FEDERIS AI

PRODUCT BRIEF

OVERVIEW AND CAPABILITIES

Federis AI acts as a modular sovereign control plane, network controller, and security proxy that orchestrates and secures AI workloads in air-gapped, on-premises, hybrid, or sovereign cloud environments. It sits between applications and GPU-backed LLM serving layers to decouple compute/storage from central governance and audit requirements.

Value Proposition

Trusted, Composable, Sovereign AI Infrastructure

Complete control over data residency, security, and regulatory compliance in on-prem, hybrid, or edge deployments.

Zero-Trust Identity & Access Enforcement

Workloads verified via cryptographic identities (SVIDs) and user access governed via central authorization controls.

Dynamic Engine-Based Routing & Policies

Access controls (RBAC/ABAC) evaluated to routing lanes based on geofencing and classification.

Pre & Post-Guardrail Verification

PII redaction and prompt-injection defense before GPU dispatch, and toxic/unsafe content filtering post-inference.

Explainability & Hallucination Mitigation

A hallucination shield that validates claims against document sources and rejects predictions with low grounding.

Tamper-Evident Ledger Archiving

Decision traces chained sequentially using SHA-256 hashes, signed dynamically, and stored under WORM retention.

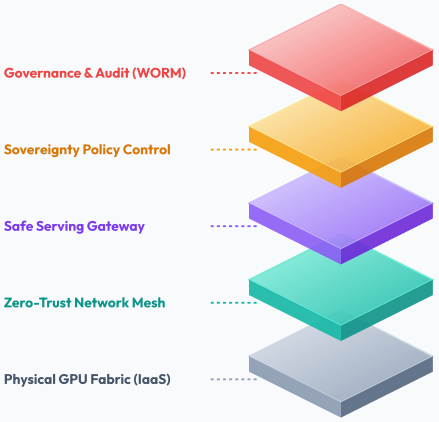


Figure 1: From Infrastructure to Sovereign Intelligence - Unified Control Stack

Product Features

- **Dynamic Sovereign Routing**
  Locks prompt dispatch to local GPU instances and RAG datasets dynamically, maintaining boundary-locked air-gapped isolation.
- **Pre-Guardrail Filtering**
  Intercepts and masks PII variables (e.g. Aadhaar, PAN card, personal identifiers) using automated regular expressions.
- **Decoupled Pluggable Adapters**
  Binds external databases, identity platforms, and vector DBs strictly via interfaces, keeping core code license-clean.
- **High-Performance Serving Gate**
  Provides queue-backed execution, model catalog reload validation, circuit-breaker routing, and autoscaling.
- **Explainable Hallucination Shield**
  Deconstructs output into clear statements, mapping each to specific local document chunks, blocking answers below validation thresholds.
- **Tamper-Evident Ledger Vault**
  Serializes each decision context into a SHA-256 hash-chained list, exporting safely to compliant Object-Lock storage systems.

© 2026 Federis AI Inc. All Rights Reserved.

FEDERIS AI

CAPABILITIES BRIEF

DETAILED CAPABILITIES

INFRASTRUCTURE & SERVING CONTROL

Autonomic Platform Engine
High-throughput orchestrator core with strict schema validation to coordinate context and classification checks.

Autoscaling & Queueing Services
Integrates work queues for workload buffering, and exports task-density metrics for auto-scaling background units.

Isolated Serving Runtimes
Supports high-performance serving runtimes, pulling weights from local read-only Persistent Volume Claims to secure air-gapped targets.

Zero-Trust Network Mesh
Establishes secure, encrypted mTLS communication cells using workload identity credentials and strict authorization rules.

CONTAINER & GITOPS SERVICES

GitOps-Driven Deployment
Declarative GitOps blueprints for fully automated installation on bare-metal and private Kubernetes.

Cluster Guardrails
Enforces runtime configuration limits and network access security boundaries using declarative safety policies.

AI PAAS SERVICES

Inference as a Service (Infra-aaS)
High-throughput gateway offering approved model admission controls, fallback paths, and circuit-breaker management.

Model Governance Registries
Model release gates, model lineage registration, and cryptographic artifact attestation.

Retrieval Services (RAG-aaS)
Binds to external vector databases, indexing target documents with regional metadata restrictions.

VALUE-ADDED & AUDIT SERVICES

Pluggable Adapter Model
Abstracts downstream databases, identity platforms, and trace sinks to ensure core code is decoupled from external licenses.

LLM Trace Observability
OTLP trace export pipeline feeding telemetry collectors and dashboards with defensive PII redaction.

