LinkTec Labs · Publications

Sovereign AI by design

Reading notice · V1 · 15 October 2026

Document type and scope: a synthesis of sources and a proposed reference architecture. This is not a peer-reviewed publication, a report of a LinkTec experiment or a client reference.

Scope correction: manuscript wording such as “deploys”, “builds”, “audit-ready”, “sovereign” and similar terms describes architecture objectives requiring project-specific verification. It attests no client deployment, operated equipment, certification or LinkTec result. Working hypotheses are neither completed research nor a commitment to publish.

Quoted quantitative results belong to their authors and evaluation settings. General claims, recommendations and thresholds are not guarantees for every workload. The reading review and checks of principal references are not experimental replication or exhaustive scientific validation.

Execution is on client premises or a cloud selected with the client. A model’s origin, size and openness determine neither execution location nor commercial rights. Versions, licences, access and dependencies require verification. No US-based model API without the client’s request.

Regulatory update: the manuscript describes a June 2026 position and is not a current legal timetable. The European Commission page consulted on 15 October 2026 reports entry into force of the AI Omnibus on 27 July 2026 and distinguishes transparency and high-risk deadlines. Check the applicable text, operator role and exceptions before drawing a conclusion. GDPR Article 17 erasure has conditions and exceptions; RAG, encryption or European hosting alone does not guarantee compliance. Provider levels and schemes cited are not LinkTec certifications. This note establishes no absolute immunity from foreign jurisdiction.

linktec.fr/recherche · This notice clarifies the scope of the following manuscript.

LinkTec LABS



RESEARCH NOTE N°03 · APPLIED AI & DATA

Sovereign AI **by design**

Building for GDPR and the EU AI Act from the architecture.

ABSTRACT

Sovereignty and compliance are not procurement checkboxes to be ticked after a system is built; they are architectural properties that must be designed in from the first line. As the EU AI Act's transparency obligations take effect and the GDPR's data-subject rights collide with how models actually store information, the question for any organisation deploying AI on real personal data becomes concrete: where does the data live, who can compel its disclosure, and can a customer's record actually be erased on request? This note separates *residency* from *sovereignty*, maps the 2026 regulatory reality precisely, and makes a single architectural argument with sharp consequences, that retrieval, not training, is the primitive that makes data-subject rights technically real. We close with the sovereign, audit-ready reference architecture LinkTec Labs deploys, completing the arc begun in Notes N°01 (context) and N°02 (memory).

01, FROM "WHERE" TO "WHO CONTROLS"

Sovereignty is an architecture, not a checkbox

For a decade, "is our data in Europe?" was treated as a settled question the moment a cloud region said *eu-west*. That era is over. The binding question in 2026 is no longer *where* data sits but *who can compel its disclosure* and *what can actually be done with it*, and neither is answered by a region selector. Roughly half of Western European enterprises are now actively re-evaluating their dependence on non-European cloud providers, and the gap they have discovered is the gap between residency and sovereignty.

The distinction is not academic. A hyperscaler region in Frankfurt satisfies *data residency*, storage and processing happen on EU soil. It does not, on its own, deliver *sovereignty*: a US-headquartered provider remains subject to the US CLOUD Act, which can compel disclosure of data regardless of where it physically resides. Sovereignty asks a harder question, whether the provider, and the legal order it answers to, can be reached by a foreign jurisdiction at all. For a hospital, a bank, or a defence supplier, that difference is the difference between a compliant system and an unlawful one.

The thesis of this note follows directly, and it is the thread that ties the LinkTec Labs series together: *compliance and sovereignty cannot be bolted on.* You cannot retrofit erasure onto a model that has memorised personal data, nor residency onto a pipeline that already shipped a prompt abroad. They are properties of the architecture or they are absent. The good news, and the substance of this note, is that the architecture that delivers them is now well understood, and it is buildable.

KEY TERMS

A shared vocabulary for the rest of this note. Specialists may skip ahead.

