A Model Is Not Sovereign Because It Is Open

Three chains of proof that decide whether an AI system remains certifiable after deployment: inspectable weights, documentable training data, governable versions.

Take Away :

- "Model sovereignty is not read in the publisher's passport. It is read in three chains of proof instructed separately: weights one can inspect and freeze, training data one can document, versions one can govern over time. Open-weight opens the first door. It does not close the genesis, nor the lifecycle."
- "Open-weight is not open-source. The OSI Open Source AI Definition 1.0 (28 October 2024) requires Data Information, Code, and Parameters, with a description detailed enough that a skilled person can recreate a substantially equivalent system. Llama 4, Mistral, and DeepSeek publish their weights without publishing their training data in a reproducible form. They are open-weight. They are not open-source within the meaning of OSAID. OLMo 2 (Allen Institute for AI) is, and remains the most robust frontier-adjacent example."
- "The AI Act and OSAID do not measure the same property. The Training Data Summary Template published by the European Commission on 24 July 2025 operationalises Article 53(1)(d) of the AI Regulation: it requires narrative transparency on training content oriented toward copyright opt-out. OSAID requires technical reproducibility allowing the recreation of an equivalent system. Compliance with the first does not produce compliance with the second."
- "A model that cannot be frozen is a model that cannot be certified. Version governance is the most neglected axis of model sovereignty, and probably the most decisive for regulated deployment. For a medical device software certified under the AI Regulation and MDR/IVDR integrating a given version of a foundation model, this exact version is part of the certified perimeter. Any substantial modification can trigger a conformity review."
- "The triad converts into a CTO arbitration matrix. For each chain (weights, data, versions), an operational question, a minimum proof, a risk if absent. This matrix transforms doctrine from observation into governance, exactly as the three-verdict grid did for the material layer in Volume 2."

- "The decisive distinction: open-weight is a commercial category, composite auditability is an engineering category. Made in Europe model is a commercial category, audited composite model sovereignty is an engineering category. Conflating the two degrades the quality of the architectural decisions made on that basis."

Key Words

- Model sovereignty open-weight OSAID 1.0 GPAI Code of Practice
- AI Act Article 53 Training Data Summary version governance
- version freeze defensive fork OLMo 2 Mistral Apache 2.0
- CTO arbitration matrix series: digital-sovereignty seriesIndex: 3 seriesTitle: "Digital Sovereignty" relatedTo:
- architectural-sovereignty sovereignty-stack-not-label
- eda-essential-complement-agentic-ai
- hexagonal-architecture-condition-governability
- ai-governance-architecture

1. Introduction

The first two volumes of this series established a thesis in two stages. Volume 1 showed that the sovereignty of execution infrastructure is not a political debate but a condition of performance capitalisation, computed by the guard and traced in the registry. Volume 2 descended into the material stack to show that sovereignty is not a label but a stack of seven layers, each calling for an autonomous arbitration verdict (acceptable, compensable, disqualifying). One layer remained whose importance both notes flagged without addressing: the model layer.

That is the object of this note. And the doctrinal question is more treacherous than it appears, because the model layer is, among the seven layers, the one where marketing produces the most illusions of sovereignty. Mistral is French, therefore Mistral is sovereign. Llama 4 is open-weight, therefore Llama 4 is free. DeepSeek V4 is under MIT license, therefore DeepSeek V4 is open. Cohere and Aleph Alpha have just merged under a sovereign banner, therefore Europe has its champion. Each of these statements is, taken literally, verifiable. And each is, taken as a conclusion, false for a regulated deployment.

The thesis of this note is sober. *Model sovereignty is not read in the publisher's passport. It is read in three chains of proof: weights one can inspect and freeze, training data one can document, versions one can govern over time. Open-weight opens the first door. It does not close the genesis, nor the lifecycle.* In a regulated environment, this distinction is not academic: it decides whether the system remains certifiable after its deployment.

Domain of validity. As in the two previous volumes, the thesis applies to AI systems deployed in regulated European environments (GDPR, HDS, NIS2, AI Act high-risk, MDR, IVDR), and for which decision traceability, auditability by notified bodies, and pluri-annual legal stability are part of the functional requirements. For unregulated uses, some of the constraints described here loosen. For regulated uses in clinical health, algorithmic justice, critical infrastructure, and prudential-stake financial services, they tighten.

