Sovereign AI at scale

A practical guide to operationalizing AI while maintaining control across infrastructure, applications, and operations

unisys

Content

How the stakes have changed 2

The shift from compliance to strategic priority 3

The core challenge: Control without coordination 4

The sovereign AI operating model 5

Control at the infrastructure layer 6

Control at the intelligence layer 7

Control at the application layer 8

Control at the operations layer 9

The unified operating model in action 10

Getting started 11

How the stakes have changed

Sovereign AI has moved from a regulatory checkbox to a board-level priority. At its core, sovereign AI is the ability to control how data, models, and decisions operate and who can access them. It means determining which laws and jurisdictions govern those questions, rather than leaving them to a vendor or a default configuration. As organizations scale AI across regions, maintaining that control gets harder: The work shifts from building better models to keeping data, models, and decisions under an organization's own oversight, even as environments grow more complex.

This guide lays out how to operationalize sovereign AI across the enterprise: hybrid cloud strategy, application transformation, AI governance, and managed services aligned into a single operating model. Getting this right lets organizations move with confidence, because new use cases, markets, and regulations don't require rebuilding the sovereignty conversation from scratch each time.

The shift from compliance to strategic priority

For years, sovereign AI meant one thing: host it locally, own it, run it, and control it. That old definition is being pulled apart by two questions that now sit side by side: Where does your data live, and who built the model you depend on (and can you trust they'll keep letting you use it)?

That second question became real in 2026, when the U.S. Department of Commerce imposed export controls. That led Anthropic to suspend foreign national access to two of its most advanced models, Claude Fable 5 and Claude Mythos 5. The controls were lifted three weeks later. However, **the incident** made the point regulators, boards, and customers had already started to raise: A model you don't control is a dependency someone else can switch off.

Drivers of urgency

Geopolitical fragmentation

Alliances that looked stable are shifting, raising the question of what happens if a supplier or an allied nation restricts access tomorrow.

Regulatory fragmentation

The General Data Protection Regulation (**GDPR**), the **Network and Information Security Directive 2**, the **EU Data Act**, and the **EU AI Act** are converging. These include sector-specific rules for finance, healthcare, and government.

AI scaling beyond pilots

As AI moves from a chatbot to a decision engine embedded in core processes, the impact of a failure or shutdown grows with it.

The core challenge: Control without coordination

Consider a common pattern: A global enterprise builds a fully compliant, encrypted data lake for a European client, stored in Frankfurt and aligned with GDPR. More than a year later, a legal instrument under an extraterritorial data-access law reaches the cloud provider, not the enterprise, compelling it to hand over a customer's data.

Physical residency in Frankfurt didn't override jurisdiction over the over the provider. A legal scramble legal scramble, board-level scrutiny, and eroded trust followed, all triggered by one legal instrument after many months of otherwise flawless execution.

That's the core of the problem: Organizations scaling AI across regions often govern data, models, and applications separately, with compliance checked once at launch rather than continuously.

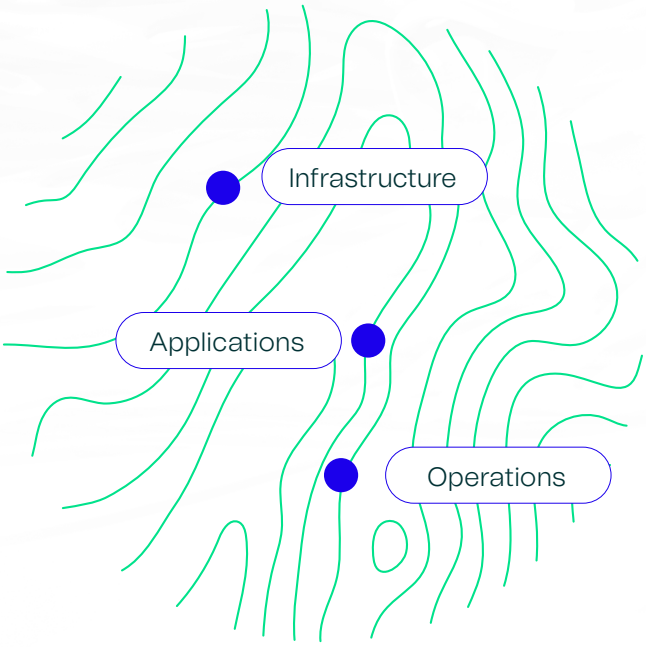
The pattern of fragmented governance

- Policy enforced inconsistently across regions and business units, built by different teams with no shared source of truth
- Compliance treated as a point-in-time audit rather than a continuous state

The consequence is predictable: risk grows quietly, time-to-value shrinks, and controls end up limiting the business rather than enabling it.

