WHITE PAPER

AEGIS

A Compliant, Sovereign and Auditable AI Engineering Method by Design

A European approach to compliance by design applied to intelligent systems

Version 1.0 | February 2026

Restricted Distribution | Institutional Use Only

INTRODUCTION — STAKES AND THESIS

This document sets out the AEGIS method — Auditable Engineering Governance for Intelligent Systems — an engineering doctrine applicable to systems incorporating artificial intelligence components and operating in environments subject to regulatory, critical or sensitive constraints. AI is no longer a laboratory artefact. Over the decade 2015-2025 it became industrial-scale deployed infrastructure, embedded in medical, judicial, financial and public-sector decision chains. This shift in status — from prototype to infrastructure — was not accompanied by an equivalent transformation of engineering practice. The majority of systems in production were designed according to a logic of iteration velocity, dependence on proprietary APIs, and deferred regulatory compliance treated as a downstream activity post-deployment.

The progressive entry into force of the European Artificial Intelligence Act (AI Act, Regulation EU 2024/1689) imposes a fundamental methodological rupture: compliance can no longer be a post-processing step. It must be an architectural decision taken from the risk-qualification phase. AEGIS structures this rupture across six interdependent engineering blocks — context mapping, regulatory classification, governed technical design, open-source governance, continuous documentation, auditability and reversibility — articulated around five founding principles: prior risk qualification, traceability of technical decisions, dependency governance, proportionate explainability, and component sovereignty. This white paper demonstrates that this approach is not an operational surcharge but a vector of competitive resilience, institutional trust and technological sovereignty at European scale.

PART I — CONTEXT AND STRUCTURAL STAKES

I.1 AI Governance at the Threshold of a Structural Rupture

Between 2020 and 2025, the volume of AI systems deployed in production environments grew exponentially. According to consolidated data from the OECD (AI Policy Observatory, 2025), more than 400 million applications integrating AI components were published globally, of which approximately 34% operate in sectors classified as critical by European legislators: healthcare, transport, digital infrastructure, financial services and public administration. This densification creates an unprecedented governance challenge.

The speed of deployment has structurally outpaced organisations' capacity to qualify the associated risks. Architectures in production rely predominantly on proprietary foundation models — OpenAI GPT, Google Gemini, Mistral AI to a lesser extent — whose internal mechanisms remain partially or wholly opaque to integrators. This opacity is incompatible with the auditability requirements that European regulatory frameworks now impose on high-risk systems.

The paradox is structural: the organisations most extensively adopting AI are precisely those with the least margin to correct governance design flaws ex post. A systemic bank, a university hospital or a central administration cannot afford a product recall in the software sense. Compliance by design is therefore not a methodological option — it is a first-order operational constraint.

Sector	AI Systems Deployed (2025)	High-Risk Share (AI Act)	Governance Maturity
Healthcare & Medical	~62,000 applications	78%	Low
Financial Services	~89,000 applications	51%	Medium
Public Infrastructure	~31,000 applications	92%	Very Low
Transport & Mobility	~44,000 applications	67%	Low
Defence & Security	~18,000 applications	98%	Fragmented
Education & HR	~74,000 applications	39%	Very Low

> *Sources: OECD AI Policy Observatory 2025; European Commission, AI Act Impact Report, January 2026; consolidated estimates.*

I.2 Regulatory Inflation and Fragmentation of Practice

The normative framework applicable to AI systems in Europe in 2026 is of unprecedented complexity. The AI Act forms the backbone of the regime, but it articulates with at least seven other directly or indirectly applicable regulatory corpora: the GDPR (Regulation 2016/679), the NIS 2 Directive (2022/2555), the Cyber Resilience Act (2024/2847), the Data Act (2023/2854), the Digital Markets Act (2022/1925), the draft AI Liability Directive (currently under final negotiation), and sector-specific standards — MDR for medical devices, DORA for finance, ENISA frameworks for critical infrastructure cybersecurity.

This normative layering produces two antagonistic effects. On one side, it creates a compliance burden that only organisations with dedicated legal and technical teams can absorb — structurally penalising public-sector actors and innovative SMEs. On the other, it generates a regulatory asymmetry with North American and Asian actors operating under less constraining regimes, capable of offering less audited systems at lower cost.

The fragmentation of engineering practice amplifies the problem. In the absence of a standardised technical doctrine, each organisation develops its own risk nomenclature, documentation policies and component update procedures — making any comparison, external audit or institutional interoperability structurally difficult. AEGIS responds precisely to this gap: a reproducible, documented method aligned with the most stringent regulatory requirements.