Subject to formal adoption of the Omnibus VII text, whose provisional political agreement was announced on 7 May 2026 and mentioned in Volume 2, the application calendar of high-risk obligations is now to be treated as moving. This modifies neither Article 53 GPAI, nor the structure of the threat model, nor the enforcement calendar of the GPAI Code of Practice from 2 August 2026.

2. Three Confusions to Dispel

Before instructing the three chains of proof, three confusions must be named. They structure the current false debate, and they produce, in architecture committees, conclusions that are inverse to regulatory reality.

First confusion: open-weight and open-source.

The Open Source Initiative published version 1.0 of the Open Source AI Definition on 28 October 2024. The text requires three cumulative components for an AI system to qualify as open-source: *Data Information* detailed enough that a skilled person can recreate a substantially equivalent system from identical or similar data; *Code* covering the entire pipeline (data preparation, training, validation, inference) under an OSI-approved license; and *Parameters* (the model's weights, and possibly intermediate checkpoints) under open terms. Against these three criteria, almost none of the foundation models commonly described as open-source in the technical press is compliant. Llama 4 publishes its weights under a license that is not OSI-approved, contains a 700-million-monthly-active-users threshold above which a separate license must be negotiated with Meta, and bears European territorial restrictions on multimodal capabilities. Mistral publishes its weights under Apache 2.0, which satisfies the *Parameters* component, but does not publish the *Data Information* component in reproducible form. DeepSeek publishes its weights under full MIT license, but exposes its training corpora through description rather than access. None of these three publishers is OSAID-compliant. OLMo 2 (Allen Institute for AI) is, because it simultaneously publishes weights, complete

code, data, training recipes, logs, and checkpoints. A few other models approach it (EleutherAI's Pythia, listed by OSI among the references aligned with the definition), but OLMo 2 remains the most robust frontier-adjacent example for a regulated deployer in 2026.

Open-weight means that the weights are downloadable. Open-source within the meaning of OSAID means that a substantially equivalent system is reconstructible. These are not the same objects. The use value of a model is in its weights; the audit value of a model is in the trajectory that produced its weights. For a regulated deployment, it is the second value that determines defensibility.

Second confusion: AI Act compliance and OSAID compliance.

The European Commission published the *Training Data Summary Template* (TDS Template) on 24 July 2025, operationalising the obligation of Article 53(1)(d) of the AI Regulation. This article requires GPAI providers to publish a sufficiently detailed summary of the content used for training. The template, in its current form, settles for high-level narration: dataset sizes, source types (web crawl, licensed data, openly licensed data), filtering and opt-out compliance measures, aggregated categories. This obligation is less demanding than that of OSAID 1.0, which asks for a description detailed enough to allow the recreation of an equivalent system by a skilled person. *The two regimes do not measure the same property.* AI Act Article 53 compliance is a narrative transparency obligation oriented toward the European market and the exercise of opt-out rights by rightsholders. OSAID compliance is a technical reproducibility obligation oriented toward the open community. Recent practice confirms that this asymmetry is exploited: a publisher can be AI Act compliant without being OSAID compliant, and that is today the case of the entire landscape of open models save a few exceptions documented by OSI of which OLMo 2 remains the most robust.

This asymmetry is not a design flaw of the regulation. It reflects an explicit political trade-off between incentivising openness and protecting publishers' commercial interests. It has a practical consequence for the regulated deployer: *the simple Article 53 compliance of the publisher does not suffice to constitute an audit chain*. The deployer must demand, contractually or technically, additional transparency elements. Or it must observe by construction that the audit chain is not closed and arbitrate accordingly.

Third confusion: publisher nationality and sovereignty of fabrication.

Volume 2 showed that the nationality of the last assembler does not summarise the sovereignty of the chain. The same rule applies to the model layer, with particular severity. A model developed in Europe by a European publisher, but trained on infrastructure whose silicon is non-European, whose high-bandwidth memory is non-European, and whose datasets contain significant fractions of non-EU-produced content, does not become a sovereign model by the sole fact of its publisher. The publisher resolves, at the

model layer, a question of applicable law, of corporate governance, and of product roadmap. It does not resolve the silicon dependency or the memory dependency, which are structurally non-European and which were the object of Volume 2.