Data residency	A guarantee that data is stored and processed within a given territory (e.g. the EU).
Data sovereignty	Residency *plus* immunity from foreign jurisdiction over the data and its operator.
Sovereign cloud	Infrastructure whose technology, operations and control sit fully under EU law; not reachable by the US CLOUD Act.
Confidential computing	Hardware trusted execution environments (TEEs) that keep data encrypted even while in use.
Sovereign RAG	Retrieval-augmented generation run on private/ sovereign infrastructure with access control and audit.
By design (Art. 25)	GDPR duty to embed data protection into systems by default, from the outset.
DPIA (Art. 35)	Data Protection Impact Assessment, required for high-risk processing.

Right to erasure (Art. 17)	A data subject's right to have their personal data deleted ("right to be forgotten").
Machine unlearning	Techniques to remove the influence of specific data from trained model weights, still approximate.
GPAI	General-Purpose AI model, governed by a dedicated chapter and Code of Practice under the AI Act.
Article 50	The AI Act's horizontal *transparency* obligations (e.g. disclosing AI interaction, synthetic content).
Audit trail	An exportable record linking each output to its inputs, sources and decisions, for accountability.

02, THE REGULATORY REALITY

What is actually due, and when

Sovereign-by-design begins with knowing precisely what the law requires, and the 2026 picture rewards precision, because the headlines have been misleading. Two regimes apply at once: the EU AI Act governs the *system*; the GDPR governs the *personal data* inside it.

The EU AI Act, after the Digital Omnibus

On 7 May 2026 the Council and Parliament reached provisional political agreement on the "Digital Omnibus on AI," confirmed at the Parliament's plenary on 16 June 2026, though, importantly, it is *not yet law* pending Council adoption and publication in the Official Journal[2]. The Omnibus postponed the heaviest obligations: high-risk systems under Annex III now apply from **2 December 2027**, and high-risk AI embedded in regulated products from 2 August 2028[2]. But the obligation most relevant to anyone shipping a generative feature was *not* delayed: the Article 50 **transparency** duties take effect on **2 August 2026**, with watermarking and synthetic-content disclosure required by 2 December 2026[1][2]. For general-purpose models, the GPAI Code of Practice, published by the EU AI Office on 10 July 2025 and signed by some two dozen providers, including the major labs and Europe's own Mistral AI, remains the primary compliance pathway[3].

> THE DATE THAT MATTERS
>
> **2 August 2026, Article 50 transparency is live.** Disclosure of AI interaction and of synthetic content is a horizontal duty that applies regardless of risk tier and was *not* postponed by the Omnibus. Treat it as the binding near-term deadline; treat the high-risk 2027 dates as the planning horizon[1][2].

The GDPR, and what the EDPB now expects of models

The GDPR is the older and, for personal data, the more demanding regime: lawful basis and transparency (Arts. 5-6), security (Art. 32), data protection by design and by default (Art. 25), and a

Data Protection Impact Assessment for high-risk processing (Art. 35)[5]. In December 2024 the European Data Protection Board issued Opinion 28/2024, the field's clearest signal yet on AI and personal data[4]. Three findings matter here: a model is "anonymous" only where it is *very unlikely* both to identify individuals and to have their data extracted by querying, assessed case by case, with the burden on the controller; legitimate interest can ground AI processing only through a strict three-step test; and, decisively for architecture, the EDPB acknowledges that honouring the right to erasure against a *trained model* is "technically complex," and points to mitigations including data-removal mechanisms and unlearning techniques[4]. Hold that last point; Section 05 turns it into a design principle.

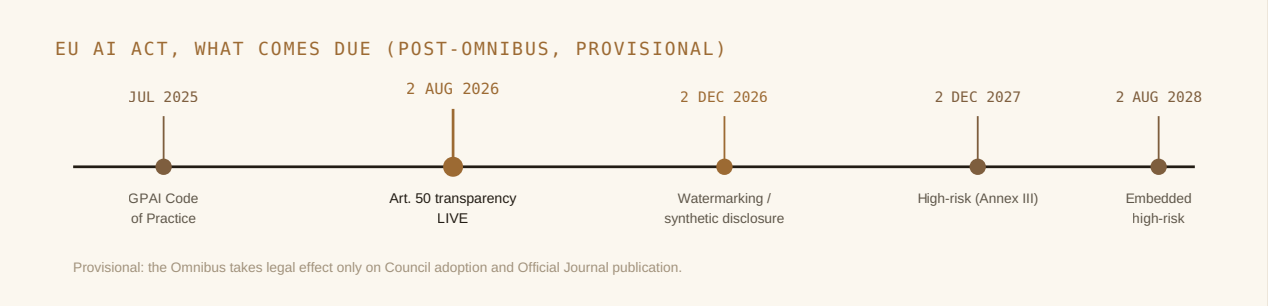


Fig. 1. The EU AI Act compliance horizon after the Digital Omnibus. Article 50 transparency (2 Aug 2026) is the binding near-term duty; high-risk obligations move to 2027-2028. Dates provisional pending publication[1][2].