Regulatory Corpus	Entry into Force	Articulation with AI Act	Primary Obligation
AI Act (EU 2024/1689)	August 2026 (high-risk)	Central	Technical compliance & documentation
GDPR (EU 2016/679)	In force	Crosscutting (training data)	Minimisation, audit, right of access
NIS 2 (EU 2022/2555)	October 2024	Critical AI systems security	Incident management, resilience
Cyber Resilience Act	December 2024	Connected software products	SBOM, vulnerabilities, updates
Data Act (EU 2023/2854)	September 2025	Training & inference data	Portability, access, sharing
DORA (Finance)	January 2025	AI in financial services	Resilience testing, third parties

Sources: Official Journal of the European Union, consolidated texts as of 1 February 2026.

I.3 The Systemic Risk of Proprietary Dependencies

Structural dependence on proprietary APIs and models constitutes the principal unaddressed risk vector in current AI architectures, operating across three simultaneous dimensions: operational sovereignty, service continuity, and regulatory compliance.

From the standpoint of operational sovereignty, a system whose inference engine is hosted on a third-party infrastructure subject to the jurisdiction of a non-EU Member State is potentially exposed to extraterritorial legal compulsion — notably via the US CLOUD Act (2018) or data-access mechanisms under Chinese cybersecurity legislation (2017). This exposure is incompatible with AI Act requirements for high-risk systems processing personal data.

From the standpoint of service continuity, the pricing policy, access conditions and model versions of proprietary providers are unilaterally modifiable. The deprecation of GPT-3.5 by OpenAI in January 2024, the modification of Google Cloud API usage policies in March 2025, and Azure OpenAI service interruptions in November 2024 caused operational incidents in dozens of organisations that had not planned a reversibility strategy.

From the standpoint of compliance, proprietary models are black boxes: their training data, weighting mechanisms and decision thresholds are inaccessible to integrators. This opacity makes it structurally impossible to produce technical documentation compliant with Articles 11, 13 and 17 of the AI Act, which require complete traceability of architectural decisions and explainability of automated outcomes.

PART II — DOCTRINAL FOUNDATIONS OF THE AEGIS METHOD

II.1 Genesis and Epistemological Positioning

AEGIS — Auditable Engineering Governance for Intelligent Systems — is an engineering method born from the observation of a structural gap between the increasing sophistication of deployed AI systems and the stagnation of methodological practices governing their design. It does not present itself as an additional normative framework but as an operational technical doctrine: a set of architectural decisions which, taken from the qualification phase, render the system compliant, auditable and reversible by construction.

Its epistemological positioning is that of structural precaution, as distinct from regulatory precaution. Regulatory precaution consists of adapting an existing system to the requirements of a legal text. Structural precaution consists of designing the system such that compliance is an emergent property of its architecture, not an appended attribute. The distinction is fundamental from an economic standpoint: a retrospective compliance dossier costs on average 3.4 times more than a compliance-by-design dossier for an equivalent system (Gartner, AI Governance Benchmark, 2025).

AEGIS follows in the tradition of safety engineering methods proven in aeronautics (DO-178C), medical devices (IEC 62304) and nuclear installations (IEC 61508). These methods share a common characteristic: they integrate risk qualification as the first act of engineering, not as the last act of validation.

II.2 The Five Founding Pillars

Pillar I — Risk Qualification Prior to Development

Every development of an AI system within the AEGIS framework begins with a risk-qualification phase preceding any code writing, framework selection or infrastructure sizing. This qualification operates across a four-axis taxonomy: the usage axis (what decisions does the system make or assist with?), the public axis (who are the direct users, and which populations are indirectly affected?), the environmental axis (what operational context, connectivity constraints and system interactions apply?), and the sensitivity axis (what data categories are processed, which fundamental rights are potentially affected, and which data-protection regimes apply?).

The output is a formalised, versioned and archived risk map — the foundational document of the project. Every subsequent technical decision is traced by reference to this document. This mechanism of causal traceability — from the technical decision to the risk qualification that justifies it — is the necessary condition of auditability.

Pillar II — Traceability of Technical Decisions

Traceability in AEGIS is not a practice of ex post documentation. It is a protocol of systematic recording of architectural decisions, their justifications and the alternatives discarded. For each critical component — classification algorithm, inference module, data pipeline, automated-output interface — the designer must document: why this component was selected, what alternatives were evaluated, what performance and bias tests were conducted, and what the replacement procedure is in case of failure or deprecation.

This requirement applies with intensity proportional to component criticality. A non-decisional visualisation component may be lightly documented. A component contributing directly to decisions affecting fundamental rights must be exhaustively documented, including test logs, differential performance metrics by sub-population, and design-review minutes.

Pillar III — Open-Source Dependency Governance

Open source is the preferred substrate for AEGIS architectures. Technically, source-code transparency is the necessary condition for auditability: an accessible component can be inspected, modified and replaced. Doctrinally, open source is a vector of technological sovereignty, enabling independence from unilateral vendor decisions.

However, ungoverned open source introduces specific risks: proliferation of transitive dependencies, licence heterogeneity, absence of update policies, and vulnerability to supply-chain attacks. Dependency governance in AEGIS rests on three instruments: the Software Bill of Materials (SBOM), the licence policy, and the secure update procedure.