This distinction is exactly homologous to the one made in Volume 2 between *sovereignty of use* and *sovereignty of fabrication* for the cloud operator layer. Conflating the two degrades the quality of architectural decisions.

3. First Chain: Auditability of Weights

Auditing a model is not auditing its weights. This formulation, which inverts common intuition, deserves development.

A model's weights are very-large-dimension numerical objects, produced by a training process whose auditability depends on the documentation of everything that contributed to computing them. The weights themselves bear no explicit trace of the data that engendered them. Having the weights enables inference, fine-tuning, red-teaming, and the post-hoc measurement of operational properties. It does not, by itself, allow answering the questions that auditability raises in the regulatory sense: *what documented biases does this model inherit, what classes of content has it seen during training, what security updates have been applied and when?*

Weight auditability, as a structured property of the system, therefore requires an attestation chain that descends below the weights. This chain comprises at least five components. *First*, the stable cryptographic identity of the weights (a hash published and signed by the publisher, whose absence transforms any audit into argumentation by trust). *Second*, the traceability of intermediate training checkpoints, which allows reconstructing the learning trajectory and locating the appearance of emergent properties. *Third*, the documentation of alignment and post-training techniques, with exposure of alignment datasets and red-teaming protocols. *Fourth*, the declaration of post-publication modifications (security patches, moderation adjustments, official fine-tunes) with strict semantic versioning. *Fifth*, the hardware attestation of the inference environment, which connects to the sovereignty port of Volume 1 and the silicon layer of Volume 2.

This chain is today partially available from some publishers and entirely opaque from others. OLMo 2 from Allen AI publishes weights, training code, datasets, recipes, logs, and intermediate checkpoints, which makes it one of the rare foundation models that simultaneously satisfies OSAID and an end-to-end cryptographic audit. Mistral exposes its weights under Apache 2.0 but does not publish all intermediate checkpoints, nor the entirety of its alignment protocols. DeepSeek exposes its weights under MIT and a technical report, but its alignment chain and the history of its post-publication

corrections are not fully exposed. Llama 4 exposes its weights under a non-OSI license with European territorial restrictions and does not publish its alignment datasets.

The decisive distinction: a model whose weights are downloadable is a queryable model, not an auditable model. Querying allows measuring the observed behaviour at the output; auditing allows tracing back to the conditions of production of that behaviour. For a medical device software, this difference is precisely what AI Act high-risk requires and what notified bodies are codifying in their evaluation grids for 2027.

4. Second Chain: Sovereignty of Training Data

If weight auditability fixes the perimeter of what can be known about a model, the sovereignty of training data fixes the perimeter of what can be asserted about its lawfulness, its representativeness, and its regulatory stability.

This second chain is where the asymmetry between regulatory regimes is most pronounced. Article 53(1)(d) of the AI Regulation, operationalised by the TDS Template of 24 July 2025, asks GPAI providers for a narrative description of the contents used for training. The Copyright chapter of the GPAI Code of Practice of 10 July 2025 adds commitments on respect for technical opt-out standards (robots.txt and future IETF specifications), non-use of notoriously problematic sources, and complaint mechanisms accessible to rightsholders. This regime is demanding in terms of copyright governance. It does not, by itself, constitute a reproducible audit chain. A publisher can be compliant while exposing only a summarised and aggregated version of its training corpora.

OSAID 1.0 demands more: not a summary description but a description detailed enough that a skilled person can recreate a substantially equivalent system. In the practice of the open community, this translates into the publication of source datasets themselves or, failing that, precise manifests (sources, crawl dates, quality filters, deduplication methods, synthetic generation prompts where applicable). This reproducibility obligation is more demanding than the AI Regulation's transparency obligation. It is precisely this difference that separates a queryable model from an auditable model.

For a regulated health deployment, the distinction is not academic. What is certified under the AI Regulation and MDR/IVDR is not the foundation model, but *a medical device software integrating a given version of a foundation model*. The clinical validation of this device requires, according to the current guidance of notified bodies and the emerging standards for the articulation between AI Act and MDR/IVDR, a characterisation of the training population (demographic, geographic, clinical distribution), an identification of documented bias sources, and a demonstration of representativeness with respect to the target population. This characterisation is not achievable from an AI Act summary; it requires OSAID-grade manifests. *A model that does not reveal its training data is not clinically auditable, it is queryable at the output*, and the notified body that must validate

the integration of the model into the device faces an audit chain gap that no measure of validation behaviour can close.