03, RESIDENCY IS NOT SOVEREIGNTY

Three rungs, one CLOUD Act

Between "our data is in the EU" and "our data is beyond foreign reach" lie three distinct rungs, and conflating them is the most common procurement error of 2026. The European Commission's own framework now makes the ladder explicit. In its €180 million sovereign-cloud procurement, awarded to four European consortiums, built around OVHcloud, STACKIT, Scaleway and, with Mistral, the Proximus group, three were certified **SEAL-3** (a fully European stack across hardware, software, capital and operations, with no non-EU interference) and one **SEAL-2** (non-European technology permitted, provided operations, control and capital remain entirely European)[6]. The forthcoming EU Cloud Services Scheme (EUCS) is expected to encode equivalent sovereignty tiers into formal cybersecurity certification, alongside the Gaia-X labelling criteria[7][8].

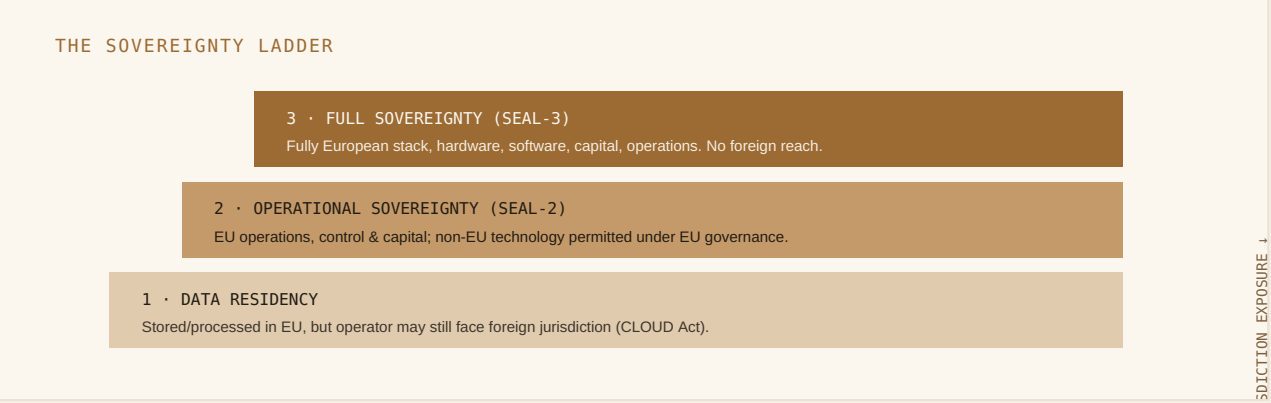


Fig. 2. Residency is the first rung, not the destination. Sovereignty rises through operational (SEAL-2) to full (SEAL-3) control; the EU's €180M procurement certifies real providers at each level[6].

The practical consequence is that open-weight, European-built models now make full sovereignty *achievable* rather than aspirational. Mistral's models, hostable on OVHcloud, Scaleway and other sovereign clouds, give regulated European institutions, banks, insurers, hospitals, public bodies, a credible path to capable AI that never leaves EU jurisdiction. Sovereignty is no longer a trade-off against capability; it is a deployment choice.

04, THE TECHNICAL SUBSTRATE

Where data lives, who can reach it, how it is protected in use

Sovereignty resolves into three technical guarantees, and a serious architecture provides all three. **Location:** data and inference run in an EU region or on-premise. **Jurisdiction:** the operator is beyond foreign compulsion, the SEAL-3 property, not merely a region label. **Protection in use:** this is the one most often forgotten. Data at rest and in transit is routinely encrypted; data *in use*, during inference, has historically been exposed in memory. *Confidential computing* closes that gap: hardware trusted execution environments, now available on data-centre GPUs, keep data and model weights encrypted even while being processed, so that neither the cloud operator nor a compromised host can read them[9].