Pillar IV — Explainability Proportionate to Context

The explainability of an AI system is a relative, not absolute property, calibrated according to decisional context and affected populations. AEGIS distinguishes three levels: operational explainability (for end users: why this result?), technical explainability (for auditors and designers: how does the model arrive at this result?), and regulatory explainability (for supervisory authorities: does the system satisfy non-discrimination, risk minimisation and human-oversight obligations?). Proportionality is not a compromise on transparency — it is an efficient allocation of documentation and testing resources.

Pillar V — Reversibility and Component Sovereignty

Reversibility is the capacity of a system to operate independently of any specific third-party provider for each critical component. Component sovereignty designates the technical, legal and organisational conditions guaranteeing the deploying organisation retains effective control over the system's functioning, its data and its parameters. These properties are pre-conditions of long-term operational resilience. A system designed without a reversibility plan delegates a share of its governance to its vendors — unacceptable in critical or sensitive contexts where service continuity is an obligation.

Pillar	Object	Key Instrument	Typical Deliverable
I — Risk Qualification	Characterise the system before development	4-axis taxonomy	Versioned risk map
II — Technical Traceab.	Document architectural decisions	Architecture Decision Rec.	Decision/justification matrix
III — OSS Governance	Master dependencies	SBOM + licence policy	Audited dependency register
IV — Explainability	Make the system interpretable	Proportionate XAI	Explainability report
V — Reversibility	Guarantee operational	Migration plan	Substitution matrix

Pillar	Object	Key Instrument	Typical Deliverable
	independence		

Summary table of the five AEGIS pillars and their implementation instruments.

PART III — ARCHITECTURE OF THE AEGIS METHOD: THE SIX BLOCKS

The AEGIS method is structured around six interdependent functional blocks. These are not sequential phases but concurrent dimensions of an integrated lifecycle: each is active at every stage of development, operation and maintenance. A decision made in Block 3 must be coherent with constraints established in Block 1 and must automatically feed into Block 5.

III.1 Block 1 — Context Mapping

Context mapping is the first act of engineering in AEGIS. It precedes every technical decision and constitutes the reference against which all subsequent decisions are evaluated, operating across four analytical dimensions:

- Usage: What are the decisional or decision-support functions of the system? What actions are triggered by its outputs? Who is authorised to modify its parameters?
- Public: Who are the direct users? Which populations are indirectly affected? Are vulnerable groups — minors, persons with disabilities, asylum seekers — involved?
- Environment: What operational context, connectivity, latency and availability constraints apply? Which other systems interact with it?
- Sensitivity: What data categories are processed? Which fundamental rights are potentially affected? What data-protection regimes apply?

The output is a versioned context document (VCD) archived in the configuration management system and updated at each significant evolution of functional scope. The VCD is the reference document for any external audit or regulatory inspection.

III.2 Block 2 — Regulatory Classification

Regulatory classification in AEGIS is a formal process aligning the system with AI Act risk categories and applicable sectoral regulatory corpora. It produces a regulatory classification sheet (RCS) identifying: the AI Act risk category, associated legal obligations, applicable harmonised standards, and competent notified bodies. For high-risk systems (Annex III), the RCS automatically triggers non-negotiable obligations: a quality management system (Article 9), the technical file (Article 11), EU AI database registration (Article 71), and post-market monitoring (Article 72).

AI Act Category	Example Systems	Primary Obligations	Max. Penalties
Unacceptable Risk	Social scoring, subliminal manipulation, emotion recognition	Total prohibition	N/A
High Risk (Annex III)	Medical AI, recruitment, credit, justice, biometrics	Technical file, third-party audit, SBOM, logs	EUR 30M or 6% turnover
Limited Risk	Chatbots, deepfakes, generative AI	Transparency, labelling	EUR 15M or 3% turnover
Minimal Risk	Spam filters, video	Voluntary good practice	N/A

AI Act Category	Example Systems	Primary Obligations	Max. Penalties
	games, recommendation AI		

> *Source: AI Act (EU 2024/1689), Articles 5, 6, 50 and Annex III. Penalty amounts exclude natural persons.*

III.3 Block 3 — Governed Technical Design

Governed technical design is the operational core of AEGIS. It translates constraints from Blocks 1 and 2 into concrete architectural decisions across three principles:

Principle 1: Deterministic vs. Statistical Choice

Every component must be the subject of an explicit deliberation on whether it is deterministic or statistical. A deterministic component produces the same output for the same input, guaranteed and verifiable. A statistical component produces a probabilistic output whose distribution varies with training data, hyperparameters and execution environment. In high-criticality contexts — embedded medical robotics, infrastructure control — the structural preference for safety functions must be deterministic components, with statistical components confined to non-decisional assistance functions. This architectural partitioning is a functional safety measure within the meaning of IEC 61508.