The predictable objection is that commercial confidentiality and copyright constraints prevent the full publication of corpora. That is correct. The doctrine does not require every model to be OSAID. It requires that *the decision to deploy a non-OSAID model in a regulated device be taken in awareness of the auditability delta*, and that this delta be compensated by other mechanisms: exhaustive independent tests on the target population, contracts of asymmetric transparency with the publisher, modification notification clauses, or a migration trajectory toward a more auditable model as the ecosystem matures. Compensation is the lucid acknowledgment of a dependency, not its denial, exactly as in the *compensable* verdict of Volume 2.

A second dimension of this chain, structurally important, is the question of synthetic data. Several recent models include in their training pipeline synthetic datasets produced by other models, themselves sometimes trained on corpora whose audit chain is not closed. This recursivity multiplies layers of opacity. Auditing a distilled model requires auditing the teacher model as well, and the conditions under which the teacher generated the synthetic data. *Training recursivity multiplies the audit chains that must be closed simultaneously.*

5. Third Chain: Governance of Versions

The third chain is the most neglected in public discourse, and probably the most decisive for regulated deployment. A model is not a fixed object. It is an object whose identity is always indexed by a version, and whose every version has its own regulatory lifecycle.

Article 53 of the AI Regulation and the Transparency chapter of the GPAI Code of Practice require documentation retained for at least ten years from the placement on the market of each version of a GPAI. This obligation has a consequence often poorly understood: *a deprecated model remains legally active for a decade after its commercial retirement*. The publisher may cease to maintain it, but cannot erase its documentary obligations, and the deployer who integrated it into a regulated device remains itself bound to be able to present, over ten years, the audit of the model in the version that was used at the time of each historical automated decision.

This temporal property interacts with a well-documented industrial reality: foundation model publishers deprecate their models at a rapid pace. DeepSeek announced the retirement of the deepseek-chat and deepseek-reasoner API endpoints for 24 July 2026, eighteen months after the commissioning of the generation they serve. Mistral renews its lineup at a rate of several major versions per year. Meta launched Muse Spark in April 2026, a sign that the Llama trajectory is not a stable line but a revisable product strategy. *The pace of model deprecation is structurally shorter than the regulatory archiving*

duration that they trigger. For a medical device software operated for ten years, the foundation model that is part of it will have been deprecated seven to eight times.

This desynchronisation produces a doctrinal statement that must be held for what it is, that is, a criterion of certifiability before being a criterion of governance.

A model that cannot be frozen is a model that cannot be certified.

If the exact version of the weights used at the time of initial conformity cannot be archived, reconstructed, or replayed identically, then conformity itself loses its object. Certification bears on a medical device software integrating a given version of the foundation model. Without the capacity to freeze, the certified system and the operational system diverge with each silent update from the publisher, and the displayed conformity becomes a documentary fiction.

Three architectural consequences follow.

1. *First*, the **version freeze for conformity**. For a medical device software that obtains a CE mark under the AI Regulation and MDR/IVDR, the exact version of the foundation model integrated at the time of initial conformity is part of the certified perimeter. Any substantial modification (version change, additional fine-tune, major security update) potentially triggers a conformity review. The deployer cannot simply follow the publisher's roadmap; it must decide, version by version, whether the update is applied, in what timeframe, and according to what re-validation protocol. This decision is an architectural primitive exposed as an external port of the system, in line with the hexagonal doctrine developed earlier in this series.
2. *Second*, the **defensive fork**. When the publisher deprecates a model whose frozen version is still in operation at regulated deployers, two scenarios coexist. The publisher maintains older versions available via API or download, or it withdraws them. For open-weight models, the defensive fork is technically practicable: the deployer downloads the exact version of the weights, archives it within its own perimeter, and continues to operate it locally after deprecation by the publisher. For purely API-based models (typically the proprietary frontiers), this fork is not possible and conformity depends on the publisher's retention policy, contractual or otherwise. *The capacity for defensive fork transforms an external model into an archivable internal component*.
3. *Third*, the **audited migration chain**. When a deployer decides to switch from one version to the next, the transition is not neutral. It requires a demonstration of conformity of the new version on the system's functional requirements, a comparison of performance on the target population, an analysis of inherited or modified biases, and depending on the regulatory context, a new evaluation by the notified body. This migration chain is an engineering discipline of its own, distinct from classic application updates. It rests architecturally on the system's capacity

to maintain two active versions simultaneously during the transition phase, to compare their outputs on representative cases, and to switch atomically with documented rollback point.

