deepset

Building Sovereign AI

A Complete Guide for Designing AI Systems for Control, Compliance, and Change in 2026 and Beyond

Table of Contents

The Shift Toward Sovereign AI

The rapid adoption of Generative AI has moved from a "competitive advantage" to a "strategic risk" discussion. Organizations are no longer asking if they can use AI, but whether they can afford to lose control of the data and logic powering it. The market has reached a tipping point: Accenture research (2025) reveals that **61% of business and government leaders** are now actively seeking sovereign technology solutions as geopolitical risks and regional compliance requirements rise. Sovereignty is no longer a fringe policy; it is the new baseline for digital competitiveness.

Key Drivers

AI Regulation Surge

Gartner predicts that by **2028, 65% of governments worldwide** will introduce formal technological sovereignty requirements. In the EU, the **AI Act** (fully enforceable by August 2026) imposes penalties of up to €15 million or 3% of global turnover for violations in high-risk sectors like banking, healthcare, and public services.

Cloud Control & Data Access Risk

Dependence on a few global providers creates a single point of failure. Enterprises are waking up to the reality that even if data is stored in a "local" cloud region, the parent company may still be subject to the **US CLOUD Act**, which grants foreign authorities legal pathways to access stored information.

Higher Costs of Closed Models

While closed APIs offer speed, they come at a "sovereignty tax." Research from MIT Sloan (2025) shows that closed models cost users **six times as much** as open alternatives. Organizations that migrate high-volume workloads to self-hosted, sovereign models can achieve **50–80% reductions in inference costs** while maintaining 90% performance parity - a gap that is closing within weeks of every new model release.

Uncontrolled AI Agents

As systems evolve toward **Agentic AI,** capable of autonomous planning and execution, the stakes for control skyrocket. Gartner forecasts that by 2029, **60% of government** agencies will use AI agents for citizen-facing transactional interactions. Without a sovereign orchestration layer, these agents become "risk amplifiers" that can inadvertently leak IP or violate privacy laws in ways a simple chatbot never could.

Introduction: Understanding Sovereign AI

AI is rapidly becoming a core component of modern digital infrastructure. Enterprises are embedding AI into customer experiences, internal knowledge systems, and operational workflows. Governments are adopting AI to modernize public services, automate administrative processes, and support policy decisions.

As AI becomes more deeply integrated into critical systems, a new question has emerged: how do organisations control the intelligence powering these systems?

Today's AI ecosystem is highly concentrated. A small number of global technology providers dominate the development of advanced models and AI platforms. While these technologies provide unprecedented capabilities, they also introduce new dependencies around data governance, regulatory compliance, and long-term technological control.

At the same time, organizations are increasingly relying on AI to interact with sensitive information such as proprietary business knowledge, customer records, operational data, or citizen services. In these contexts, maintaining visibility and control over how AI systems operate becomes essential.

These shifts are forcing organizations to confront a fundamental reality: AI is no longer just a tool, it is infrastructure that must be governed. This has led to the emergence of **Sovereign AI**.

Sovereign AI refers to the ability of an organization or a nation to **design, deploy, and operate AI systems under its own control.** This includes maintaining authority over how data is accessed, how models are used, how systems are deployed, and how AI behavior is monitored and governed over time.

Importantly, sovereign AI does not require organizations to build every component of the AI stack themselves.

Common Misconceptions around AI Sovereignty

Misconception	Reality - The Sovereign Approach
"We must build our own LLM."	Teams should be able to swap our LLM without breaking our workflows.
"Sovereignty means isolation."	Sovereignty means strategic interdependence, choosing where data is processed.
"It's only a government issue."	Enterprises face identical risks of vendor lock-in and regulatory fines.

Why Sovereign AI is Hard to Build

Building sovereign AI is at its core, a challenge of systems design. Fragmented AI stacks can no longer reconcile strict data residency and sovereign AI requirements with the high-speed performance required for production.