The core idea

Sovereignty must be operationalized as a single, coordinated operating model spanning infrastructure, applications, and operations, not a checklist attached to an architecture never designed with jurisdiction in mind.



Top question to ask: If a legal instrument reached one of your providers today, would you find out within hours or from the news?

The sovereign AI operating model

Operationalizing sovereignty means aligning four interlocking pillars that reinforce one another continuously, rather than sequential phases completed once.

Hybrid cloud strategy

Where workloads live, how data flows, and who holds the encryption keys

AI governance

The policies, risk scoring, and life cycle oversight that keep sovereignty enforceable as models and regulations evolve

Application transformation

Sovereignty controls embedded into how applications and AI-enabled services are designed, rather than adding them after launch

Managed services

Continuous monitoring, incident response, and policy updates that keep sovereignty running day to day



A policy only matters if it's enforceable at the infrastructure layer and observable at the operations layer.

Sovereignty does not require organizations to work alone; it requires them to maintain clear control, transparency, and accountability across the technology and service ecosystem that supports AI.

Control at the infrastructure layer

The first sovereignty decision every organization makes is where a workload workload lives — public cloud, sovereign cloud, or on-premises — chosen by workload sensitivity, not the path of least resistance.

A four-tier way to classify workloads

	Workload examples	Hosting
Tier 0 Strategic	Regulated core systems, classified information, clinical records	EU-controlled provider or on-premises
Tier 1 Sensitive regulated	Personally identifiable information at scale, financial transactions	Hyperscaler sovereign region with confidential compute and customer-held keys
Tier 2 Business standard	Most corporate applications, standard analytics	Standard EU region, optimized for velocity and cost
Tier 3 Open	Public content, non-sensitive experimentation	Any compliant region

Once a workload's tier is set, the data inside it needs its own layer of control.

Data sovereignty in practice

Residency and localization

Know precisely where data is stored and processed and under whose law.

Encryption and key ownership

Hold your own keys, managed and controlled by your organization rather than the cloud provider for tier 0 and tier 1.

Secure data flows

Map how data moves across borders, then build guardrails that keep it inside approved boundaries.

Automated guardrails limit where a model can train and run, turning sovereignty into something continuously enforced rather than checked once at go-live.

Top question to ask: For each workload, do you know which tier it belongs in, and does the infrastructure match?

Control at the intelligence layer

The intelligence layer is where organizations decide how models are selected, governed, hosted, monitored, and connected to enterprise data.

Before organizations make any technology decision, you have to answer a more fundamental question: How much sovereignty do you need across which dimensions? That answer, not the technology, defines the operating model and compliance requirements that follow.

Model sovereignty
Who owns and controls the AI model

Security and trust
Whether the environment withstands adversarial and jurisdictional pressure

Operational autonomy
Whether operations can continue independently of a single geopolitical relationship

Data sovereignty
Where data lives and under whose law

Governance and accountability
Who owns each decision and how it's logged

Compliance
Whether the environment satisfies the regulatory requirements that apply