The decisive distinction: *the version of a model is not a detail, it is the regulatory object*. The doctrine of *governance-as-architecture* developed earlier in this series posited that governance requirements are not documentary layers separable from the system; it extends here to the model layer with an additional requirement: the version is the unit of regulatory discourse, and the architecture must make it explicitly manipulable.

6. CTO Arbitration Matrix

A useful doctrine cannot remain descriptive. It must produce a decision instrument tenable in an architecture committee. The triad of three chains converts into a matrix with four columns: the sovereignty plane concerned, the operational question the CTO must be able to formulate, the minimum proof to require from the publisher or to produce by construction, and the risk that materialises if the proof is missing.

Chain	CTO Question	Minimum proof	Risk if absent
Weights	Can I inspect and freeze the effective model?	Signed checksum, local archive license, downloadable artefact, auditable fine-tune	Opaque dependency
Data	Can I defend the genesis of the model?	Enriched TDS, OSAID-grade Data Information, independent audit of sources, documentation of biases	Unassignable bias
Versions	Can I maintain conformity over time?	Contractual freeze, EOL notice, validated migration protocol, defensive fork	Certification rupture

This matrix is the concrete instrument of the triad. It allows, for each foundation model considered, the production of a three-line arbitration sheet that the architecture committee can sign or refuse. When the minimum proof is obtained or producible, the chain is green. When it is not, the corresponding risk must be named, formally accepted by the decision-maker, or compensated by an explicit alternative mechanism. When it

can be neither obtained nor compensated, the model does not cross the door of the regulated system.

The matrix does not select a vendor; it forces traceability of the decision. This is its proper architectural value, and this is exactly the role the three-verdict grid (acceptable, compensable, disqualifying) plays for the material layer of Volume 2.

7. Three Test Cases on the Matrix

The matrix is not validated as a market panorama. It is validated on concrete test cases. Three trajectories currently available to a regulated European deployer in 2026 serve here as instruction of the matrix, not as a vendor ranking.

Mistral is the European publisher of reference for open-weight foundation models. The current lineup (Large 3 published in December 2025 under Apache 2.0, Small 4 published in March 2026 under Apache 2.0) satisfies the *Weights* chain through public availability of artefacts and a license allowing defensive fork. On the *Data* chain, Mistral will have to publish, like any GPAI provider operating on the European market, the information required by Article 53 of the AI Regulation and the TDS Template; this does not mean OSAID-grade publication of training data. On the *Versions* chain, exposure via Hugging Face makes contractual freeze practicable, but the renewal pace requires the deployer to put in place an explicit audited migration protocol. Training compute rests on NVIDIA accelerators, which places Mistral in the composite sovereignty situation of Volume 2: publisher under European law, non-European silicon dependency in transition. *On the matrix, Mistral is compensable*.

OLMo 2 from Allen Institute for AI is the structural exception. The model fully satisfies the three chains: public weights under Apache 2.0, complete training code, published source datasets, recipes, logs, and accessible intermediate checkpoints. It is today the most robust frontier-adjacent foundation model qualifying as OSAID-compliant by construction. The trade-off is real: OLMo 2 remains below the performance frontier of larger-scale proprietary models for the most demanding tasks, and its publisher is American, which places the question of legal sovereignty back within the perimeter of Volume 1. *On the matrix, OLMo 2 is acceptable* for deployments where frontier performance is not the discriminating criterion.

Cohere and Aleph Alpha, since the announcement on 24 April 2026 of their merger, constitute a transatlantic entity whose planned execution stack runs through STACKIT, the sovereign cloud operated by Schwarz Digits. The Cohere lineup is not open-weight; the models are accessible by API. On the matrix: *Weights* chain not satisfied by download (defensive fork not practicable outside ad hoc contractual arrangement); *Data* chain not exposed to OSAID standard; *Versions* chain under control of the merged publisher, partially compensated by the guarantee of execution on European sovereign

infrastructure. *On the matrix, this profile is compensable on infrastructure, not covered on model auditability without a dedicated contract*.