Operational & Scaling Challenges: The "Velocity Paradox"

The "Distributed Debt": Architectural Complexity

Compliance requires operating across clouds, on-prem, and regions.

The Challenge: AI systems must stay portable and interoperable across this fragmented landscape.

The Struggle: Without the right tooling, teams face model lock-in, data residency risks, and growing operational complexity.

The "Glue Code" Trap: Fragmented Tooling

The current AI ecosystem is a collection of brilliant but isolated parts (vector DBs, monitoring agents).

The Challenge: Integrating these disparate components into a cohesive way to respect compliance constraints.

The Struggle: 80% of teams' time writing "brittle glue code" where components don't share context, and governance cannot be enforced consistently.

The Moving Target: Regulatory "Feature Creep"

Regulation is not a one-time event; it is a shifting landscape of overlapping frameworks like the EU AI Act & GDPR.

The Challenge: Building a system that is flexible enough to adapt to new technical standards without a total re-architecture.

The Struggle: The real difficulty is the constant engineering refactoring required to stay compliant as "high-risk" definitions and audit requirements evolve.

The "Agent Sprawl" Risk: Safeguarding Autonomous Control

Agentic systems expand the surface area for risk. Agents invoke dynamic decision-making and autonomous tool use.

The Challenge: Containing an agent so it doesn't take actions or access data that violates regional policy.

The Struggle: Without a **Central Control Plane**, agents amplify risk, lacking real-time guardrails to prevent crossing sovereign boundaries mid-task

The Regulatory Landscape: The "Cold Start" Governance Problem

Sovereignty is no longer a strategic choice; it is a legal mandate. For any AI system, regardless of use case, compliance has shifted from a policy document to a real-time engineering requirement. If your architecture cannot programmatically "prove" its behaviour, it is non-compliant by default.

1. The "Forensic Traceability" Mandate (EU AI Act & ISO 42001)

Regulators now require "logging of the system's functioning throughout its lifecycle."

- The Requirement: For any output, e.g. a medical summary, a legal draft, or a customer response, you must be able to produce a timestamped audit trail showing the exact model version, the specific data chunks retrieved, and the system prompts used at that moment.
- The Building Challenge: When a regulator asks, "Why did the AI say this on Oct 12th?", teams struggle to reconstruct the state of a dynamic RAG pipeline from weeks prior.

2. "Data Provenance" & Right to Explanation (GDPR / EU Digital Omnibus)

Sovereignty isn't just about where data sits; it's about where it came from.

- The Requirement: Under the 2026 "Secure by Design" baseline, if a user challenges an AI-generated decision, a "meaningful explanation" of the logic involved must be provided.
- The Building Challenge: In a fragmented stack, the "Chain of Custody" is often broken. Developers find it nearly impossible to link a specific LLM output back to the original, sovereign data source (the "Grounding Source") in a way that is legally defensible.

3. "Boundary Enforcement" (NIST AI RMF 2.0)

AI systems must strict data boundaries (e.g., keeping PII or CUI to a specific region).

- The Requirement: You must demonstrate "Technical Guardrails" that prevent a model from "bleeding" data across unauthorized boundaries, even if the user asks it to.
- The Building Challenge: This requires runtime governance. Intercepting and checking every data retrieval and model response adds significant latency and complexity.

4. The "Human-in-the-Loop" (HITL) Requirement for High-Risk Triage

- The Requirement: Under the EU AI Act, systems categorized as "high-risk" (HR, Finance, Critical Infrastructure) cannot be fully autonomous without a human override function.
- The Building Challenge: Triggering timely and accurate manual review workflows is a massive development hurdle that most "prototypes" never solve.

The Four Pillars of Sovereign AI

To build sovereign AI systems, organizations must maintain control across four key dimensions of the AI lifecycle.

