June 2025

initiative for applied artificial intelligence

AI Act Governance: Best Practices for Implementing the EU AI Act

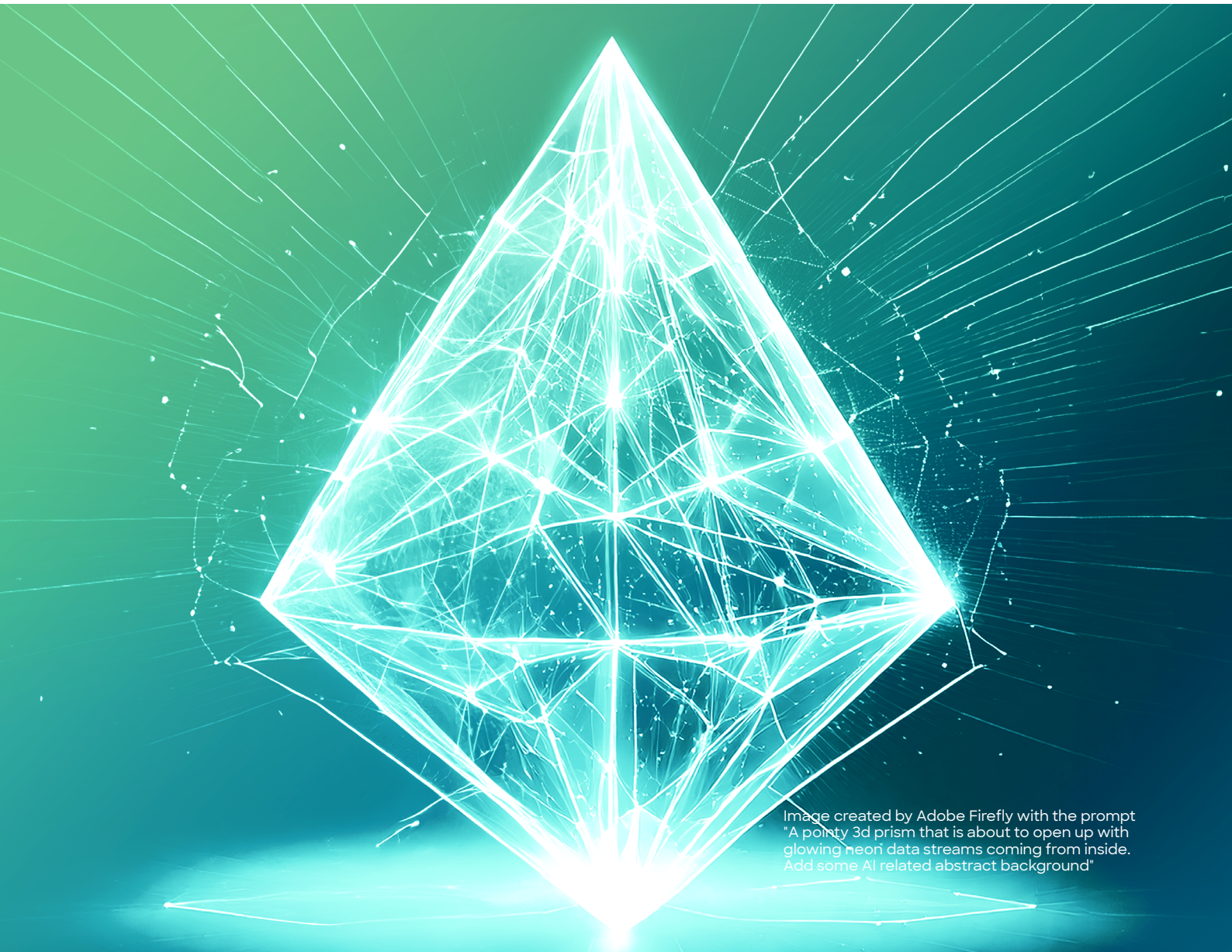


Image created by Adobe Firefly with the prompt "A pointy 3d prism that is about to open up with glowing neon data streams coming from inside. Add some AI related abstract background"

Contents

Introduction 4

Section 1: A Quick Introduction to the EU AI Act 7

- What is the AI Act? 8
- How to Identify Obligations 8
- Technical Obligations Per Risk Class and Role 10

