal doctrine developed in earlier articles. A system is said to be sovereign not because it is deployed in a sovereign environment, but because it requires, by construction, a qualified environment to function. This requirement is exposed as an explicit external port, verified at startup, verified at each sensitive operation, and triggering a structured refusal (the WITHHOLD doctrine of the *StandardDecisionPolicy*) when the condition is not met. The practical consequence: a properly architected clinical AI system must be able to refuse to run on a non-HDS-qualified cloud, and this refusal is a positive functionality, not a degradation.

A system that consents to run anywhere is a system that does not know where it is legal.

Second principle: sovereignty is composed, not bought. No single provider supplies the complete stack. The typical stack for a healthcare AI system in France combines a SecNumCloud-qualified IaaS (S3NS, OVHcloud, Outscale, NumSpot or equivalent), a compliant orchestrator (Kubernetes operated under sovereign control, qualified CaaS container), an HDS-qualified data store, foundation models whose supply chain is documented and AI-Act-compatible (which makes problematic, for the most sensitive uses, recourse to closed models whose training data and fine-tunes are not auditable), and an application layer whose event traceability is ensured by an event-driven architecture compatible with the logging requirements of Article 12 of the AI Act. This composition is not a commercial product. It is architectural work, indissociable from systems engineering, and it is the proper subject of the industrial architect.

Third principle: sovereignty, treated as a port, becomes a computable state of the runtime; and it decomposes into a *gate function*. If sovereignty is exposed as an admissibility condition tested at execution time, then it ceases to be an annual documentary declaration and becomes a synchronous primitive: at each sensitive

operation, the system answers the question "does this operation have the right to take place here, now, with this workload identity, on this hardware". The answer is binary. The instrument chain already exists in the cloud-native stack: *policy-as-code* (OPA, Kyverno, Cedar) to express rules declaratively and testably, Kubernetes *admission control* to refuse deployment of a workload on a non-qualified node, *workload identity* cryptographically attested and bound to a certified environment, *confidential computing* with remote attestation (Intel TDX, AMD SEV-SNP, ARM CCA) proving to the application layer that the hardware and hypervisor are in an uncorrupted state, *sovereign posture verification* assembling these signals into a continuous attestation. The hexagonal sovereignty port is therefore not an abstract architectural concept: it has an implementation programme already existing in productive stacks, and it transforms compliance from a declarative property into an observable property.

Fourth principle: to the gate function is added the *ledger function*. Where the gate answers "can one, here, now", the ledger answers "can one, afterwards, prove". The two functions are distinct: the gate is synchronous and Boolean, the ledger is asynchronous and historical. A workload may be correctly blocked by the gate without any useful trace being produced; and conversely, an exhaustive trace can be produced without any admission decision being taken. The two ports are jointly necessary, and they cannot be conflated in a single primitive. The AI Act requires automatic event logging throughout the entire lifetime of the system (Article 12), technical documentation conforming to Annex IV, and the conservation of logs allowing audit and explanation of decisions. SecNumCloud 3.2 requires intrusion testing throughout the qualification life cycle. HDS v2 requires annual audits. The infrastructure of the ledger is different from that of the gate: event-driven architecture with append-only immutable logs, cryptographic traceability of key events, qualified storage for the conservation durations imposed (ten years for certain healthcare regimes), documented and stable event schemas. It is here that EDA, addressed in the previous note, ceases to be a choice of elegance and becomes the technical condition of regulatory traceability.

The distinction that cuts, and that condenses the four principles: *a system is not configured to be compliant, it is architected to be compliant; and its compliance is not attested, it is computed by the gate and traced by the ledger*.

6. Articulation with prior doctrine: a structural homology

This thesis does not fall from the sky. It is part of a chain of articles published under the Twingital Institute label over several months and of which it constitutes the regulatory complement. *Governance-as-architecture* posited that governance requirements are not documentary layers separable from the system but structural constraints that build it. *Event-driven architecture as essential complement to agentic AI* posited that traceability, asynchrony, and event persistence are not optimisation choices but structural conditions

of agentic composition. *Hexagonal architecture applied to clinical AI systems* posited that reliability ports and adapters (Out-of-Distribution, Calibration, Audit, Refusal) are first-order components, not wrappers added a posteriori.