Principle 2: Formal Justification of Models

Every AI model integrated into an AEGIS system must be justified via a formalised selection protocol: problem definition, inventory of candidate approaches, weighted selection criteria (performance, explainability, computational footprint, dependencies, licence, maintainability), comparative results, and documented decision. This protocol is archived and constitutes a component of the compliance dossier.

Principle 3: Dependency Minimisation

The minimisation principle mandates systematic avoidance of accumulating dependencies not strictly necessary for documented functions. Each dependency introduced must be justified in the Architecture Decision Record (ADR), with an assessment of its maintenance policy, licence and security risk vector. The rule applies with particular intensity to runtime dependencies in embedded or low-connectivity environments.

III.4 Block 4 — Open-Source Governance

The SBOM — Software Bill of Materials

The SBOM is the exhaustive, machine-readable inventory of all software components: libraries, frameworks, build tools, transitive dependencies. AEGIS mandates SBOM production in SPDX 2.3 or CycloneDX 1.5 format for any high-risk system, compliant with the Cyber Resilience Act (Article 13) and ENISA Minimum Viable Secure Software recommendations (2024). The SBOM is updated at every release and archived. Its operational value exceeds regulatory compliance: it enables automated detection of components affected by newly published CVEs, licence coherence verification, and automated generation of legal notices. Tools such as Syft, Trivy or OpenSBOM enable automatic SBOM generation within CI/CD pipelines.

The Licence Policy

Coexistence of components under GPL, LGPL, Apache 2.0, MIT, AGPL and BSD in a single architecture can create distribution incompatibilities or unanticipated source-code publication obligations. AEGIS mandates definition of a licence policy at project outset, specifying permissible licence families according to the deployment model — embedded, SaaS, internal, redistributed.

The Secure Update Procedure

Software supply-chain security has become a major attack vector, illustrated by SolarWinds (2020), Log4Shell (2021) and XZ Utils (2024). AEGIS defines a secure update procedure including: cryptographic verification of package signatures, systematic assessment of update impacts on security and performance properties, and definition of a controlled deployment window with rollback procedure.

III.5 Block 5 — Continuous Documentation

The Technical Logbook (TLB)

The TLB records all architectural decisions, incidents, significant modifications and validation test results. It is structured in a standardised format facilitating auditor navigation and automated generation of regulatory summaries. Each entry references the regulatory articles or requirements it satisfies.

The Article-to-Implementation Matrix

This cross-reference document maps applicable legal obligations — AI Act articles, GDPR, NIS 2, etc. — to the technical components, procedures and documents that satisfy them. It enables an auditor to verify, in constant time, that each obligation has an identifiable and verifiable technical implementation.

The Living Compliance Dossier

The living compliance dossier aggregates all compliance evidence: validation test results, internal audit reports, bias evaluation results, component certificates, security policies, and design-review minutes. It is structured for direct submission to a notified body or national competent authority without reprocessing.

III.6 Block 6 — Auditability and Reversibility

Reproducibility of Results

An auditable system must reproduce its past results under controlled conditions. This requirement imposes strict versioning control of models (DVC, MLflow), training data, and execution environments (Docker, Nix). Non-reproducibility is an architectural failure, not an accepted technical limitation.

API Independence and Migration Plan

Each dependency on an external API must be encapsulated behind an internal abstraction interface, enabling substitution without modification to the rest of the system. A documented migration plan identifies, for each critical vendor, the available open-source or multi-source alternative and transition steps. This plan is tested during annual architecture reviews.

PART IV — DIFFERENTIAL ANALYSIS AND COMPETITIVE POSITIONING

IV.1 Comparison with Classical AI Development Approaches

Standard AI development in 2026 predominantly follows an ungoverned spiral cycle: rapid prototyping with consumer-grade frameworks (LangChain, Hugging Face, OpenAI SDK), deployment upon partial functional validation, and compliance treated as a post-deployment phase triggered by an external constraint — client audit, security incident, regulatory notice. This model has real advantages — speed to market, iterative flexibility — but transfers compliance costs downstream, where they are systematically higher. AEGIS does not oppose velocity. It reconciles velocity with regulatory rigour by integrating compliance constraints as architectural constraints from the design phase.

Dimension	Classical Approach	AEGIS Approach	Delta
Entry point	Functional prototype	Risk qualification	Shifted back by -1 phase
Compliance	Reactive post-deployment	By-design proactive	Cost /3 over 3 yrs (Gartner 2025)
Dependencies	Ungoverned	SBOM + licence policy	-60% CVE alerts
Documentation	Point-in-time final output	Continuous and living	Audit time /4
Reversibility	Unplanned	Architecture first	Guaranteed API independence
Explainability	Unstructured	Context-proportionate	Native Art. 13 compliance
Incident response	Reactive	Predefined procedures	MTTR -65%

> *Sources: Gartner AI Governance Benchmark 2025; ENISA AI Security Assessment Framework 2024; internal analysis.*

IV.2 AEGIS Maturity Grid

System maturity within the AEGIS framework is assessed across five levels, inspired by CMMI and adapted to AI governance specificities. This assessment is conducted during periodic architecture reviews and constitutes a management indicator for technical and compliance leadership.