Section 2: Challenges to Operationalize the EU AI Act in the Enterprise 11

- Challenges to Operationalize the EU AI Act 12
- Responding to These Challenges: The appliedAI Working Groups 13
- Working Group Outcomes 13

Section 3: The AI Act Governance Pyramid 15

- What is the AI Act Governace Pyramid? 16
- The AI Act Governance Pyramid Benefits and Limitations 17
- Layer 1: Orchestration Layer 18
- Layer 2: Integration Layer 20
- Layer 3: Execution Layer 21
- Interactions between Layers 23
- Putting it All Together 24

Section 4: Operationalizing Requirements for Providers of High-Risk AI Systems 27

Methodology: An Iterative Approach 28

Overview of Identified ISO/IEC Standards 29

Best Practices for Implementing Requirements of High-Risk AI Systems under the AI Act 31

A Brief Overview on Transparency Obligations and Low Risk AI Systems 47

Section 5: ML Skills Profiles under the AI Act and Getting Started Today 49

ML Skill Profiles under the AI Act 50

Getting Started Today: Bridging the Gap 52

Outlook 53

References 54

Authors, Contributors and Acknowledgements 57

About appliedAI Initiative GmbH 58

Join the appliedAI Partnership 59

Introduction

Integrating regulatory requirements into corporate strategy not only enables compliance but also strengthens companies' competitive advantage [1], generating value through the use of compliant artificial intelligence. However, operationalizing the EU AI Act, in particular, requirements for high-risk AI systems, presents significant operational challenges for enterprises [2, 13].

appliedAI and its partners identified three key challenges: First, companies lack clear guidelines to operationalize the AI Act, particularly how best to orchestrate tasks across the enterprise. Second, the delayed publication of harmonized standards discourages enterprises from exploring high-risk AI systems and prevents technical and legal stakeholders from having a shared understanding of compliance. Finally, companies are unsure about the necessary skill profiles and the first steps they can already take to operationalize the AI Act. As a result, we see a risk that AI innovation and competition in Europe might slow down in response to the AI Act [3].

To address these challenges, this whitepaper captures solutions that appliedAI and its partners developed together through two parallel working groups. We introduce the AI Act Governance Pyramid framework, a structured approach for operationalizing the AI Act by orchestrating stakeholders across enterprise layers. We then compile technical and governance best practices for the AI Act's requirements for high-risk AI systems, including references to available international standards. Finally, we updated appliedAI's ML Skill Profiles framework [4] taking the EU AI Act into consideration, and provided a guide about what companies can start doing today to prepare to operationalize the AI Act.

During 2024, appliedAI has been hosting multiple working group sessions with key corporate partners [5] to explore the challenges and emerging best practices in implementing AI governance under the EU AI Act, as well as considering the AI standardization landscape to strengthen its operationalization.

This whitepaper is based on insights from professionals in the industry and is intended for heads of AI, heads of AI governance, compliance officers, AI project managers, product owners, and engineers who seek to operationalize the AI Act within their organization.

The paper starts with laying out the challenges that occur when companies start operationalizing the EU AI Act. Subsequently, we discuss how these challenges are mitigated by the AI Act Governance Pyramid. We then proceed with an in-depth discussion of best practices to implement each high-risk requirement, and conclude with considerations on the necessary skill profiles and how companies can start doing today to prepare to operationalize the AI Act.

“There is a tendency for some organizations to say we can't do anything right now with regards to the AI Act because standards will only be ready at the beginning of 2026. But that's not quite true. There is a lot that organizations can do already today towards implementing the AI Act.”

Sebastian Hallensleben

Chair of Joint Technical Committee (JTC) 21,
AI CEN and CENELEC

“We are moving towards the implementation of the high-risk AI systems requirements of the AI Act and our first analysis shows that certain articles are already part of the available QMS. A best practice is to set the status quo of your current internal AI infrastructure, calculate the delta with the Requirements of the AI Act and just implement the missing pieces and wrap process and policies around.”

Araceli Alcala

RA Manager | RA SME for Artificial Intelligence,
Carl Zeiss Meditec AG

Section 1:
A Quick Introduction to the EU AI Act

A Quick Introduction to the EU AI Act

What is the AI Act?

The AI Act [6] is a product safety regulation designed to protect health, safety and fundamental rights. It regulates the design, development and use of general purpose AI (GPAI) models [7] and AI systems [8].