These guarantees do not require sending every workload to a sovereign enclave. The mature pattern is a *routed* architecture: a classifier splits traffic so that anything touching personal or regulated data is handled by a private model on sovereign, confidential infrastructure, while non-sensitive workloads may use a global frontier model for raw capability. Sovereignty becomes a property applied where it is legally required, not a blanket tax on performance.

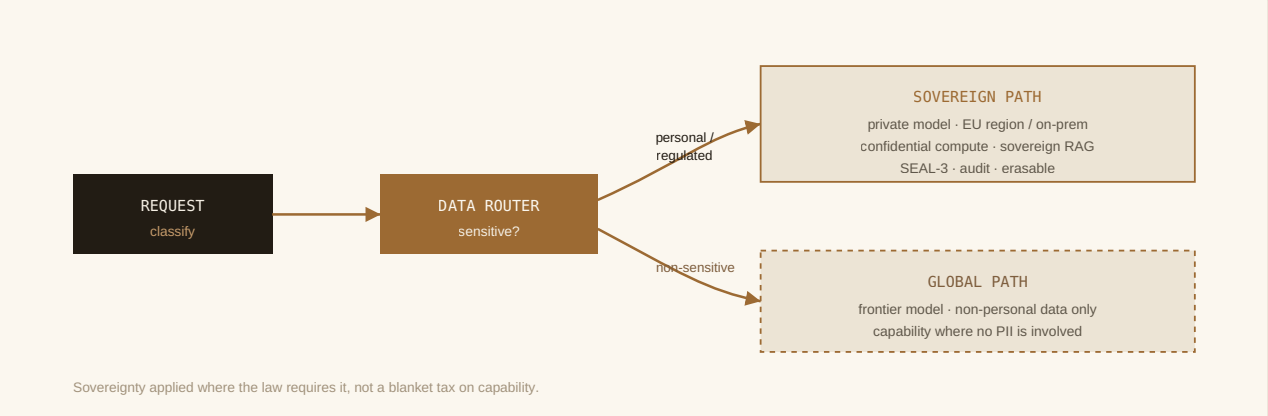


Fig. 3. Routed sovereignty. A classifier sends personal and regulated data down a sovereign path (private model, confidential compute, sovereign RAG) and reserves global frontier models for non-sensitive work[9].

05, WHY RETRIEVAL IS THE SOVEREIGNTY PRIMITIVE

The argument that changes the build

Here is the architectural insight that turns compliance from a burden into a design choice, and it is the through-line of this entire series. The GDPR grants a data subject the right to have their personal data erased. Against a model that has *learned* that data into its weights, this right is, in the EDPB's own words, "technically complex": the information is diffused across billions of parameters, and machine unlearning remains approximate, unverifiable, and an open research problem[4]. You cannot prove you have forgotten someone you trained on.

But personal data held in a *retrieval store*, and in the typed agent memory of Note N°02, is different in kind. It is addressable. It can be located by subject, corrected when wrong, and deleted on request, with a record that it was. The same is true of the synthetic context that enters a prompt at inference and is discarded after. *The right to erasure is intractable in weights and trivial in a store.* The conclusion is an engineering instruction with legal force:

> Don't train on personal data you may have to forget. Retrieve it. Sovereign RAG turns an unsolved machine-learning problem, unlearning, into a solved data-engineering one: delete the record.

This reframes the entire series. Note N°01 argued that retrieval keeps the *signal-to-token ratio* high; Note N°02 argued that typed, governed memory keeps continuity *affordable*; this note adds the decisive third reason, that keeping personal data *out of the weights and in a governed store* is what makes the data subject's rights, and therefore lawful deployment, technically achievable at all. Context architecture, agent memory, and sovereignty are not three topics. They are one architecture, seen from three sides. Retrieval is what makes it compliant by construction.