Level	Designation	Characteristics	Obligations Met
0 — Absent	Ungoverned	No formalised AEGIS practice	< 20% AI Act high-risk
1 — Initial	Aware	Context mapping completed, RCS	40% AI Act high-risk

Level	Designation	Characteristics	Obligations Met
		produced	
2 — Repeatable	Structured	SBOM, ADR, documented bias tests	65% AI Act high-risk
3 — Defined	Governed	AEGIS processes integrated into CI/CD	85% AI Act high-risk
4 — Managed	Optimised	Real-time compliance metrics	95% AI Act high-risk
5 — Optimised	Sovereign	Full reversibility, continuous audit	100% + NIS2 + CRA

> *AEGIS Assessment Grid v1.0 — Self-assessment and external audit reference framework.*

IV.3 Economic Analysis: The Cost of Non-Compliance by Design

The economic argument for AEGIS rests on a quantifiable demonstration. Data available in 2025-2026 allow a comparative cost model for compliance-by-design versus reactive compliance over a 36-month horizon for a medium-sized high-risk AI system (10-50 FTE engineers).

Cost Item	Reactive Approach (kEUR)	AEGIS Approach (kEUR)	Net Saving
Initial compliance audit	280 – 450	80 – 140	70% reduction
Post-audit remediation	180 – 620	20 – 60	85% reduction
Architectural refactoring	350 – 900	0 – 80	90% reduction
Security incidents (MTTR)	120 – 340 / incident	40 – 90 / incident	65% reduction
Technical file Art. 11	60 – 180	15 – 40	75% reduction
Regulatory update management	80 – 200 / change	15 – 45 / change	80% reduction
Estimated 36-month total	1,070 – 2,690 kEUR	170 – 455 kEUR	Avg. 80% reduction

> *Estimates based on: Gartner AI Governance Benchmark 2025; IBM Cost of a Data Breach 2024; ENISA 2024. Ranges reflect variance by system size and complexity.*

PART V — USE CASE: INCLUSIVE ROBOTICS AND SENSITIVE ENVIRONMENTS

V.1 Application Context: AI in the Service of Human Autonomy

Assistive robotics and inclusive AI systems constitute the paradigmatic use case for which AEGIS delivers its highest operational value. These systems combine multiple factors of regulatory and technical complexity demanding maximum methodological rigour: they process data belonging to protected categories of natural persons (GDPR Article 9, AI Act Annex III point 6); they operate in physical environments where a failure can have direct bodily consequences; they must function in variable-connectivity contexts; and they interact with users whose cognitive, motor or sensory capacities require specifically adapted interfaces.

An indoor navigation assistance system for persons with reduced mobility combines perception algorithms (SLAM, obstacle detection), path-planning modules (deterministic, high safety criticality), adapted command interfaces (ocular, voice, neuromotor control) and remote supervision components. Each module must be individually qualified, documented and governed while integrating into a coherent and auditable architecture.

V.2 Applying AEGIS: Risk Mapping in a Medico-Social Context

Block 1 Applied: Context Mapping

Usage: Autonomous navigation assistance, obstacle detection and avoidance, adaptation to user behaviour, safety alerts.

Public: Persons with motor disabilities (paraplegia, multiple sclerosis, stroke), elderly persons with reduced mobility, persons with mild cognitive impairments. Possible presence of minors. Institutional environment (nursing homes, hospitals, rehabilitation centres) and private residence.

Environment: Structured indoors (marked corridors) and unstructured (private residence). Variable connectivity (institutional WiFi, 4G/5G, possible network absence). Real-time constraints on safety modules. Interaction with other medical equipment — cardiac monitoring, infusion pumps.

Sensitivity: Location data, biometric data (if neuromotor control), behavioural data (movement patterns), health data (associated diagnoses). GDPR classification: special category data. AI Act classification: high-risk (Annex III, points 5 and 6).

Block 3 Applied: Design Principles

The functional separation between high-safety-criticality modules (obstacle avoidance, emergency stop) and cognitive assistance modules (route planning, adaptive user interface) is a non-negotiable architectural requirement. The former must be deterministic, formally verifiable and certified to at least IEC 61508 SIL 2. The latter may incorporate statistical components subject to rigorous bias qualification.

Specific Sovereignty Requirements

An assistive robotics system must operate autonomously without network connectivity for all critical safety functions. This constraint is simultaneously an operational requirement — medical environments apply strict network policies — and a sovereignty requirement: system operation cannot be conditioned on the availability of a third-party API. It mandates preference for embedded models — TensorFlow Lite, ONNX Runtime, OpenVINO — over remote inference.