Horizontal Regulation

- It applies to all AI applications in all domains and sectors
- It applies to all member states
- Note: Other laws, like the GDPR and sector specific laws, continue to apply as well

Intended Purpose

- The AI Act's application depends on the intended purpose of the AI system
- The AI Act does not govern the technology itself, except for general purpose AI models

Risk Proportional

- AI Act follows a risk-based approach
- The higher the risk of an AI system, the more obligations actors need to meet

New Legislative Framework (NLF) Inspired

- The Act only defines essential requirements, while technical standards define compliance activities
- Implementation monitored by EU and member states authorities

How to Identify Obligations

Enterprises can determine their obligations based on three characteristics: Their role in relation to an AI system, the risk class of an AI system, and whether they integrate general purpose AI models.

1) The Roles

The AI Act takes a value chain approach towards distributing obligations. In other words, in addition to the risk class of an AI system, the obligations an enterprise might face depend on their role in relation to an AI system. While there are several roles under the AI Act, the role of provider and deployer are the most important. In simple language, the provider is an entity that builds or sells a system, while the deployer is an entity that uses the system under its authority.

Provider Article 3(3)

Develops AI System

Provider means a natural or legal person, public authority, agency or other body that **develops an AI system** or a general purpose AI model (or has them developed)

+ and

Places on the Market

Places them on the market or puts the system into service under its own name or trademark, whether for payment or free of charge

→

Deployer Article 3(4)

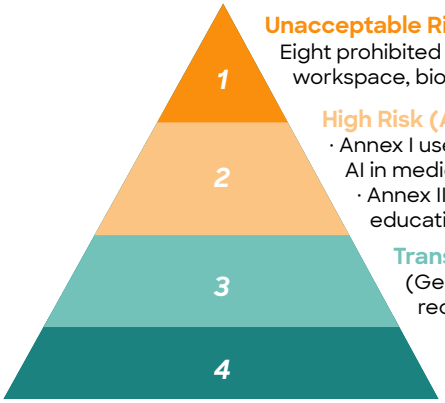
Uses it under its Authority

Deployer means any natural or legal person, public authority, agency or other body **using an AI system under its authority** except where the AI system is used in the course of a personal non-professional activity

A Quick Introduction to the EU AI Act

2) Risk Classification for AI Systems

For the most part, the obligations that an enterprise must meet are defined by the intended purpose of their AI system, and not by the inherent AI capabilities. Therefore, the obligations are proportional to the risk that the intended use of an AI system might pose to the health, safety, and fundamental rights of Europeans. The Act defines four "classes" of risk that an AI system might pose as reflected in the figure below.



Unacceptable Risk (Art. 5)
Eight prohibited use cases, *e.g.* emotion recognition in the workspace, biometric categorization

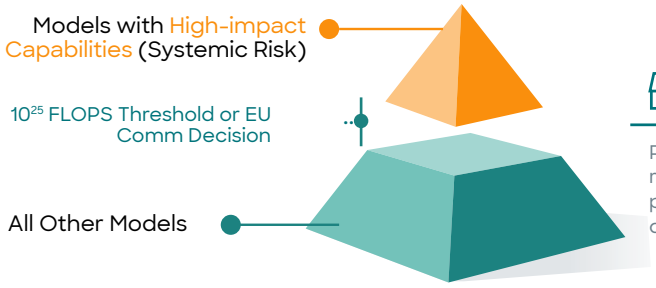
High Risk (Art. 6)
· Annex I use cases : Products or safety components of products, *e.g.* AI in medical devices or vehicles
· Annex III use cases: Standalone AI systems *e.g.* recruiting, education, critical infrastructure

Transparency Obligations (Art. 50)
(General purpose) AI systems and biometric or emotion recognition systems *e.g.* chatbots, advertisements

Low Risk (Art. 95)
All other AI systems

3) Using GPAI Models

The AI Act also creates rules for so called general purpose AI model (GPAI Models), typically understood foundation models such as LLMs. For the most part, these rules are for actors who train these models, such as big tech corporations like OpenAI, Alphabet, etc. These rules depend on two characteristics - whether the model poses a systemic risk to the EU and if it does not pose a systemic risk, whether it was released under an open-source licence or a proprietary licence. 'General-purpose AI model' means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market Art.3(63).



Placing on the Market

Placing on the market can include release through libraries, application programming interfaces (APIs), as direct download, or as physical copy.