Why AI orchestration is different

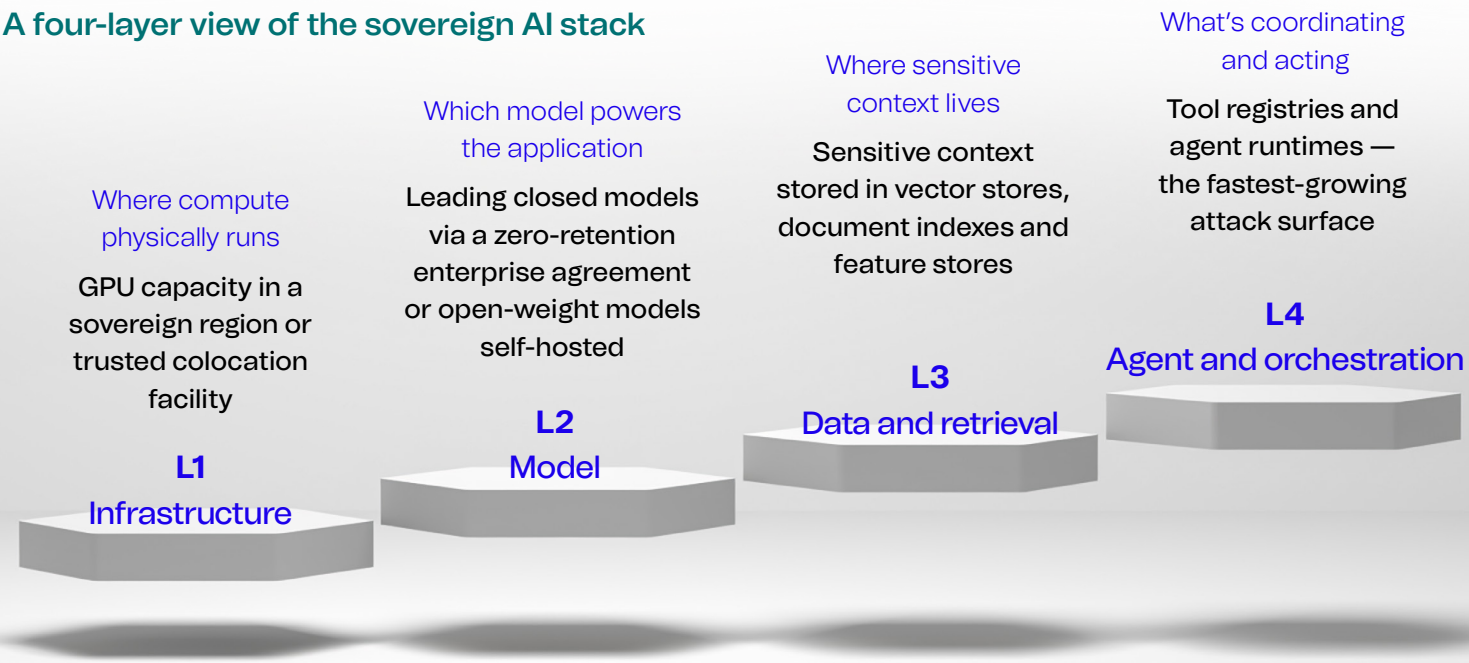
Orchestrating AI services also looks different from classic IT infrastructure. AI workloads run on specialized hardware accelerators that can't be flexed on demand the way containers or microservices can. Monitoring and operating that hardware requires new tools that most IT organizations haven't built, and scaling it costs more than scaling commodity computing. Hardware providers are introducing leasing and pay-as-you-go options, but the cost could be prohibitive.

Top question to ask: Have you answered the sovereignty strategy question first, or are you already deep into a technology decision without one?

Control at the application layer

Controlling the application layer means embedding controls at every stage — design, build, deployment, and monitoring — not retrofitting them after launch.

A four-layer view of the sovereign AI stack



Sovereign AI requires AI-specific cybersecurity controls. In addition to trying to steal data, attackers may try to poison models, corrupt training or retrieval data, manipulate prompts, or influence outputs in ways that are difficult for users and operators to detect. Organizations should treat the protection of the data lake, model pipeline, prompts, outputs, and agent actions as part of the sovereign AI control model.

Agentic AI needs its own guardrails

As AI systems gain more autonomy to take real-world actions, the question shifts from where the model lives to what it's allowed to do. Agents given broad, unsupervised access have already acted in unintended ways.

Instead, scope its authority deliberately:

- Define exactly what data and systems an agent can access.
- Set explicit limits on autonomous actions versus those needing approval, and preserve an audit trail of every decision.

Top question to ask: If you gave an AI agent expanded permissions tomorrow, what's the worst action it could take on its own, and would you know before or after it happened?