06, SOVEREIGN RAG, CONCRETELY

From principle to control

"Sovereign RAG" is not a product; it is a set of controls, each mapping a legal duty to a concrete mechanism. Retrieval must be **access-controlled**: the store enforces per-document permissions so the model can only surface what the asking user is entitled to see, retrieval inherits the organisation's authorisation model rather than bypassing it. Personal data must be **minimised and guarded**: PII detection and redaction at ingestion and at egress, and a top-k context that carries only what the task needs. And every interaction must be **auditable**: an exportable trace from prompt, through the exact sources retrieved, to the output, the evidence an auditor or a data subject can be shown.

LEGAL DUTY	ARCHITECTURAL CONTROL
Data protection by design (Art. 25)	Access-controlled retrieval, PII redaction, sovereign hosting as defaults [5]
Data minimisation (Art. 5)	Top-k retrieval and context budgeting, no bulk history in the prompt
Right to erasure (Art. 17)	Delete from store / memory; keep personal data out of model weights [4]
Security of processing (Art. 32)	Confidential computing, encryption in use; encryption at rest & in transit [9]
Transparency (AI Act Art. 50)	AI-interaction disclosure; synthetic-content marking from 2 Aug 2026 [1]
Accountability (Art. 5(2))	Exportable, source-attributed audit trail from query to answer
Impact assessment (Art. 35)	DPIA wired into the deployment process, not appended after

07, A REFERENCE ARCHITECTURE

Sovereign by design, as LinkTec Labs builds it

The controls compose into one stack with an explicit governance plane running beneath it. This is the architecture that unifies the series: the context layer of Note N°01 and the memory layer of Note N°02, hosted and governed so that sovereignty and compliance are structural, not aspirational.

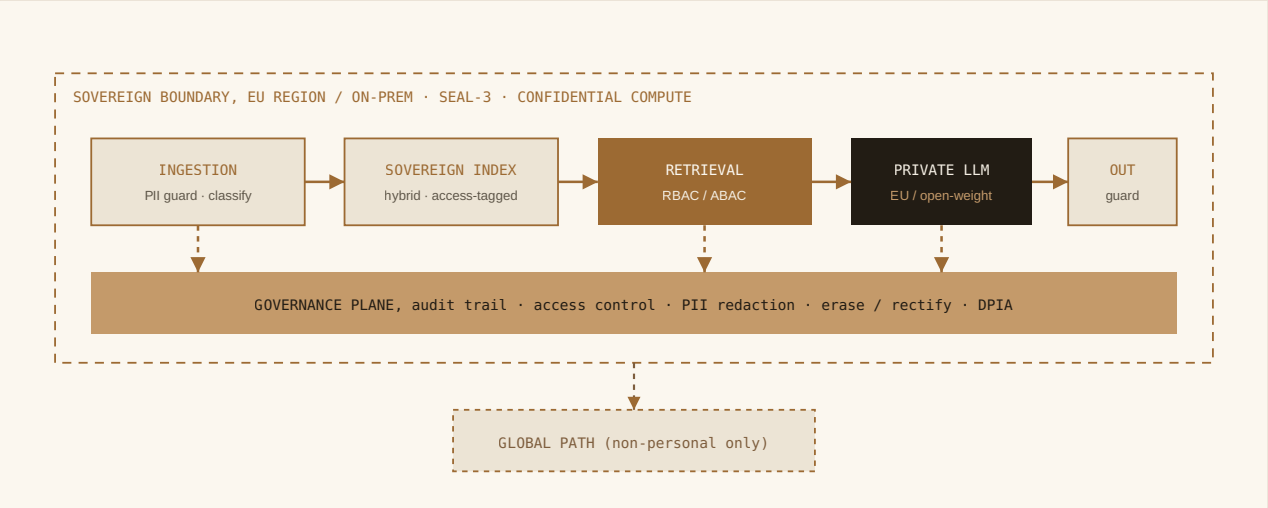


Fig. 4. LinkTec Labs sovereign reference architecture. Everything touching personal data stays inside a sovereign boundary (EU/on-prem, SEAL-3, confidential compute); a governance plane makes every record auditable, rectifiable and erasable. Non-personal work may exit to a global path.