Proprietary

Free Open Source (FOS)

Most enterprises do not train their own general purpose AI models. Instead, they integrate these models into their own systems. Under the EU AI Act, such AI systems are known as general purpose AI systems. Enterprises should take note of two important points. First, the use of general purpose AI systems must be classified according to the same 4-tier risk class framework. Second, fine-tuning or modifying the underlying general purpose AI model might result in more obligations.

The EU Commission is expected to publish the final GPAI guidelines in the summer of 2025 [9].

2-tier classification

GPAI Model → **GPAI System**

GPAI Model Provider → **Downstream Provider**

- 'General-purpose AI system' means an AI system which is based on a general-purpose AI model and which has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems **(Art.3(66))**
- It is important to note that although AI models are essential components of AI systems, they do not constitute AI systems on their own. AI models require the addition of further components, such as for example a user interface, to become AI systems. **(Recital 97)**

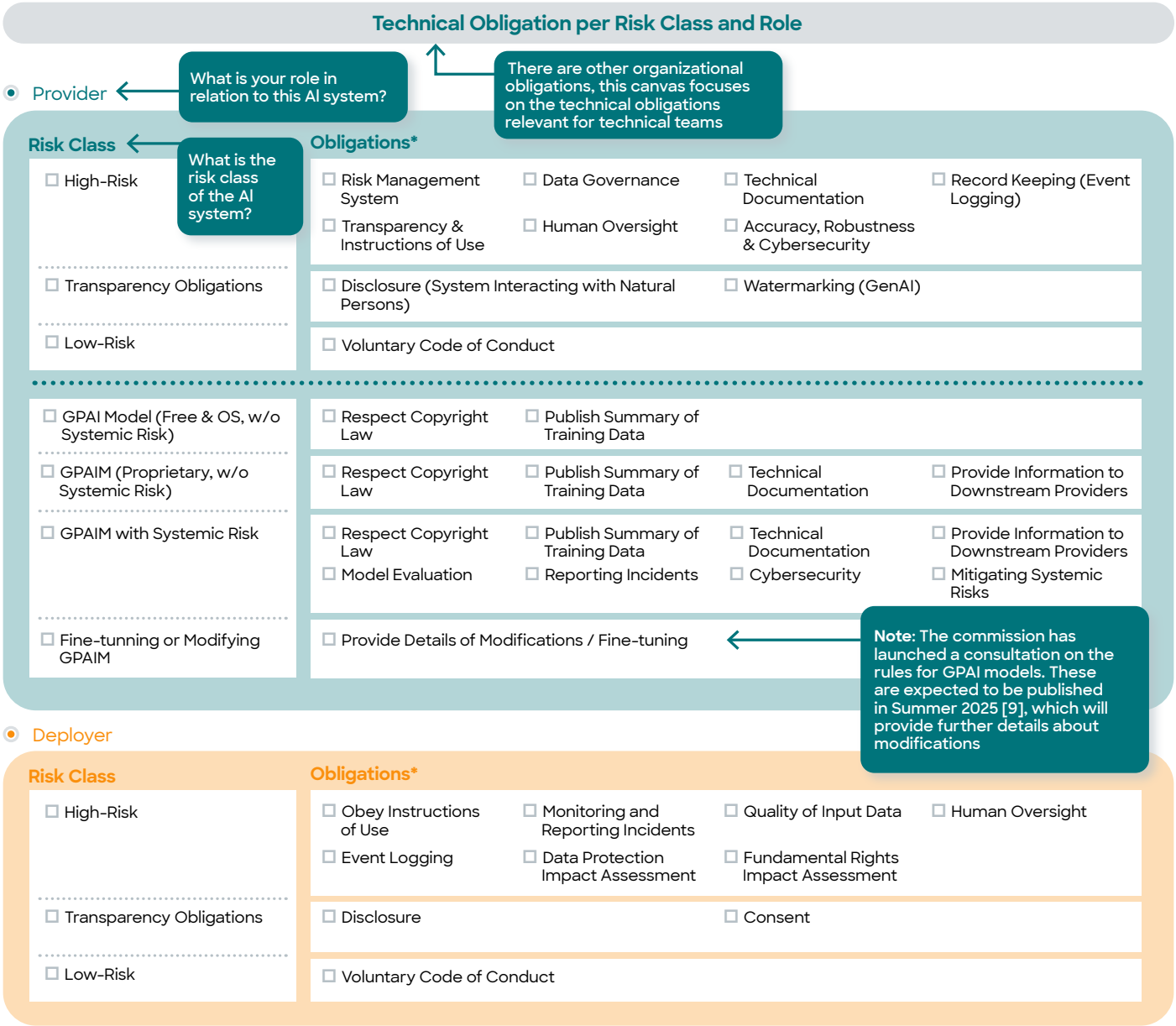
4-tier classification

A Quick Introduction to the EU AI Act

Technical Obligations Per Risk Class and Role

Considering the roles, the risk class of an AI system, and the use of general purpose AI models, an enterprise might face any of the following technical obligations listed in the table below.

Disclaimer: We focus here only on the technical obligations, i.e., those relevant to designing, building, and using an AI system. There are more obligations for providers and deployers, such as responding to incidents and co-operating with authorities, that we exclude from the overview below.



This whitepaper focuses on the technical requirements for providers of high-risk AI systems as listed in **chapter III, section 2 of the AI Act.** As we will discuss in subsequent sections of this paper, the AI Act only defines the high-level requirements. The concrete activities that will give companies a presumption of conformity with these requirements will be defined by harmonised technical standards. The standards [10] for the EU AI Act are currently being developed by CEN-CENELEC JTC 21 [11] and are expected to become available early 2026.

Section 2: Challenges to Operationalize the EU AI Act in the Enterprise

Challenges to Operationalize the EU AI Act in the Enterprise

The introduction of the EU AI Act creates several challenges within enterprises: **(1) the absence of detailed guidelines** for operationalizing the AI Act in order to orchestrate tasks across the enterprise, **(2) the need to understand and translate legal obligations into technical requirements**, and **(3) understanding what the required skill profiles are and how companies can start preparing** today.