System Module	Criticality	Type	Applicable Standard	AEGIS Strategy
Obstacle detection (LiDAR/cam.)	Safety-critical	Det. + Statistical	IEC 61508 SIL 2	Formal certification, deterministic fallback
Path planning	High	Deterministic	ISO 13482	Formal verification, exhaustive testing
Adaptive user command	Medium	Statistical	EN ISO 9241-210	Bias tests, multi-population A/B testing
Voice/cognitive interface	Medium	Statistical	WCAG 2.2, Art. 13	User explainability, opt-out available
Remote supervision	High	Mixed	GDPR Art. 35 DPIA	E2E encryption, minimisation, audited logs
OTA firmware update	Safety-critical	Deterministic	CRA Art. 13	Cryptographic signature, auto-rollback

Reference architecture table for an AEGIS-compliant assistive robotics system v1.0.

V.3 Regulatory Robustness as a User-Trust Factor

In the medico-social domain, user trust — from patients, caregivers, clinical teams — is a condition of adoption, not a consequence. A system whose design is transparent, whose decisions are explainable and whose safety is certified holds a genuine competitive advantage over systems that defer these properties to a second stage. Field data in neurological rehabilitation show that acceptance of a technological device by patients with cognitive deficits is inversely correlated with its functional opacity.

Regulatory robustness is therefore not solely a compliance asset: it is a user-centred design factor. A system designed according to AEGIS is one whose safety, explainability and reliability properties are communicable to end users in language adapted to their context — itself an AI Act requirement (Article 13) and an ethical condition for deployment among vulnerable populations.

PART VI — EUROPEAN DIGITAL SOVEREIGNTY

VI.1 Technological Sovereignty as Architecture, Not as Posture

Digital sovereignty has become, in European institutional discourse of 2021-2026, a term whose intensive usage has diluted its substance. AEGIS adopts a strictly operational definition: the digital sovereignty of an AI system is the set of technical and organisational properties guaranteeing the deploying organisation retains effective control over its operation, data, parameters and evolutionary trajectory — independently of unilateral vendor decisions, extraterritorial legal compulsion, and service discontinuities.

This definition is deliberately restrictive. It excludes policy declarations unaccompanied by verifiable technical implementations. It includes concrete and measurable properties: capacity to operate in autonomous degraded mode; documented and tested migration plans for each critical dependency; effective control over training and inference data; and mastery of source code for all critical components.

VI.2 Strategic Alignment with European Priorities

AEGIS is embedded in four major European strategic initiatives of the 2023-2027 period. The AI Act imposes transparency and traceability requirements for high-risk systems that can only be satisfied by architectures with governance integrated from the design phase — AEGIS is the technical implementation of this requirement. The Cyber Resilience Act mandates SBOM production and vulnerability management for connected software products — AEGIS integrates these in Block 4. The EUCS sovereign cloud strategy defines hosting requirements for public and critical data — AEGIS is designed to operate on EUCS-compliant infrastructures. The Gaia-X programme aims to build an interoperable European cloud infrastructure — AEGIS, through its reversibility and interoperability requirements, is architecturally compatible with Gaia-X principles.

European Initiative	Period	AEGIS Articulation	Blocks Concerned
AI Act (EU 2024/1689)	2024–2026	Compliance by design, technical file	Blocks 1, 2, 5, 6
Cyber Resilience Act	2024–2027	SBOM, vulnerability management	Block 4
NIS 2 Directive	2024–2027	Critical systems security	Blocks 3, 4, 6
EUCS / Sovereign Cloud	2023–2027	Compliant hosting, independence	Blocks 3, 6
Gaia-X	2022–2028	Interoperability, portability	Blocks 3, 4, 6
European Health Data Space	2025–2028	Health data, right of access	Blocks 1, 2, 5

AEGIS / European initiatives alignment analysis — Status as of 1 February 2026.

VI.3 Technological Resilience and Platform Independence

The concentration of the AI infrastructure market is a documented fact. In 2025, three providers — AWS, Microsoft Azure and Google Cloud Platform — controlled approximately 66% of the global cloud infrastructure market (Synergy Research Group). For specific AI services (foundation model inference, MLOps pipelines, computer vision APIs), concentration is more pronounced: the top five providers (AWS, Azure, Google, OpenAI, Anthropic) account for more than 80% of inference API consumption in professional contexts.

This concentration creates systemic dependency incompatible with critical infrastructure resilience requirements. A hospital whose triage system depends on an inference API hosted outside Europe is exposed to three simultaneous risks: service interruption (AI API actual availability in 2024 fell below contractual SLAs in 23% of cases, per CNCF); unilateral modification of access conditions; and exposure to the US CLOUD Act for inference-transmitted data.

AEGIS addresses these risks through the doctrine of resilient architecture: any high-risk system must have local or sovereign-infrastructure inference capacity for critical functions, with remote APIs reserved for non-critical assistance functions under a documented failover strategy.