These three test cases illustrate the central doctrinal observation: *none closes the triad without concession*. The instrument does not elect a vendor, it instructs the decision.

8. Three Verdicts for the Model Layer

Volume 2 established a three-verdict grid for the material layer. This grid applies to the model layer with a proper instantiation.

1. **Acceptable.** The concession on the chains exists, but it is technically substitutable at reasonable cost and timeframe, and it does not create an asymmetry of power that the publisher can exploit unilaterally. Portability to another model is documented and bounded. The dependency exists, but it is symmetrical in the sense that rupture would be costly to both parties.
2. **Compensable.** The concession is not substitutable in the short term, but it can be covered by verifiable contractual or operational mechanisms. Compensation takes the form of four cumulative arrangements: the framework contract exposing the publisher's commitments beyond the AI Act minimum; the systematic defensive fork (weights archived locally for each version deployed in regulated production); the re-validation protocol triggered by any substantial modification; and the documented migration trajectory toward a more auditable model as the ecosystem matures. Compensation is the lucid acknowledgment of a dependency, not its denial.
3. **Disqualifying.** The concession can interrupt, alter, or render indefensible a critical system without realistic recourse. Typical example: a model whose publisher can, by silent update, modify the behaviour of the deployed version without notification, without the possibility of contractual rollback, and without a public manifest of modifications. For a medical device software, this property makes MDR/IVDR conformity impossible to maintain over time. Another example: a model whose hardware attestation architecture is conditioned by an authority external to the deployer, without an isolation clause against the regulatory evolution of the issuing jurisdiction, which then joins the disqualification described at the silicon layer in Volume 2. *If the concession is disqualifying, the architecture must be revised, not softened by discourse*.

9. Why the "Made in Europe Model" Paradigm Falls Short

The European debate on foundation model sovereignty slides toward a binary logic homologous to the one criticised in Volume 2 for the material layer: European model or non-European model. This grid is insufficient for two symmetrical reasons.

A model developed in Europe by a European publisher, but trained on infrastructure whose silicon is non-European, whose high-bandwidth memory is non-European, and whose datasets contain significant fractions of non-EU-produced content, does not become a sovereign model by the sole fact of its publisher. The nationality of the last integrator does not summarise the sovereignty of the model's fabrication chain. Buying a European model does not manufacture a European model in the sense of the three chains.

Conversely, the use of a non-European model in an architecture that has strong local governance (defensive fork, local weights archive, re-validation protocol, additional transparency contract), an audit chain closed by the downstream side (exhaustive independent tests on the target population, runtime conformity monitoring), and a migration trajectory toward a more auditable model, can be far more defensible than an opaque "European" model whose three chains have not been instructed.

The decisive distinction: *open-weight is a commercial category, composite auditability is an engineering category. Made in Europe model is a commercial category, audited composite model sovereignty is an engineering category.* These are not the same objects, and conflating them degrades the quality of architectural decisions.

This homology with Volume 2 is not accidental. It reflects the fact that the doctrinal pattern of composite sovereignty, posited for the material layer, is strictly transposable to the model layer because it describes the same problem structure: a stratification of dependencies, each of which calls for its own arbitration verdict, and whose composition produces, or does not produce, a defensible auditability.

10. Conclusion

The false dilemma "European model versus American model" is the analogue, for the model layer, of the false dilemma "performance versus sovereignty" that Volume 1 deconstructed for the cloud layer. Both suppositions share the same vice: they treat sovereignty as a binary attribute of an object, when it is a structured property of an architecture.

Model sovereignty is not read in the publisher's passport. It is read in three chains of proof: weights one can inspect and freeze, training data one can document, versions one can govern over time. Open-weight opens the first door. It does not close the genesis, nor the lifecycle. In a regulated environment, this distinction is not academic: it decides whether the system remains certifiable after its deployment.

The triad that structures this note converts into a CTO arbitration matrix: three questions, three minimum proofs, three risks. This is not a scoring grid; it is a discipline of decision traceability that transforms the evaluation of a foundation model into a systematic instruction of the three chains, followed by a verdict (acceptable, compensable,

disqualifying) proper to each deployment. This discipline is demanding, but it is the only one technically defensible for building regulated AI systems without falling into either the illusion of open-weight as a guarantee of sovereignty, or the ideological rejection of models whose fabrication chain is not entirely European.