Challenges to Operationalize the EU AI Act

Lack of Clear Guidelines for AI Governance

Lack of an AI Act Governance Framework: Although there is a vast landscape of AI governance frameworks [12], they often do not consider regulatory requirements from the AI Act. The absence of a framework with well defined guidelines to operationalize the AI Act in the enterprise leads to fragmented approaches towards compliance within corporates.

Interdisciplinary Task Orchestration: The EU AI Act is a regulation that affect both technical and non-technical stakholders. As a result, there is a need to orchestrate tasks across stakeholders from different disciplines in the enterprise including legal, infrastructure and AI developer roles.

The Need to Understand and Translate Legal Obligations into Concrete Technical Requirements

Standards Availability: The required AI Harmonized standards are expected to become available by early 2026 [13], giving companies less time than expected to translate and understand these standards. In the absence of harmonized standards, the AI Act by itself is not concrete enough to enable compliance.

Communication Challenges: it is a challenge for technical stakeholders to translate and interpret the legal requirements. On the other hand, it is a challenge for non-technical roles to evaluate technical evidence of compliance.

Uncertainty about Which Skill Profiles are Necessary and Where to Start

Roles and Responsibilities: The implementation of the AI Act requirements demands a close collaboration between technical and non-technical stakeholders. There is uncertainty not only on governance guidelines but also on what skill profiles or clusters of skill profiles are necessary to effectively operationalize the AI Act.

Where to Start? Finally, there is uncertainty about where to start and in which order the obligations should be implemented.

Challenges to Operationalize the EU AI Act in the Enterprise

Responding to These Challenges: The appliedAI Working Groups

To respond to the challenges outlined previously, appliedAI coordinated two parallel working groups with over 15 corporate partners across multiple sessions in 2024.

The first working group, focused on AI Act Governance, interpreted legal expectations for building high-risk AI systems and mapped these requirements to existing ISO/IEC standards. The goal was to create concrete implementation steps for each requirement and outline expected artifacts for oversight.

The second working group, named MLOps under the AI Act, focused on translating these legal obligations into technical requirements. The goal was to develop technical best practices for the implementation steps within each article.

Both working groups worked in tandem and focused on the interaction between these two types of stakeholders, aiming to identify governance challenges and develop best practices to address them.