Four principles govern it. **The boundary is explicit**, personal and regulated data never leave a sovereign, confidential perimeter, and that perimeter is a design artefact, not a hope. **Retrieval inherits authorisation**, the store enforces who may see what, so the model cannot launder access. **Personal data stays out of the weights**, it lives in a governed store where erasure is real, never trained in where it is not. And **governance is a plane, not a patch**, audit, rectification, erasure and DPIA run underneath every component, producing the evidence that turns "we are compliant" into something demonstrable. The first three earn trust; the fourth proves it.

08, PRACTITIONER'S PLAYBOOK

What to do on Monday

- → **Separate residency from sovereignty.** An EU region is not immunity from the CLOUD Act, check the operator's jurisdiction (SEAL-2 vs SEAL-3)[6].
- → **Diarise 2 August 2026.** Article 50 transparency is the binding near-term duty; the high-risk dates are 2027-2028[1][2].
- → **Keep personal data out of the weights.** Retrieve it; don't train on what you may have to erase[4].
- → **Make retrieval access-controlled.** The store must enforce per-document permissions, not the prompt[5].
- → **Guard PII in and out.** Detect and redact at ingestion and egress; minimise context to top-k.
- → **Protect data in use.** Confidential computing / TEEs, not just encryption at rest and in transit[9].
- → **Route by sensitivity.** Sovereign path for personal data; global capability only where no PII is involved.

→ **Build the audit trail first.** Prompt → sources → output, exportable, it is your DPIA evidence and your erasure proof.

Compliance you can demonstrate is an architecture. Sovereignty you can prove is a boundary. Neither can be added afterwards, and that is precisely why they are an advantage.

LinkTec Labs designs sovereign, audit-ready AI systems for organisations that cannot ship their data abroad, built on context architecture (N°01), governed memory (N°02), and the sovereignty-by-design principles above. This note closes the foundational arc of our 2026 research series. **linktec.fr / labs**

REFERENCES

Regulatory, technical & primary sources

[1] European Commission. **Regulatory framework on Artificial Intelligence (EU AI Act).** Shaping Europe's Digital Future, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

[2] Gibson Dunn. **EU AI Act Omnibus Agreement, Postponed High-Risk Deadlines and Other Key Changes.** 2026 (analysis of the Digital Omnibus on AI, provisional agreement 7 May 2026), https://www.gibsondunn.com/eu-ai-act-omnibus-agreement-postponed-high-risk-deadlines-and-other-key-changes/

[3] EU AI Office / European Commission. **General-Purpose AI Code of Practice** (published 10 July 2025), https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice

[4] European Data Protection Board. **Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models.** 17 Dec 2024, https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf

[5] European Union. **Regulation (EU) 2016/679, General Data Protection Regulation (GDPR)** (Arts. 5, 17, 25, 32, 35), https://eur-lex.europa.eu/eli/reg/2016/679/oj

[6] European Commission. **Commission advances cloud sovereignty through strategic procurement** (€180M sovereign-cloud framework; SEAL-2 / SEAL-3). 17 Apr 2026, https://commission.europa.eu/news-and-media/news/commission-advances-cloud-sovereignty-through-strategic-procurement-2026-04-17_en

[7] ENISA. **European Cybersecurity Certification Scheme for Cloud Services (EUCS).**, https://www.enisa.europa.eu/topics/cloud-and-big-data/european-cloud-services-scheme

[8] Gaia-X European Association for Data and Cloud. **Framework, labels & criteria.**, https://gaia-x.eu/

[9] NVIDIA. **Confidential Computing** (trusted execution environments for data-centre GPUs; encryption in use)., https://www.nvidia.com/en-us/data-center/solutions/confidential-computing/

[10] Anthropic. **Effective Context Engineering for AI Agents** (governance, citations & auditability primitives). 2025, https://www.anthropic.com/engineering/effective-context-engineering-for-ai-agents

[11] Lewis et al. **Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.** NeurIPS 2020. arXiv: 2005.11401, https://arxiv.org/abs/2005.11401

© 2026 LinkTec, LinkTec Labs. Research Note N°03. Figures are original schematics by LinkTec Labs. Regulatory dates reflect the Digital Omnibus position as of June 2026 and are provisional pending Official Journal publication; nothing here is legal advice. Quantitative and legal claims are attributed to their primary sources; secondary analyses are indicated as such in-text.