deepset

A Practical Guide

EU AI ACT · REGULATION (EU) 2024/1689

The EU AI Act: A Guide for AI Teams

Learn how deepset can support your journey to the high-risk compliance deadline.

deepset

Table of Contents

SECTION 01

What Is the EU AI Act?

The EU AI Act **(Regulation (EU) 2024/1689)** is the world's first comprehensive legal framework for regulating artificial intelligence. Adopted by the European Parliament and the Council on **13 June 2024** and published in the Official Journal on **12 July 2024**, it establishes harmonised rules for the development, placing on the market, putting into service, and use of AI systems across the EU.

Its purpose is twofold: to promote uptake of human-centric, trustworthy AI, while ensuring a high level of protection for health, safety, and fundamental rights: including democracy, the rule of law, and environmental protection.

> The EU AI Act is a shift from Experimental AI to Sovereign AI. For enterprises, compliance is no longer just a legal hurdle, it is a competitive advantage. deepset provides the orchestration layer that transforms a raw LLM into a robust, production-ready AI system.
>
> DEEPSET PERSPECTIVE

Key Definitions

AI System

A machine-based system that operates with varying autonomy and may adapt after deployment, inferring from inputs how to generate predictions, content, recommendations or decisions that influence physical or virtual environments.

Provider

The entity that develops or commissions an AI system and places it on the market or puts it into service under its own name or trademark.

Deployer

Any person or organisation using an AI system under its authority: excluding purely personal, non-professional use. Most enterprises are deployers.

EXPLORE Read the **full text of the EU AI Act**, in the AI Act Explorer.

Risk-based approach. Rather than regulating all AI equally, the Act categorises use cases by the risk they pose to individuals and society, and applies obligations proportionate to that risk, targeting effort where potential harm is greatest while leaving low-risk innovation untouched.

SECTION 02 · 03

Who Does It Apply To?

The Act has **broad territorial and personal scope**. AI systems affecting people within the EU are subject to its rules regardless of where the provider or deployer is based; the same extraterritorial logic the GDPR established.

Actor	Description
Providers	Entities placing AI systems on the EU market or putting them into service, whether established inside or outside the Union.
Deployers	Persons or organisations using AI systems under their authority within the EU (excluding personal, non-professional use).
Importers & Distributors	Those bringing AI systems bearing a third-country provider's name onto the EU market, or making them available downstream.
Product Manufacturers	Companies placing products on the market with an embedded AI system under their own name or trademark.
Third-Country Operators	Providers and deployers based outside the EU where the *output* of their AI system is used within the Union.

The Risk Classification Framework

Four tiers, with obligations escalating in proportion to the potential for harm.

Unacceptable Risk
BANNED · ART. 5

Banned outright — social scoring, subliminal manipulation, untargeted facial-recognition scraping. *(Article 5)*

High Risk
COMPLIANCE BY DEC 2027

Heavy obligations + conformity assessment — AI in hiring, credit scoring, law enforcement, critical infrastructure. *(Art. 6–49)*

Full list of high-risk use cases: Annex III

Limited Risk
TRANSPARENCY · ART. 50

Transparency obligations — chatbots, deepfakes must disclose AI involvement. *(Art. 50)*

Minimal Risk
NO SPECIFIC OBLIGATIONS

Unregulated — spam filters, AI in video games, inventory management, most low-stakes tooling.

ASSESS Check which risk tier **your AI use case or system** falls into with the EU AI Act Compliance Checker.

SECTION 04

Penalties & Timeline

Three-Tier Penalty Structure (Article 99)

€35M or 7% global turnover	**Prohibited Practices** Violations of Article 5: social scoring, banned biometric ID, subliminal manipulation.	Whichever is *higher*. The heaviest tier, reserved for unacceptable-risk breaches.
€15M or 3% global turnover	**High-Risk Obligations** Non-compliance by providers, deployers, importers, distributors, or notified bodies; Art. 50 transparency.	Whichever is *higher*. Main target of the December 2027 deadline.
€7.5M or 1% global turnover	**Incorrect Information** Providing incomplete, incorrect or misleading information to notified bodies or competent authorities.	Whichever is *higher*. SMEs and start-ups receive a softer cap (whichever is *lower*).

Compliance Timeline

Timeline as amended by the Digital Omnibus on AI **(Regulation (EU) 2026/1744)**, published in the Official Journal on 24 July 2026 and in force since 27 July 2026 - the first formal amendment to the AI Act.

Feb 2, 2025
Prohibitions on **Unacceptable Risk** systems take effect, and AI literacy duties.

Aug 2, 2025
General Purpose AI governance rules & systemic-risk obligations kick in.

Aug 2, 2026
Transparency obligations (Art. 50) apply: chatbots and AI-generated content must disclose AI involvement.

Dec 2, 2027
Full compliance for High-Risk AI.

Aug 2, 2028
Obligations for AI **embedded in regulated products**.

How fines are calibrated. Authorities weigh the nature, gravity and duration of the infringement; the number of people affected and degree of harm; any prior fines for the same breach; the operator's size, market share and benefit gained; cooperation with regulators; and the technical and organisational safeguards in place.