Pillar		Focus
Data Control	→	Governing how enterprise data is accessed and used in AI systems
AI System Architecture	→	Designing and deploying modular AI systems to be portable and adaptable
Model Sovereignty	→	Enabling flexible model choice, portability, and independence from single vendors
Operations & Governance	→	Monitoring, evaluating, and managing AI systems over time

Pillar 1 — Data Control

Data is the most important asset in most AI systems. For enterprises and public sector organizations, internal knowledge - documents, records, databases, and proprietary information - is often the primary source of value in AI applications. Maintaining control over how this data is accessed and used is therefore essential.

AI systems must interact with enterprise data in ways that ensure:

- Data stored and processed **within trusted jurisdictions** to meet compliance requirements
- Sensitive information remains protected
- Access permissions are respected
- Outputs can be traced back to source information

Pillar 2 — AI System Architecture

Performant, production-grade sovereign AI application share architectural characteristics.

1. Modularity

Modern AI applications are not single models responding to prompts. They are complex systems composed of multiple interacting components including:

- Enterprise data sources
- AI Models
- External tools, memory, API
- Monitoring & Evaluation

The way these components are connected determines how teams maintain control over their AI systems, a modular approach enables replacement without redesigning the entire system.

2. Deployment flexibility

AI systems can run across different environments depending on organizational requirements. Common deployment models include:

- **Cloud,** where systems run on externally managed infrastructure and services
- **On-premise deployments,** where systems run entirely within internal infrastructure
- **Private cloud or VPC environments** - run in isolated cloud infrastructure with greater control and security
- **Hybrid architectures** - internal systems integrate with external models or services

Pillar 3 - Model & Technology Sovereignty

Sovereign AI systems depend on the ability to choose, adapt and replace the models that power them, ensuring organizations retain control, flexibility and independence over their AI infrastructure. In today's AI ecosystem, teams have access to a range of model options:

- Open Source Foundation Models
- Commercially Hosted Models
- Domain-Specific Models
- Internally Fine-Tuned Models

Each option offers different advantages for performance, cost, transparency, and control. Heavy reliance on a single model provider can create strategic limitations such as dependence on a vendor's pricing model, API availability, or update schedule.

AI systems should allow models to be **swapped or upgraded without requiring fundamental architectural changes**. Different tasks may benefit from different models. One model may excel at summarization while another performs better for reasoning or classification tasks. Operating truly sovereign AI systems requires **control at the model layer.**

Pillar 4 — Operational Governance

Building an AI system is only the first step. Maintaining control over that system as it operates and evolves is equally important. Operations and governance focus on how AI systems are **managed, monitored, and improved over time.** Several capabilities are critical here:

Observability

Teams must be able to monitor how AI systems behave in production, including model outputs, system performance, and usage patterns.

Human-in-the-Loop (HITL):

In high-stakes environments (e.g., issuing a government permit), the AI should not act autonomously. Sovereign systems bake HITL checkpoints into the workflow.