The homology is not accidental. The same structural *pattern* runs through the entire series, and it takes the generic form of a distinction between a partial signal and the real systemic property. *Benchmark ≠ production: a score on a validation set is not a guarantee of behaviour in deployed distribution. Monitoring ≠ governance: an observation dashboard is neither an arbitration device nor a mechanism of imputability. Certification ≠ sovereignty: an audit paper is not a continuous property of the runtime. Localisation ≠ legal immunity: a geographical coordinate is not a capitalistic status.* Each time, the same error: taking an observable *proxy* for the property it only partially indicates, and conflating the measurement of an attribute with its reality.

The thesis of architectural sovereignty therefore extends the series exactly along the homological axis. It treats regulatory compliance as a structural constraint whose satisfaction is an architectural property of the system, not a deployment option. It introduces a fourth port to the hexagonal reliability ports: a *sovereignty port* that continuously verifies the compliance of the execution environment with applicable regulatory requirements, and that triggers a structured refusal in case of violation. This port is not a regulatory gadget superimposed: it is of the same nature as the OOD or calibration ports, because it bears on the conditions of validity of the system, not on its performance. And it participates in the same programme: *to translate into computable architectural properties what the community usually treats as documentary commitments*.

A final homology, already signalled in the introduction as a delimitation of the domain, deserves to be reformulated as a systemic prolongation: sovereignty is stratified by layer, and the architectural rigour of one layer does not mechanically entail the rigour of the underlying layers. The present doctrine treats the cloud infrastructure layer, because enforceable regulatory instruments exist there today. It does not treat the model layer, nor the hardware layer (Nvidia/CUDA concentration, HBM dependency on Samsung-SK Hynix-Micron, TSMC concentration on advanced nodes), nor the energy layer (powering of datacentres in the strained European electricity mix). A cloud sovereignty rigorously architected, but combined with an unexamined dependency on the underlying material or energy layers, remains a logically bounded sovereignty. This doctrine therefore does not exhaust the subject; it treats the sub-question for which a dated regulatory framework exists, and proposes an instrumentable architectural target. The other layers fall under other tools and other doctrines, and their treatment belongs to subsequent notes in this series.

7. Inadequacy of the market paradigm

The current market paradigm, dominated by American hyperscalers and the associated "Cloud de Confiance" offerings (Bleu, S3NS), proposes a partial answer to the problem but does not resolve it. The distinction between the two models now in competition on the French market deserves to be made explicit.

The model called *sovereignty-by-control*, validated by the SecNumCloud 3.2 qualification of S3NS on 17 December 2025, rests on a legal and operational construction: the operator is a French-law company entirely controlled by a European shareholder (Thales, in the case of S3NS), operates on a technology stack under licence (Google Cloud), with French localisation of datacentres and operations, and contractual guarantees of isolation from the American technology parent. This model has been recognised viable by ANSSI, which thereby settled a legal debate of several years. The model called *pure technological sovereignty*, defended by OVHcloud, Outscale, Scaleway, and NumSpot, rests on an entirely European stack: European technology, European operator, European capital. The two models now coexist and each has its zone of validity.

For the architect, the consequence is the following: the choice between the two models is not reducible to an ideological preference or a direct cost calculation. It depends on the type of data and the type of use. For routine-exploitation health data, without massive territorial aggregation, under documented governance, both models can suit. For territorial predictive medicine systems aggregating nominative cohorts at large scale, or for SecNumCloud-related use cases bearing information-classification constraints reaching back to the General Inter-Ministerial Instruction 1300 or equivalent, the pure technological sovereignty model offers a supplementary guarantee that may be necessary. The choice must be motivated by a documented risk analysis, not by a cost intuition.