SECTION 05

From “Black Box” AI to Compliance-Ready

For customers, the EU AI Act is **not a barrier but a quality standard** in the era of digital sovereignty. The Act distinguishes the **Provider** (who creates the AI) from the **Deployer** (the customer/user). Haystack provides infrastructure that primarily supports Deployers in their compliance obligations (Arts. 26–27) and in evidencing the high-risk requirements of Articles 8–15.

As organizations face immediate regulatory deadlines, the core challenge of AI adoption is no longer performance, it is control. Most enterprise AI implementations fail compliance audits because they rely on opaque, black-box systems where **data handling is invisible, model behaviour is untraceable, and vendor lock-in prevents architectural changes**.

DEFINITION

Sovereign AI

The practice of designing, deploying, and operating artificial intelligence while maintaining **complete, autonomous control** over four key vectors:

Data	Models	Infrastructure	Governance
Your sources, your residency	Open-weight or proprietary, swappable	VPC or self-hosted	Audit trails & oversight

The Haystack Enterprise Platform is a Sovereign AI Platform designed specifically to give organizations this level of control. Rather than forcing you to send proprietary data or citizen metrics to third-party endpoints, Haystack acts as a fully compliant orchestration layer. It can be deployed entirely within your own sovereign environment: a private VPC or self-hosted.

With the Haystack Enterprise Platform, deepset provides the orchestration, transparency, and oversight tools that wrap around any model: OpenAI, Azure, Llama, Mistral, to ensure that AI teams can verify that their most business-relevant use cases are legally deployable. While the model provides the intelligence, the Haystack Enterprise Platform is designed to support you in fulfilling the core EU AI Act requirements even for high-risk use cases and help you to transform "Black Box" AI into governed, sovereign systems by providing the mechanisms for continuous qualitative review, human oversight, and transparency.

SECTION 06

How Haystack Supports Compliance

Mapping the core legal mandates of the EU AI Act for High-Risk systems to specific capabilities in the Haystack ecosystem.

Requirement	Summary of Mandate	Haystack Enterprise Platform's Role as Enabler
Accuracy, Robustness & Cybersecurity Art. 15	High-risk AI systems must achieve appropriate levels of accuracy, robustness, **and cybersecurity**, and perform consistently across their lifecycle. Accuracy levels and the relevant metrics must be declared in the instructions for use (15(3)). Resilience to attacks such as data and model poisoning and adversarial examples sits under the cybersecurity section (15(5)).	**Evaluation Module (measurement, not declaration).** The Evaluation Module lets you measure accuracy and retrieval quality repeatably, generating the evidence base for the 15(3) declaration, but it does not itself constitute a declared accuracy level, which the provider must author.
Human Oversight Art. 14	High-risk systems must be designed so that qualified humans can monitor operations, detect anomalies, remain aware of automation bias, correctly interpret outputs, and at any point override, reverse, or halt the system entirely.	**Agents & HITL (Human-in-the-loop).** HITL approval gates before agent tool execution and operator-controlled pipeline undeployment map to Art. 14(4)(d)–(e). Search History, component-level logs and source-referenced answers support 14(4)(a)–(c). Automation-bias awareness (14(4)(b)) is an interface/training matter the platform does not address.
Data Governance Art. 10	Training, validation and testing datasets must meet quality criteria and follow documented governance practices. Under Art. 10(6), systems that do not use model-training techniques are subject to paras. 2–5 for their testing datasets only.	**Privacy-First Inference.** A retrieval-based architecture narrows Art. 10 to testing datasets under 10(6), rather than the full training-data regime; it does not exempt systems from Art. 10, and any fine-tuning or model training in the pipeline reopens the full article. Haystack supports versioned, inspectable evaluation datasets to evidence testing-set quality.
Transparency Art. 13	High-risk systems must be designed so their operation is sufficiently transparent for deployers to interpret outputs appropriately, accompanied by clear instructions covering the system's intended purpose, accuracy metrics, known limitations, and the human oversight measures in place.	**Modular Traceability.** Supports 13(1) interpretability: every step (Retrieval → Context Engineering → LLM) is logged and inspectable, and source-document attribution lets a deployer trace an answer to its retrieved evidence.
Record Keeping Art. 12	High-risk systems must automatically log events throughout their lifetime, including usage periods, input data, match results, and the identities of human verifiers, to enable traceability for post-market monitoring, risk identification, and incident investigation.	**Enterprise Logging.** Component-level tracing and centralized logging capture pipeline execution, retrieval steps, tool calls and model responses. Retention duration, log completeness and tamper-evidence remain provider configuration decisions.

SECTION 06 · CONTINUED

How Haystack Supports Compliance

Provider documentation duties, post-market monitoring, and the obligations that fall on you as a deployer.