Evaluation

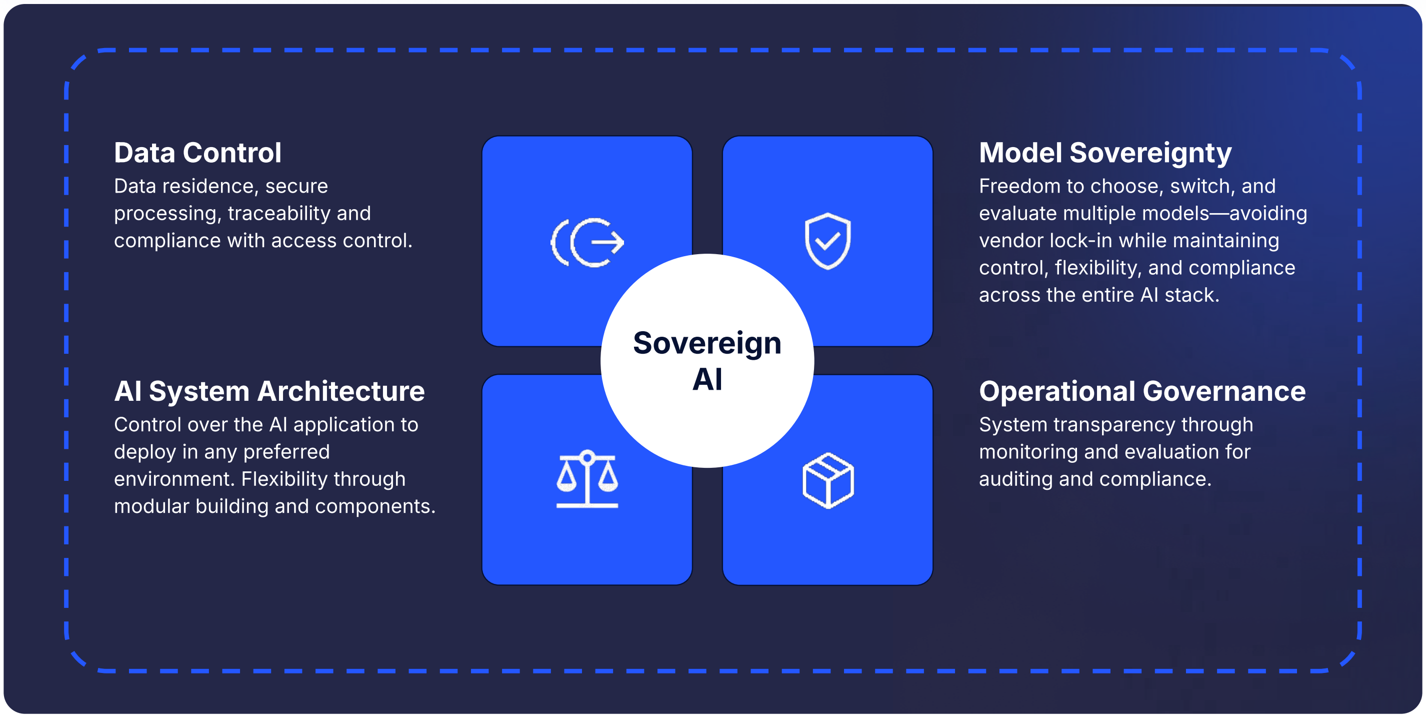
AI systems must be tested regularly to ensure they meet reliability, safety, and performance standards.

Governance

Organizations must enforce policies that ensure AI systems operate in accordance with regulatory requirements and internal standards.

Lifecycle management

Managing updates, versioning, and changes to models and pipelines in a controlled, auditable way.



Data Control
Data residence, secure processing, traceability and compliance with access control.

Model Sovereignty
Freedom to choose, switch, and evaluate multiple models—avoiding vendor lock-in while maintaining control, flexibility, and compliance across the entire AI stack.

AI System Architecture
Control over the AI application to deploy in any preferred environment. Flexibility through modular building and components.

Operational Governance
System transparency through monitoring and evaluation for auditing and compliance.

AI Orchestration as the Foundation for Achieving and Scaling Sovereign AI

Gartner Predicts 2026: AI Sovereignty lists the following recommendations for approaching the design and build of Sovereign AI systems:

- Design model agnostic workflows using orchestration layers that enable switching between LLMs across regions. Use abstraction, routing, and standardised prompt templates to reduce vendor dependence and adapt to local compliance.
- Ensure your AI governance, data residency, and model tuning practices can meet country-specific legal, cultural, and linguistic requirements.
- Monitor AI legislation, data sovereignty rules, and emerging standards that may affect where and how you can deploy AI models and process users data.

Organizations must maintain control across key dimensions:

Pillar		Core Question
Data Control	→	How to control and govern data as AI systems scale?
AI System Architecture	→	How are AI systems deployed and integrated?
Model & Technology Sovereignty	→	How can we ensure AI systems aren’t tied to a specific model?
Operational Governance	→	How do you maintain AI governance as regulations and requirements evolve?

However, these pillars cannot exist independently. Sovereign AI systems requires a mechanism for coordinating how data, models, and workflows interact.