Working Group	Goal Description	Deliverable	Target Group
AI Act Governance	Create concrete implementation steps for each requirement and outline expected artifacts for oversight.	AI Act governance framework implementation steps for high-risk requiremetns	AI and data governance roles, head of AI governance, compliance roles, DPO, etc.
	High-Level Policy ↕ Documentation & Evidence		
MLOps under the AI Act	Develop technical best practices for the implementation steps for each article.	Technical best practices to implement high-risk requirements	Technical roles: heads of AI, DE, DS, data and ML architects, ML-MLOps engineers

Working Group Outcomes

1. To respond to the lack of clear guidelines for AI governance, we developed the AI Act Governance Pyramid **(Section 3)**.
2. To respond to the delay in the publication of harmonized standards, we developed implementation steps and technical best practices for high-risk AI systems **(Section 4)**.
3. To respond to the uncertainty about which skill profiles are necessary and where to start, appliedAI updated their ML Skill Profiles framework [4] taking the EU AI Act into consideration, and provided a guide about what companies can start doing today to prepare to operationalize the AI Act **(Section 5)**.

"By structuring responsibilities into distinct functional layers—governance, infrastructure, and AI system development—the AI Governance Pyramid fosters vital cross-functional collaboration among our legal, technical, and business teams, ensuring a cohesive approach to operationalizing the EU AI Act."

Dirk Wacker
AI Lead,
Giesecke+Devrient GmbH

Section 3: The AI Act Governance Pyramid

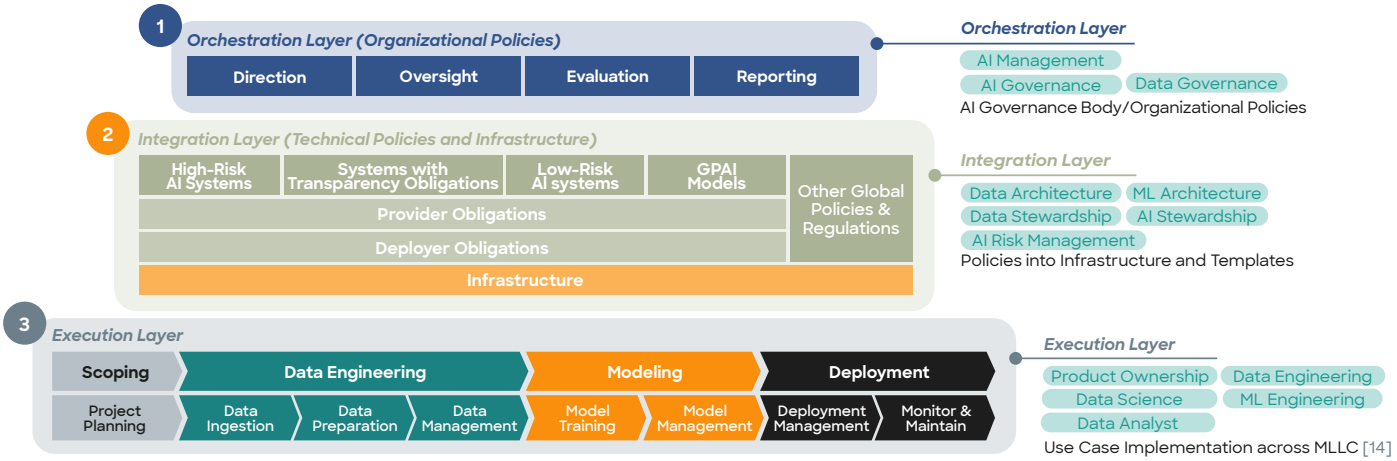
The AI Act Governance Pyramid

In the previous section, we established that the EU AI Act introduces new challenges for enterprises. Among the most pressing challenges is the lack of clear guidance on how to orchestrate legal requirements across the distinct stakeholders involved in their day-to-day practices. Although there are a variety of AI governance frameworks, few are explicitly designed with the AI Act's operational demands in mind [12]. To mitigate this gap, we introduce the AI Act Governance Pyramid.

What is the AI Act Governace Pyramid?

The AI Governance Pyramid is a practical AI governance framework that simplifies the implementation of the AI Act by organizing compliance responsibilities into three functional layers: Orchestration (driven by top level leadership e.g., management and AI governance functions), Integration (driven by Data-AI architecture and Data-AI stewardship functions), and Execution layer (driven by AI project teams, e.g., Data Science functions).

These layers are not related to individual roles but to a set of functions, related skills, or "hats" that may span over multiple stakeholders or be consolidated depending on the size and structure of the organization.