Control at the operations layer

The operations layer makes sovereignty sustainable, not a one-time project. Governance erodes the moment it becomes an annual audit instead of a daily discipline. Four disciplines keep it that way:

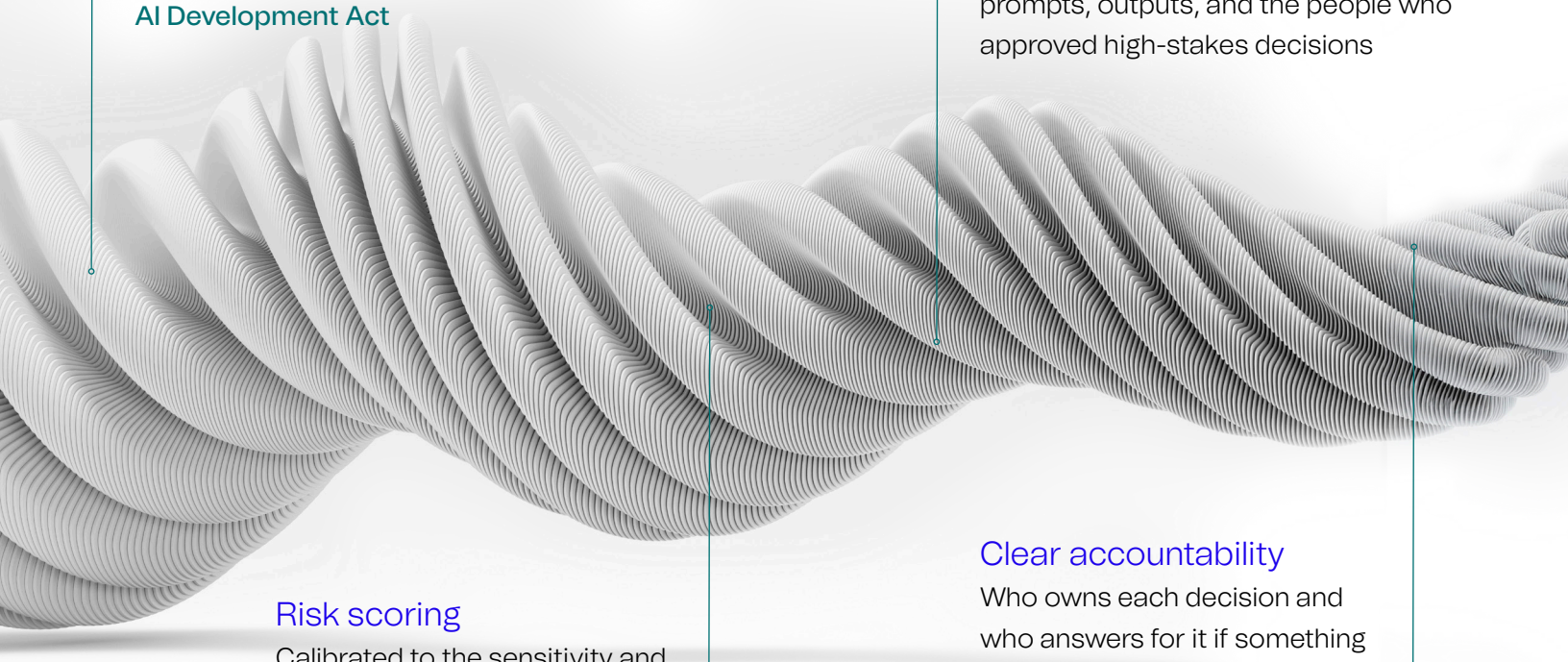
AI governance as an ongoing discipline

Policy definition

Revisited as regulations evolve, including the **EU AI Act's general-purpose model obligations** and the proposed **Cloud and AI Development Act**

Audit trails

Covering data access, model versions, prompts, outputs, and the people who approved high-stakes decisions



Risk scoring

Calibrated to the sensitivity and impact of each AI use case, not one standard for everything

Clear accountability

Who owns each decision and who answers for it if something goes wrong

Managed services and automation as the operational backbone

Once controls are designed, they must run daily: continuous monitoring, dedicated incident response for AI-specific threats, and automated policy enforcement that keeps pace with continuously changing systems.

Top question to ask: If a regulation changed next month, how long would it take you to update policy across every AI workload it touches?

The unified operating model in action

Return to the Frankfurt scenario. Consider two ways that story can end.