That mechanism is **AI orchestration** - this is the essential control layer that enables teams to:

- connect AI systems to enterprise data securely
- switch between models without redesigning applications
- manage system workflows and tool integrations
- enforce monitoring, governance, and compliance policies
- architect for deployment optionality

Without orchestration, organizations often lose control over how AI systems operate. With orchestration, they gain a unified architecture that enables sovereignty across the entire AI lifecycle.

Enabling Sovereign AI with the Haystack Enterprise Platform

Building sovereign AI systems requires more than powerful models—it requires **control, modularity, and transparency across the entire AI stack.** Organizations must be able to govern how data enters AI pipelines, control system behavior, deploy systems in secure environments, and demonstrate compliance with regulatory requirements.

The Haystack Enterprise Platform provides the **orchestration layer that bridges the gap between AI experimentation and sovereign, production-grade deployments.** Through modular pipelines, governed data access, and flexible deployment options organizations can build AI systems that remain adaptable, secure, and under their control.

Haystack's **component-based orchestration architecture** enables organizations to design custom AI workflows while remaining independent of specific model providers or infrastructure environments.

This orchestration layer acts as the **control plane for AI systems,** allowing teams to define:

- how enterprise data is retrieved and filtered
- which models are used for specific tasks
- how tools and APIs are invoked
- how outputs are evaluated and monitored

Enterprise Knowledge Infrastructure for Sovereign AI

In sovereign environments, knowledge ingestion and management must be governed with the same rigor as the AI systems that use it.

The Haystack Enterprise Platform supports this through:

- governed ingestion pipelines
- versioned knowledge indexes
- controlled context assembly for AI pipelines

By managing knowledge as independent infrastructure, decoupled from applications, organizations can prevent data drift, duplication, and inconsistency across systems.

This approach ensures that AI applications always operate on trusted and up-to-date knowledge sources.

Secure Enterprise Data Integration

For sovereign AI systems, controlling how data enters AI workflows is essential.

Haystack provides secure integration with enterprise data stores, enabling organizations to build retrieval-augmented systems that leverage internal knowledge while maintaining strict governance over data access.

Sophisticated RAG & context engineering workflows allow teams to:

- connect to internal document repositories and databases
- enforce role-based access control over retrieved content
- ground model responses in verifiable enterprise knowledge

This architecture ensures that AI systems can leverage proprietary data while preserving data ownership, security, and traceability.

Flexible Deployment Across Sovereign Environments

Sovereign AI systems often need to run in environments that meet strict security and regulatory requirements.

The Haystack Enterprise Platform supports flexible deployment models, including:

- on-premise infrastructure
- private cloud or VPC environments
- hybrid deployments combining internal and external systems

For organizations operating in highly regulated or classified environments, Haystack pipelines can be deployed within air-gapped or sovereign infrastructure, ensuring sensitive data never leaves controlled environments.

This deployment flexibility allows organizations to maintain control over data residency, network boundaries, and compliance requirements while still advancing their AI capabilities.

Production-Grade Governance & Operational Controls

Moving AI systems into production requires strong operational governance.

The Haystack Enterprise Platform extends the open-source framework with **production-grade lifecycle management and governance capabilities,** including:

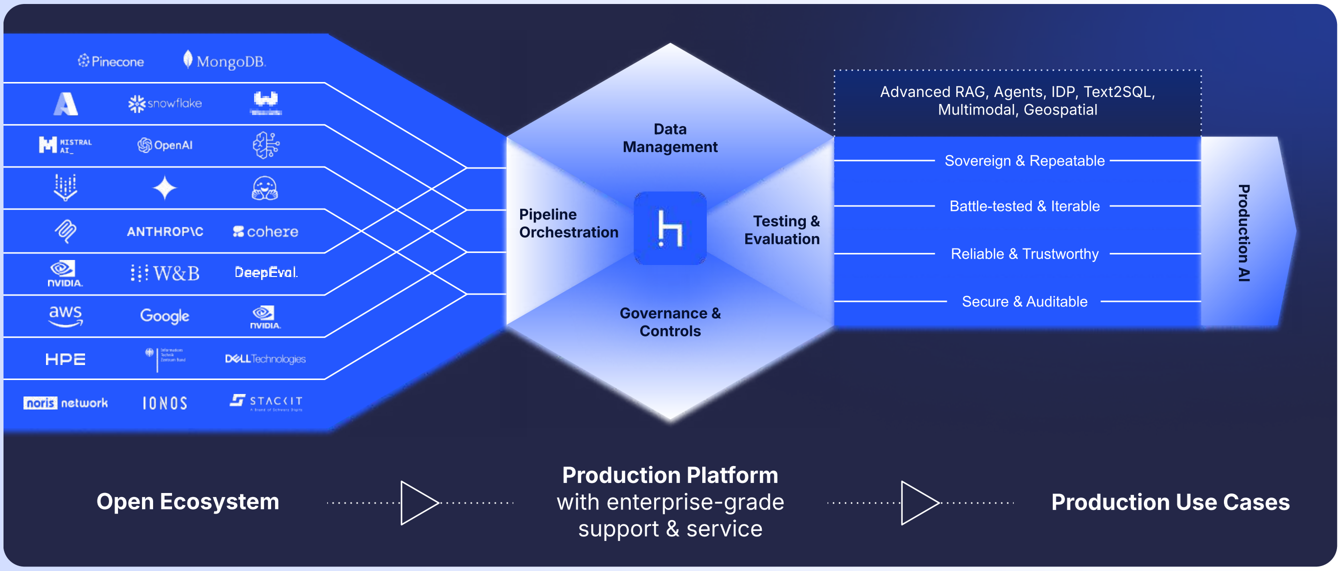
- pipeline versioning and CI/CD integration
- role-based access control for pipeline components
- runtime monitoring and observability
- evaluation frameworks for testing model performance
- audit trails across the AI development lifecycle

These capabilities provide the operational evidence required for regulatory compliance and internal governance, particularly in environments subject to frameworks such as the EU AI Act or sector-specific regulations.

Haystack Enterprise Platform for Sovereign AI

By combining **modular workflows, governed data access, flexible deployment, and production-grade governance,** Haystack enables organizations to build AI systems that remain secure, adaptable, and compliant as they scale.

Rather than locking teams into proprietary platforms, Haystack provides the open and extensible foundation required to design AI systems that organizations can **fully understand, control, and operate within sovereign environments.**



Sovereign AI in Practice

Haystack is used by organizations operating in highly regulated and sovereign environments, including government agencies, defense organizations and enterprises.

credX

Bringing control, efficiency and transparency to real estate AI

- **Goal:** Enable faster, more reliable analysis of complex real estate transactions—while maintaining full control, transparency, and trust over data and workflows.
- **Solution**: credX built a production-grade AI system using the Haystack Enterprise Platform, creating an application that processes thousands of pages of unstructured documents, extracts 200+ key data points with ~99.5% accuracy, and delivers fully traceable, source-backed insights.
- **Impact**: credX can deploy AI on their own terms, integrating domain expertise, enabling audit processes while reducing analysis time by up to 80% and enabling analysts to focus on higher-value decisions.

Beauftragt durch:

Bundesministerium
für Forschung, Technologie
und Raumfahrt

Modernizing public sector funding applications and approvals

- **Goal:** The German Federal Ministry of Research, Technology, and Space (BMFTR) manages several million euros in annual project funding, adhering to complex regulations contained in a 200-page Project Funding Manual.
- **Solution:** AI assistant that helps application analysts answer questions about project funding rules in natural language, meeting high data protection and usability requirements, complying with all aspects of C5 cloud security criteria, and providing responses in under three seconds.
- **Impact:** The AI assistant has become an essential daily tool for staff, resolving 1000 queries each week without additional staff resources, reducing issue resolution time, and enhancing transparency via verifiable source citations and references.