The predictable objection is the argument of European integration through the future EUCC: why concern oneself with SecNumCloud if a unified European certification is being finalised. The objection ignores two facts. The first: as long as the EUCC regulation is not adopted and has not designated its "high" level as equivalent to SecNumCloud, the French qualification remains the applicable instrument. The second: the joint ANSSI-BSI declaration of March 2026 on cloud sovereignty criteria suggests a Franco-German alignment that consolidates, rather than replaces, the high-level standard. To wait for the EUCC is to wait for a normalisation that establishes itself *through* SecNumCloud, not *against* it.

8. Illustrative instances

The instances that follow are implementation grounds, not general proofs. They are drawn from use cases and reference architectures, and they have no vocation to publicly

document the certification status of effectively deployed systems. They show what, in practice, the architectural translation of the three planes of sovereignty looks like for healthcare AI systems in regulated European environments.

OCTOPUS: real-world ambispective study on an mNSCLC BRAF V600E cohort, n=184 patients distributed across five European countries, treated according to differentiated therapeutic sequences. Digital-twin pipeline trained on 299 nominative clinical features extracted from 59 SAS datasets and 37 SDTM domains, with conditional synthetic generation, TSTR validation (95.2% on downstream tasks, which does not prove general statistical indistinguishability but constitutes a strong indicator of operational fidelity for simulation tasks), and SurvTRACE counterfactual simulator for analysing therapeutic sequences with competing events. On the three planes: data localisation in the EEA, HDS qualification of the execution environment for the nominative phases, immunity to extra-Community laws required for the cross-country aggregation phase. The choice of cloud operator was not a performance arbitration; it was a regulatory feasibility condition.

Sentinelle IA / PREDICARE: territorial predictive medicine programme, aggregation of heterogeneous data (PMSI, the French national hospital activity database; SNDS, the French national health data system; biology; exposome) at the scale of a territorial basin, predictive models for the early detection of clinical transitions. Structural constraint: cross-source aggregation at large scale is exactly the scenario in which legal sovereignty becomes critical, because extraterritorial requests on massive nominative bases are precisely what SecNumCloud 3.2 addresses. The retained stack could not rest on components exposed to the CLOUD Act, regardless of their sectoral HDS certification. The system is conditioned by this requirement, and this conditionality is exposed as an external port to the system.

ToxTwin: predictive system of molecular toxicity, deployed in freemium mode with quotas and calibrated validation, hosted on French sovereign infrastructure: servers with GPU and VRAM on our premises running Ubuntu 24.04. A simpler case than the two preceding ones because the data are not nominative. Nevertheless, AI Act Annex III, in its reading by certain regulators, may classify predictive healthcare tools as high-risk systems according to their use in the clinical decision pipeline, and use in support of pharmaceutical R&D raises the question of articulation with IVDR for associated diagnostic modules. Decision traceability, registration in the European database (Article 71), and Annex IV documentation are architectural constraints that the system integrates by construction, not by ulterior patch.

These three instances do not prove the generality of the thesis. They show that, at different scales and sensitivities, concrete architecture obeys the same principles: sovereignty is treated as a port of the system, explicitly composed, and auditably traced.

9. Discussion and authentic limits

This doctrine calls for several limits that must be formulated explicitly, on pain of losing the credibility of the argument.

First limit: the opportunity cost of non-adoption of rapid innovations. American hyperscalers have a technological lead on certain specialised AI services (proprietary foundation models, dedicated hardware accelerators, integrated MLOps tool ecosystems). Choosing an entirely sovereign stack may lead to a delay in access to certain cutting-edge capabilities. This limit is real. It is treated by composition: use of European or open-weight models compatible with audit requirements for the sensitive layers, and strict isolation of non-sensitive layers if they rely on exposed components. It is not treated by denial.

Second limit: regulatory uncertainty on the AI Act calendar. The probable adoption of the Digital Omnibus postponing the application of Annex III to 2 December 2027 reduces short-term pressure. An organisation may be tempted to defer its compliance investments. This calculation is short-termist: the date is postponed, the substantive requirements are not. To build an AI system in 2026 without compliance architecture is to plan a massive refactoring for 2027. The cost of a retrofit systematically exceeds the cost of an initially compliant architecture.