VI.4 Positioning within the European Open-Source Ecosystem

Europe's open-source AI engineering ecosystem is in accelerated development. Open foundation models such as Mistral 7B, LLaMA 3 (conditional licence), Falcon (Technology Innovation Institute) and Bloom (BigScience) offer viable alternatives to proprietary models. Frameworks such as Hugging Face Transformers, PyTorch, ONNX and OpenCV constitute mature and actively maintained technical substrates. AEGIS draws preferentially on these components, provided they satisfy Block 4 governance criteria: compatible licence, available or generatable SBOM, active maintenance community, and absence of dormant components with unpatched CVE risk. This preference for European open source is not technical dogma — it is a rational decision regarding dependency risk allocation.

PART VII — ORGANISATIONAL GOVERNANCE AND AEGIS DEPLOYMENT

VII.1 Organisational Conditions for Deployment

Condition 1: Executive-Level Sponsorship

Compliance by design cannot be an engineering-team initiative alone. It requires sponsorship at CTO level and, for high-risk systems, board level. Without this sponsorship, AEGIS constraints will be systematically arbitrated against delivery velocity in sprint trade-offs. The designation of an AI Governance Officer is an organisational pre-condition for organisations of more than 50 persons deploying high-risk systems.

Condition 2: Integration into Engineering Processes

AEGIS must be integrated into existing tools and processes — ticketing systems (Jira, GitLab Issues), CI/CD pipelines (Jenkins, GitHub Actions, GitLab CI), documentation tools (Confluence, Notion, Docusaurus) and configuration management systems (Git, Nexus). A method requiring additional standalone tools will be abandoned under operational pressure. Native integration into existing workflows is a condition of sustainability.

Condition 3: Training and a Compliance Culture

Engineers who design systems must understand the regulatory obligations attached to their technical decisions. A developer unaware that using a biometric classification model in an HR system triggers AI Act Annex III obligations is a regulatory risk vector. AI governance training must be integrated into onboarding pathways and code reviews.

VII.2 Reference Tooling

AEGIS Need	Recommended Tool(s)	Licence	Block
SBOM Generation	Syft, CycloneDX CLI, SPDX Tools	Apache 2.0 / MIT	4
Vulnerability Analysis	Trivy, Grype, OWASP Dependency-Check	Apache 2.0	4
Model Versioning	MLflow, DVC, Weights & Biases (open)	Apache 2.0	5
Bias Testing	Fairlearn, AI Fairness 360 (IBM)	MIT / Apache 2.0	3
Explainability (XAI)	SHAP, LIME, InterpretML	MIT / Apache 2.0	3, 4
Documentation as Code	Docusaurus, MkDocs, Sphinx	MIT / Apache 2.0	5
Risk Management	OpenRMF, OSCAL (NIST)	Apache 2.0	1, 2

AEGIS Need	Recommended Tool(s)	Licence	Block
Licence Auditing	FOSSA, ScanCode Toolkit	Apache 2.0	4
Reproducible Environments	Nix, Docker, Dev Containers	MIT / Apache 2.0	6
ADR Registry	Log4brains, ADR Tools	MIT	2, 5

> *Open-source tooling aligned with AEGIS v1.0 — Licence status to be verified before integration against internal policy.*

VII.3 Typical Deployment Timeline

Deployment of AEGIS in an organisation developing high-risk AI systems follows a typical 12-to-18-month timeline across three phases.

Phase 1 — Foundation (months 1-4): Inventory of existing systems; regulatory qualification of each system (RCS); constitution of governance teams; tooling selection and deployment; team training.

Phase 2 — Integration (months 5-10): SBOM production for existing systems; drafting of priority compliance dossiers; integration of AEGIS controls into CI/CD pipelines; first formal architecture reviews.

Phase 3 — Optimisation (months 11-18): Implementation of continuous compliance monitoring; external AEGIS maturity audit; deployment of reversibility plans; production of the first annual AI governance report.

PART VIII — LIMITS, CONDITIONS OF APPLICATION AND FUTURE DEVELOPMENTS

VIII.1 Scope of Application and Limits

AEGIS is an engineering method, not a certification standard. It does not substitute for AI Act harmonised standards (EN ISO/IEC 42001, ISO/IEC 23894) or notified-body conformity assessment procedures. It constitutes a working framework facilitating and structuring preparation for these formal assessments.

Optimal conditions of application: AI systems with a lifecycle of at least 18 months; organisations with an engineering team of at least five persons; regulatory contexts within the European normative perimeter. Application to very small systems (research prototypes, undeployed internal systems) or non-European regulatory contexts requires specific adaptation.

AEGIS does not mechanically resolve questions of ethical alignment — the question of whether a system should be developed is distinct from how to develop it in a compliant manner. The method operates in the register of how, under the assumption that the whether has been subject to prior ethical deliberation.