Requirement	Summary of Mandate	Haystack Enterprise Platform's Role as Enabler
Risk Management Art. 9	Providers must run a continuous risk management process throughout the system's entire lifecycle, identifying and mitigating risks under both intended use and foreseeable misuse, with testing against predefined metrics before market release.	**Testing apparatus for Art. 9(6)–(8).** The Evaluation Module and prompt explorer supply the systematic, repeatable pre-release testing against defined metrics that Art. 9(6)–(8) requires.
Quality Management Art. 17	Providers must maintain a documented quality management system covering compliance strategy, design and development procedures, testing, data management, risk management, post-market monitoring, incident reporting, and an accountability framework.	**Structured Infrastructure (partial).** Modular pipelines, versioned components, reproducible configs and standardized workflows support roughly four of the thirteen Art. 17(1) elements: design/development control (c), verification and testing (d), technical specifications (e) and data management (h).
Technical Documentation Art. 11 + Annex IV	Technical documentation must be drawn up before market placement and kept current, covering the Annex IV elements: general description, design specifications, system architecture, development process, validation and testing.	**Declarative Pipeline Configuration.** Pipeline YAML, versioned components and reproducible configurations produce a machine-readable, diffable record of system architecture, design choices and component versions.
Post-Market Monitoring Art. 72	Providers must establish and document a proportionate post-market monitoring system that actively and systematically collects and analyses performance data across the system's lifetime, based on a written monitoring plan forming part of the Annex IV technical documentation.	**Enterprise Observability.** Centralized logging of queries, tool calls, model responses and agent actions provides the data-collection layer the monitoring system runs on, and supports incident investigation.
Deployer Obligations Art. 26	Deployers of high-risk systems must use them in accordance with the instructions for use, assign trained and competent human oversight, ensure input data is relevant and sufficiently representative, monitor operation, report serious incidents and retain the automatically generated logs for at least six months (longer where other EU or national law requires).	**Deployer-Ready Logging.** Enterprise Logging generates the log trail Art. 26(6) requires deployers to retain, and Search History review supports the ongoing monitoring duty under 26(5). Retention scheduling and the assignment of oversight personnel remain organisational decisions of the deployer.
Fundamental Rights Impact Assessment (FRIA) Art. 27	Bodies governed by public law, private entities providing public services, and certain financial-sector deployers must complete a fundamental rights impact assessment before first use of a high-risk system: describing the deployment context, affected persons, risks of harm, human oversight measures, and mitigations, and notify the market surveillance authority.	**Evidence Base for the FRIA.** Declarative pipeline configuration documents the deployment process, and component-level traceability plus HITL gates supply the concrete oversight and mitigation measures the assessment must describe. The FRIA itself is an organisational document the deployer must author.

SECTION 07

4 Steps to a Sovereign Pipeline

1

Data Retrieval & Filtering

Use a modular **RAG** architecture. Filter and define the data the system is allowed to access.

2

Data Immunity

No data is ever used for training by deepset's infrastructure. Your **IP and PII stay** within your sovereign environment.

3

Monitoring & Evaluation

Use the **Evaluation Module** and search-history review to perform qualitative assessments of the AI's reasoning trace.

4

Human Intervention Gates

Enable the **Human-in-the-Loop** component in agent pipelines. Require manual signatures for sensitive tool calls.

Your Rollout Checklist

Five immediate organisational steps to transition smoothly into full EU AI Act readiness before the **December 2027 high-risk deadline**.

- [x] **Audit and Categorize**
  Identify current and pipeline AI use cases and map them against the Act's four risk tiers (Unacceptable, High, Limited, Minimal).

- [x] **Establish Data Boundaries**
  Implement a modular architecture, like Retrieval-Augmented Generation, to keep sensitive IP and PII strictly within your controlled environment.

- [x] **Implement Human-in-the-Loop Safeguards**
  Embed mandatory human intervention gates into agentic workflows - enabling high-risk tool calls to require a manual signature before execution.

- [x] **Centralize Logging and Observability**
  Set up unified tracking for every query, tool call, and model response to provide the automated audit trail regulators require.

- [x] **Build AI Literacy**
  Train staff operating or overseeing AI systems.

deepset

SECTION 08

Compliance is an Architecture Decision.

The EU AI Act's requirements: accuracy, transparency, human oversight, data governance, risk management, ongoing monitoring are not just legal boxes to tick. They are the engineering practices that separate production-grade AI from ungoverned experimentation.

The Act defines *what* enterprises must achieve. Haystack provides the *how*: a modular orchestration layer that wraps around any model to deliver traceable retrieval, logged reasoning, human-intervention gates, and continuous evaluation.

The enterprises that treat the EU AI Act as a quality standard rather than a burden will be the ones best positioned to deploy AI at scale with confidence, accountability, and sovereign control.

Contact us →

AUTHORED BY

Dr. Anna Gründler
Senior Director Strategy, deepset

deepset

deepset supports enterprise and public sector organizations to build and run AI systems.

REVIEWED BY

Dr. Lasse Milinski
Chief Solutions Officer, eagle LSP

eagle lsp

eagle LSP supports with EU AI Act compliance, AI governance and risk classification.

This article was written with the support of AI and reviewed by our editorial team and subject matter experts.