Third limit: the residual contestability of the Data Privacy Framework. The DPF survived the Latombe appeal in September 2025, but the political conditions in the United States (modifications of the PCLOB, of the FTC, of the composition of the DPRC) suggest that its stability is not acquired. Schrems himself indicated in February 2026 that the Commission could suspend the agreement of its own motion before any new appeal. An architecture that rests on the DPF for its data transfers to the United States inherits a non-negligible continuity risk. This limit does not overturn the thesis, it hardens it: it adds a reason to avoid unmastered transfers towards unstable jurisdictions.

Fourth limit: the sovereignty of the upper and underlying layers. The infrastructure stack can be sovereign without the model stack being so, and conversely. A European foundation model trained on a non-sovereign cloud is no more sovereign than an American model trained on OVHcloud. The sovereignty of models is a subject in the process of structuring at the European level (Mistral, Aleph Alpha, various consortia) but it is not settled. And beyond the model, the hardware layer (Nvidia concentration, HBM dependency, TSMC foundries) and the energy layer remain points of major vulnerability that the doctrine presented here does not address. This limit signals that the cloud infrastructure layer, while regulatorily instrumentable today, is only one layer in a stack whose complete sovereignty requires other tools, other calendars, and other doctrines.

Fifth limit: the practical transition of existing systems. Every organisation has an existing IT and AI patrimony that was not designed under these constraints. The passage

to a sovereign architecture does not happen by decree. It requires a transition plan that ranks systems according to their sensitivity, their regulatory exposure, and their migration cost. This transition is an architectural subject in its own right, deserving a dedicated article. The present note posits the architectural target, not the trajectory.

Sixth limit: the risk of political appropriation. The doctrine of architectural sovereignty can be instrumentalised by economic actors who push the sovereign argument for protectionist rather than technical reasons. This appropriation is a risk for the credibility of the argument. The counter-fire is the requirement of precision: not to conflate the three planes, not to substitute political rhetoric for regulatory analysis, not to validate an offering because it is French but because it is qualified. Architectural rigour is, here too, the best antidote to ideology.

10. Conclusion

The false dilemma "performance versus sovereignty" presupposes that the two are on the same axis and that one can arbitrate linearly between them. They are not. Performance is measured within the perimeter of admissible architectures. Outside this perimeter, there is no arbitration: there is no system. And even where deployment remains temporarily possible by tolerance, the performance obtained is not capitalised: it is recorded as a conditional operational debt, callable at the next regulatory event.

Three planes to separate, three planes to validate, three planes to trace. *Localisation says where the bits are. Qualification says what one has the right to do with them. Legal sovereignty says who can, in last resort, compel the operator to surrender them.* None of the three planes substitutes for the others. None is an adjustment variable. And each addresses a threat that the other two do not neutralise: unmastered jurisdictional transfer, lack of operational mastery, coercive extraterritorial capture.

In perimeters subject to qualification obligations, a clinical AI system deployed on a non-qualified infrastructure ceases to be juridically admissible to exploitation: its performance, real though it may be, cannot be capitalised, and its operational existence remains suspended on a regulatory tolerance whose deadline does not depend on the architect. Compliance, in this perspective, is no longer an annual documentary declaration: it becomes an observable state of the runtime, computed at the moment of the gate and traced in the ledger, refusable when the conditions of admissibility are no longer met.

The false dilemma being lifted, the architectural question that opens is no longer that of the choice between cloud providers. It is that of the exact form that the sovereignty port must take in a productive stack: which gate runtime primitives, which cryptographic attestations, which event schemas for the ledger, which structured-refusal thresholds, which explicit articulation with the sectoral MDR and IVDR regimes for healthcare

systems, which transition architecture for existing patrimonies, and beyond, how to address the model, hardware, and energy layers that remain here outside the field. This open question is the proper subject of the industrial architect, and it constitutes the natural sequel to this doctrine. The debate on digital sovereignty changes nature for those who wish to deploy. It does not close: it shifts.