The European 2026 landscape confirms this analysis. Mistral consolidates its independent frontier trajectory while remaining dependent on non-European silicon and memory. The Cohere-Aleph Alpha merger traces a path of transatlantic composite sovereignty backed by the STACKIT cloud, explicitly renouncing isolated frontier competition. OLMo 2 demonstrates that complete OSAID auditability is technically achievable, at the price of a concession on frontier performance. None of these trajectories closes the triad for all deployments. *None ever will*, because model sovereignty, like that of infrastructure and that of the material stack, is a composite property that is built, by deployment, and not a pure state that one would choose.

A SecNumCloud certification immunises the operator, not the silicon it operates; the silicon operates a model, not the auditability chain that makes it defensible; and a model that cannot be frozen is a model that cannot be certified. The extension of architectural sovereignty doctrine to the model layer does not change the nature of the criterion, it instantiates it on a new layer of the stack.

Volume 4 of this series will address the energy layer, the last of the four layers identified in Volume 1 and signalled as an authentic limit of the doctrine. The sequence will then be complete, and the architectural arbitration of which the triad is the instrument will be able to be instructed in full knowledge of the entire stack: infrastructure, fabrication, model, and the energy that powers them all.

Sources and References

1. **Open Source Initiative**: *The Open Source AI Definition 1.0*, 28 October 2024. Defines the components required for an AI system to qualify as open-source: Data Information, Code, Parameters. opensource.org/ai/open-source-ai-definition

2. **AI Office, European Commission**: *General-Purpose AI Code of Practice*, final version published 10 July 2025. Three chapters: Transparency, Copyright, Safety and Security. Application 2 August 2025 (new models), enforcement 2 August 2026; models placed on the market before 2 August 2025 have until 2 August 2027 to comply. digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

3. **European Commission**: *Training Data Summary Template*, published 24 July 2025, operationalising Article 53(1)(d) of the AI Regulation for GPAI providers. Requires a high-level narrative description of training content oriented toward copyright opt-out.

4. **Council of the European Union**: *Artificial Intelligence: Council and Parliament agree to simplify and streamline rules*, press release of 7 May 2026. Provisional political agreement on the Omnibus VII package, subject to formal adoption of the final text. The calendar is interpreted by some legal sources as a postponement of Annex III obligations to 2 December 2027 and Annex I obligations to 2 August 2028; the definitive wording remains to be monitored. consilium.europa.eu

5. **Mistral AI**: product page Mistral Large 3, made available in December 2025 under Apache 2.0 license, Mixture-of-Experts architecture with 675 billion total parameters and 41 billion active per token, trained on NVIDIA H200 GPUs. mistral.ai

6. **Mistral AI**: product page Mistral Small 4, published in March 2026 under Apache 2.0 license, Mixture-of-Experts architecture with 119 billion total parameters and 6 billion active per token. mistral.ai

7. **Meta AI**: Llama 4 published on 5 April 2025 under the Llama 4 Community License, not approved by the Open Source Initiative. 700-million-monthly-active-users threshold clause. European territorial restrictions on multimodal capabilities. ai.meta.com/blog/llama-4-multimodal-intelligence/

8. **DeepSeek**: official announcement of the retirement of the deepseek-chat and deepseek-reasoner API endpoints for 24 July 2026, a concrete example of the commercial lifecycle of a foundation model used in production.

9. **Allen Institute for AI**: OLMo 2 family, identified by OSI as compliant with OSAID 1.0. Weights, training code, datasets, recipes, logs, and checkpoints published under Apache 2.0. allenai.org

10. **EleutherAI**: Pythia family, also identified by OSI among the references aligned with OSAID 1.0, reference instruments for interpretability research and the reproducibility of training trajectories.

11. **Cohere and Aleph Alpha**: announcement of 24 April 2026 of the merger between Cohere (Toronto) and Aleph Alpha (Heidelberg), with execution on the STACKIT sovereign cloud operated by Schwarz Digits. Aleph Alpha's strategic pivot away from frontier model competition documented in 2024.

12. **Latham & Watkins**: *EU AI Act: GPAI Model Obligations in Force and Final GPAI Code of Practice in Place*, 30 July 2025. Analysis of the fines regime (up to 15 million euros or 3% of global annual turnover, Article 101) applicable to GPAI non-compliance from 2 August 2026. lw.com