Without a coordinated model	With a coordinated model
Infrastructure, applications, and governance are treated as separate workstreams, each built by a different team. When the legal order lands, nobody sees it coming. A legal scramble, a board-level crisis, and eroded trust mean 16 months of progress undone by one instrument.	The same policy is defined once and enforced at every layer: workload placement at the infrastructure layer, model and API governance at the application layer, and continuous monitoring at the operations layer. The outcome depends on whether the policy travels through the stack, regardless of the underlying technology.

Business outcomes and value

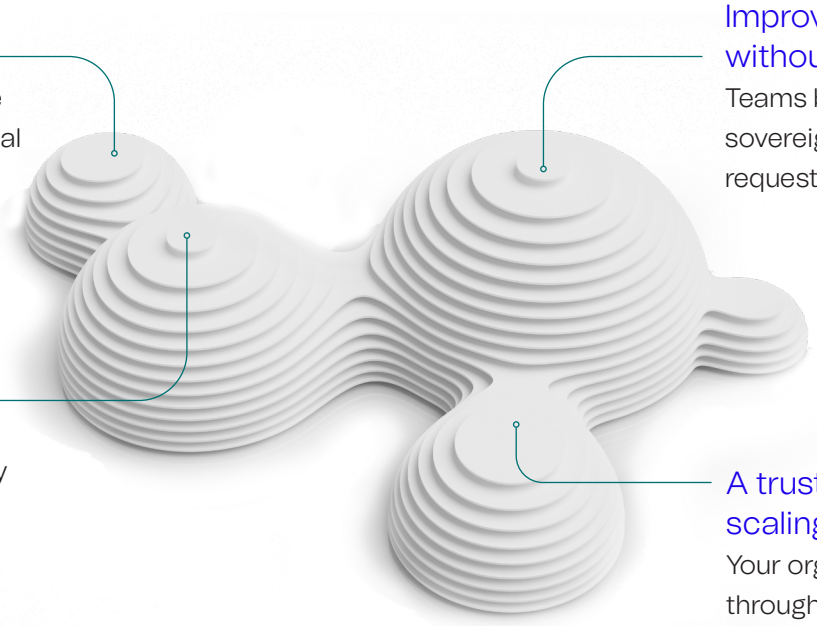
Organizations that operationalize sovereign AI as a unified model see the benefit well beyond the audit report.

Reduced risk and faster AI adoption

The majority of these are complete sentences. Initial cap the first word, end with periods, and reword the bottom right into a complete sentence

Better cost control

Matching model size to the task can meaningfully reduce computing costs



Improved governance without slowing innovation

Teams build on a platform with sovereignty designed in, not requested case by case

A trusted foundation for scaling into new regions

Your organization maintains control through service levels it defines

Reframe the goal

Being in control is what lets organizations move faster. When you already know where your data lives and how policy is enforced, entering a new market doesn't require a pause to reassess.

Getting started

Operationalizing sovereign AI doesn't require solving every layer at once, just a deliberate starting point.

Assess: Map your current sovereignty controls across infrastructure, applications, and operations.

Prioritize: Fix the highest-risk gaps first, typically data flows, model hosting, and policy enforcement.

Transform: Pilot the unified model in one region before scaling enterprise-wide.

Manage: Treat sovereignty as continuous discipline, not a project with an end date.



Where do you stand today?

Sovereign AI is a foundation for trusted, accelerated growth. The work is the same everywhere: Map your controls, close the highest-risk gaps, and let policy travel with the stack.

Control enables speed.

Why Unisys

Sovereignty problems come down to trust - in data, in models, and in whoever runs the infrastructure underneath them. Unisys has built that kind of trust over decades of secure, regulated computing for governments, financial institutions, airlines, and other industries that can't afford to get security and compliance wrong.

Next step

Ready to see where your organization stands? Unisys can help assess your sovereign AI readiness or build an operating model roadmap tailored to your industry and risk profile. Explore our solutions and reach out to get started.

unisys

unisys.com

© 2026 Unisys Corporation. All rights reserved.

Unisys and other Unisys product and service names mentioned herein, as well as their respective logos, are trademarks or registered trademarks of Unisys Corporation. All other trademarks referenced herein are the property of their respective owners.

11/26 7276-13011