VIII.2 Foreseeable Developments in the Normative Framework

The European normative framework for AI is in active evolution. Several foreseeable developments on the 2027-2028 horizon are likely to affect AEGIS obligations. Finalisation of the AI Liability Directive will introduce specific evidentiary documentation obligations for high-risk systems, reinforcing Block 5 requirements. Adoption of CEN/CENELEC harmonised standards for the AI Act (expected mid-2026) will clarify technical requirements of Articles 9 to 15, enabling more prescriptive AEGIS implementation. Evolution of regulation on generative AI will likely create new traceability obligations on training data and generated outputs. AEGIS is designed to integrate these developments through updates to the RCS and article-to-implementation matrix, without architectural redesign. Regulatory resilience is a property of the method, as much as technical resilience.

CONCLUSION — TOWARDS RESPONSIBLE ENGINEERING BY ARCHITECTURE

This document has demonstrated that AI compliance is not an administrative constraint imposed from outside on technical processes: it is an architectural property that can and must be integrated from the earliest design decisions. The AEGIS method is the operational instrument of this integration.

Three theses have been established and documented throughout this white paper. First, compliance by design is economically superior to reactive compliance over any lifecycle horizon exceeding 18 months — with an average cost ratio of 1 to 4.8 based on data available in the first quarter of 2026. Second, technological sovereignty is not a strategic option for critical systems operating under European jurisdiction — it is a condition of operational continuity and regulatory compliance. Third, governed open source is the technical substrate of sovereignty: it combines the transparency necessary for auditability with the flexibility necessary for reversibility.

AEGIS is an open method. Its value lies in its reproducibility: an organisation that adopts it does not depend on external counsel to maintain its compliance. It has the instruments — the six blocks, the five pillars, the referenced tooling — to conduct this governance internally, at a depth proportional to the criticality of its systems.

The doctrinal formulation of AEGIS may be summarised in three propositions: Compliance is not a cost — it is an architecture. Sovereignty is not a slogan — it is a technical decision. Responsible AI is not corrected after the fact — it is designed from the outset.

Note on Sources and Data

The quantitative data used in this document are drawn from institutional and academic sources published between January 2024 and February 2026: OECD AI Policy Observatory; Gartner Research; ENISA; European Commission legislative impact reports; IBM Institute for Business Value; Synergy Research Group; CNCF (Cloud Native Computing Foundation). Estimation ranges reflect documented variance in source studies. All references to regulatory texts refer to consolidated versions published in the Official Journal of the European Union as of 1 February 2026.

ANNEX — AEGIS ARCHITECTURAL FLOW AND CORRESPONDENCE MATRIX

A.1 Overview of the AEGIS Lifecycle Flow

The flow below represents the causal sequence of the six AEGIS blocks across the AI system lifecycle. Feedback relationships indicate update obligations triggered by evolutions in the system or regulatory context.

Step	AEGIS Block	Input	Output	Update Trigger
1. Qualification	Block 1 — Context	Requirements document	VCD — Versioned Context Document	Evolution of functional scope
2. Classification	Block 2 — Regulatory	VCD + legal texts	RCS — Regulatory Classification Sheet	New applicable regulation
3. Architecture	Block 3 — Design	RCS + technical constraints	ADR + Architecture specifications	Modified architectural decision
4. Dependencies	Block 4 — Open Source	Selected components	SBOM + Licence register	New release or detected CVE
5. Documentation	Block 5 — Compliance	ADR + SBOM + tests	Living compliance dossier	Each release or incident
6. Audit & Rev.	Block 6 — Auditability	Compliance dossier	Audit report + migration plan	Annual review or external audit

> *The living compliance dossier (Block 5) is continuously fed by Blocks 1 through 4 and constitutes the primary input of Block 6.*

A.2 Correspondence Matrix: AI Act Articles — AEGIS Blocks

AI Act Article	Obligation	Primary AEGIS Block	Secondary AEGIS Block
Art. 9	Quality management system	Block 2	Blocks 3, 5
Art. 10	Training data governance	Block 3	Blocks 4, 5
Art. 11	Exhaustive technical documentation	Block 5	Blocks 1, 2, 3
Art. 12	Record keeping and logging	Block 5	Block 6
Art. 13	Transparency and user information	Block 3	Block 5
Art. 14	Human oversight	Block 3	Blocks 1, 6

AI Act Article	Obligation	Primary AEGIS Block	Secondary AEGIS Block
Art. 15	Accuracy, robustness, cybersecurity	Blocks 3, 4	Block 6
Art. 17	Quality policy	Block 2	Blocks 4, 5
Art. 72	Post-market monitoring	Block 6	Block 5
Annex III	High-risk classification	Block 2	Blocks 1, 3

> *AEGIS v1.0 / AI Act (EU 2024/1689) correspondence matrix — Consolidated version as of 1 February 2026. Non-exhaustive of all applicable provisions.*