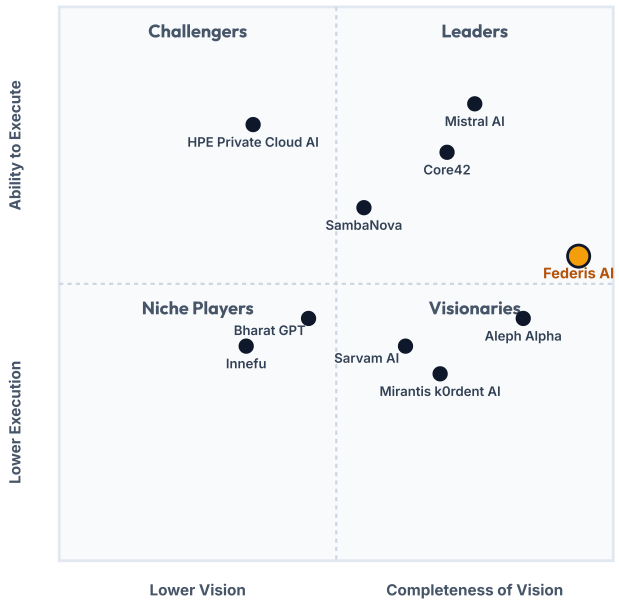
WORM Audit Archiver
Sequentially chained SHA-256 records with HSM signers, written to Object-Lock compliant object storage for compliance.

© 2026 Federis AI Inc. All Rights Reserved.

FEDERIS AI

MARKET MAPPING

SOVEREIGN AI MAGIC QUADRANT



Magic Quadrant Placement

Federis AI is positioned as the absolute leader in completeness of vision for sovereign deployments. Unlike legacy platforms, its native micro-architectured pipeline ensures hard geofencing, regulatory claim auditing, and strict open-source safety.

Sovereign Control Capabilities

- Open Standard Identity & Access (OIDC/OAuth2)
- Declarative Sovereign Access Policies (RBAC/ABAC)
- Cryptographic Workload Identities (SVIDs/mTLS)
- Metadata-Restricted Vector Search Datastores
- High-Performance Inference Execution Gateways

© 2026 Federis AI Inc. All Rights Reserved.

F FEDERIS AI

COMPARATIVE MATRIX

CAPABILITY COMPARISON TABLE

Feature / Capability	Federis AI	Mistral AI	Core42	Aleph Alpha	Sarvam AI	HPE / Mirantis	SambaNova	Bharat GPT / Innefu
Native Ground-Up Stack	● Yes	● Partial	● Partial	● Partial	● Partial	● No (Assembled)	● Yes (HW focus)	● Partial
Hardware-Based Geo-Locking	● Yes	● No	● No	● No	● No	● No	● No	● No
Cryptographically Safe Logging & Audit	● Yes	● Basic	● Basic	● Yes	● Basic	● Via Add-ons	● No	● Basic
Built-in Compliance & Policies	● Yes	● API level	● Yes	● Yes	● Regional	● Via Add-ons	● Basic	● Regional
Native Guardrails (Pre/Post-Proc)	● Yes	● Add-on	● Add-on	● Yes	● Add-on	● BYO	● BYO	● Basic
Air-Gapped Deployment (Zero Call)	● Yes	● Yes	● Yes	● Yes	● Hybrid	● Yes	● Yes	● Yes
Primary Vision / Market Focus	**Native Sovereignty**	Open Weight Enterprise	National Infrastructure	EU Lawful AI	Indian Language Enterprise	Enterprise Cloud / MLOps	Inference Infrastructure	Public Service / Defense

© 2026 Federis AI Inc. All Rights Reserved.

F FEDERIS AI

OPERATIONS BRIEF

GETTING STARTED

1. PLANNING AND SOVEREIGNTY SCOPING

Our teams collaborate with your compliance and platform architects to build a sovereign boundary profile, mapping data sensitivity classifications (e.g. secret, top secret) to authorized physical server topologies.

2. DECLARATIVE GITOPS BOOTSTRAPPING

Engineers deploy target configurations using declarative GitOps manifests. The setup establishes system cells, configures workload identities, and spins up local vector endpoints.

3. OPERATIONS & MAINTENANCE LIFECYCLE

Choose from flexible enterprise support packages. From standard developer updates to fully managed zero-trust operational pipelines, our engineers keep systems updated and running smoothly 24/7.

DEPLOYMENT SPECS

Control Plane Cluster:
- 3x HA Nodes
- Each: min 8x vCPUs, 32GB RAM, 200GB SSD

Inference Cluster:
- 2x Redundant GPU Nodes
- Each: min 16x vCPUs, 64GB RAM, 1TB model cache PVC, and dedicated enterprise GPU lanes

Kubernetes Target:
- Native K8s v1.28+ with strict namespace routing and mTLS configurations.

See Federis AI In Action

Simplify regulatory compliance and zero-trust delivery for your sovereign AI workloads today.

Ashok Gautam | +91-9810800861 | ashok@federis.in

© 2026 Federis AI Inc. All Rights Reserved.