AIRBUS

Transforming real-time planning and decisioning

- **Goal:** Enhance mission-critical decision-making for the German Armed Forces through AI, ensuring military-grade reliability aligned to NATO standards.
- **Solution:** AI for Tactical Chat in Simulation Systems (KITCH), integrating agentic LLMs, reinforcement learning, and RAG for real-time tactical support while mitigating hallucinations. Through a satellite feed and chat interface teams are able to get actionable, context-aware guidance in evolving scenarios to make critical mission decisions for personnel safety.
- **Impact:** Supports multi-perspective reasoning and adaptive situational awareness while significantly reducing the risk of hallucination or misleading information.

Sovereign AI as a Strategic Capability

Sovereign AI is no longer just a matter of national policy, it is rapidly becoming the blueprint for the next generation of enterprise digital infrastructure.

As AI models move from experimentation into core business operations, organisations are facing a new reality: the systems that generate insight, automate decisions, and interact with critical data must remain **secure, governable, and highly controllable.**

At the same time, regulatory frameworks are evolving and geopolitical dynamics continue to shape how AI technologies are deployed and governed. The ability to **operate AI systems independently and responsibly** is emerging as a strategic capability. Organisations that succeed in this transition will treat sovereignty not as a constraint, but as a **design principle for AI architecture.**

The Foundations of Sovereign AI

Building sovereign AI systems requires control across several key domains.

Data Control

AI systems must integrate enterprise knowledge without compromising security or compliance. Sovereign data pipelines typically include:

- secure integration with enterprise data sources
- RBAC for retrieved content
- filtering and policy enforcement mechanisms
- traceability and citation of source information

System Architecture

Modern AI applications are complex systems composed of retrieval pipelines, models, tools, and applications.

AI orchestration provides the mechanism that keeps these systems modular and adaptable, separating workflow logic from specific models or infrastructure providers. This allows organizations to evolve their systems as technologies change without rebuilding applications from scratch.

Deployment Flexibility

Sovereign AI systems must often operate across a variety of infrastructure environments, including private clouds, VPCs, and secure on-premise deployments.

By abstracting AI workflows from underlying infrastructure, orchestration enables organizations to deploy the same systems wherever security, compliance, or operational requirements demand.

Operations & Governance

Finally, sovereign AI requires strong operational oversight, this enables teams to understand how systems behave in production:

- monitor model performance
- evaluate system reliability
- enforce governance policies
- maintain audit trails for compliance

This transparency is essential for maintaining trust in AI-driven decisions.

How AI Orchestration Enables Sovereign AI

To solve the complexity of sovereign AI, organizations must shift their focus from individual models to the **orchestration layer.**

AI orchestration acts as the **control plane for AI systems,** coordinating how information flows between enterprise data, models, and external tools.

It enables sovereignty in three critical ways:

- **Decoupling Logic from Models:** If an application's intelligence is embedded inside a proprietary model API, the system becomes dependent on that provider. When workflows are defined in the orchestration layer, models can be replaced or upgraded as technology evolves.
- **Governed Data Flow:** Orchestration determines exactly how data enters AI workflows: controlling what information is retrieved, filtered, and passed to models.
- **Traceability:** By defining the full pipeline, from data ingestion to final output, AI orchestration provides a unified audit trail for the entire AI lifecycle.

The Path Forward

Organisations that embrace **modular orchestration architectures** will be best positioned to innovate rapidly while maintaining control over their most valuable assets:

- their **data**
- their **system logic**
- their **future AI capabilities**

As AI becomes core infrastructure, orchestration makes it possible to build sovereign AI systems that are powerful, transparent, and **fully under organizational control.**

By separating system logic from infrastructure and model dependencies, Haystack allows organizations to maintain long-term architectural flexibility while avoiding platform lock-in. To build sovereign AI systems, organizations must maintain control across key dimensions.

READY TO BUILD SOVEREIGN AI SYSTEMS AT SCALE?

Book a Demo Explore Haystack