Layers Overview

The pyramidal shape illustrates the **distribution of responsibilities**, **accountability levels of a corporate AI governance structure,** and the **scale of processes and frameworks used** throughout the organization.

Orchestration Layer

The first layer is responsible for harmonizing overall AI governance and management. It comprises high-level stakeholders who are tasked with setting the direction, providing oversight, evaluating, and reporting.

Integration Layer

This intermediate layer encompasses the infrastructure and AI stewards teams responsible for integrating policies into technical infrastructure and documentation templates to be reused across use cases. They also implement necessary controls within an enterprise's technical architecture.

Execution Layer

The final layer consists of business units that apply the defined policies to develop and deploy trustworthy AI use cases. They are responsible for providing evidence of compliance to the relevant stakeholders, before deploying AI systems into production.

By aligning AI responsibilities into functional organizational structures, the AI Act Governance Pyramid enables companies to build workflows that streamline end-to-end accountability, from legal interpretation and policy definition to compliant technical implementation and evidence generation. This framework is a valuable companion to other framework propositions (e.g., NIST AI RMF 1.0) [15], or standards on AI governance (e.g., ISO/IEC 38507:2022) and AI management systems (e.g., ISO/IEC 42001:2023) [16].

The AI Act Governance Pyramid

The AI Act Governance Pyramid Benefits

The proposed pyramid offers a **scalable**, **modular**, and **AI Act-first** approach that prioritizes compliance with the AI Act with a **clear separation of responsibilities** across interdisciplinary teams.

The Pyramid is AI Act-first

- The AI Act governance pyramid is AI Act-first. The AI Act requirements that companies must operationalize were identified and mapped to each layer of the pyramid, as well as relevant international standards.
- It is designed to be particularly helpful for teams starting to extend their governance processes with the AI Act requirements, enabling them to identify potential gaps, align internal processes, and build knowledge.

It offers a clear separation of functions

- The AI Act governance pyramid provides a standardized structure that separates accountability into functional layers, facilitating the operationalization of the AI Act across interdisciplinary teams.
- It offers a separation from broad governance oversight at the top, enablement in the middle, and implementation at the base helping to avoid gaps or duplicated work, optimizing current governance structure alignment, enhancing compliance efforts and streamlining communication.

It provides modularity and flexibility

- The pyramid provides a modular framework with building blocks that can adapt to an organization's set of obligations under the EU AI Act without disrupting the overall governance structure.
- It can be extended with policies from domain-specific regulations to be compatible with existing regulatory requirements.
- It is a flexible framework for all company sizes (function-based, not person-based). A smaller company might have a small team wearing all the hats, while larger companies could have dedicated teams for each layer.

It is easily scalable and enables observability

- It is a framework that enables scalability, as the infrastructure layer acts as a multiplier by reusing tools and policy templates across use cases.
- It enables observability for governance teams to access and review compliance evidence gathered by technical teams. It speeds up time-to-compliance and enables continuous evaluation and improvement.

The AI Act Governance Pyramid Limitations

We developed this framework with a strong focus on operationalizing the EU AI Act. Despite its modular structure allowing for potential extension to incorporate other regulations, such as sector-specific laws or global AI standards, it does not currently contain policies beyond those outlined in the AI Act. This includes areas such as intellectual property, data protection regulations, among others.

Layer 1: Orchestration Layer

Layer Definition

Why is it necessary?

Some companies already rely on existing standards for high-level AI governance and AIMS implementation (e.g., ISO/IEC 38507:2022 and ISO/IEC 42001:2023, respectively) [16]. However, to meet the requirements of the EU AI Act, organizations will have to close the gap between these existing governance practices, which are primarily at the organizational level, and a product-first approach that is required to satisfy the requirements of the AI Act. Without a structured approach, organizations risk inconsistent implementation, unclear accountability, and non-compliance.

What is it?

The Orchestration Layer is the organizational level that manages the overall AI governance process by setting the direction, overseeing and evaluating its implementation, and reporting outcomes to key stakeholders. The strategy and policy decisions in the orchestration layer guide the processes of the Integration and execution layers.

Layer Components and Regulatory AI Act Connection

These are the high-level component and tasks of the proposed Orchestration layer:

Direction

Aligns organizational AI practices with regulatory requirements by proposing a regulatory strategy, policies, processes, and infrastructure to close the gap between existing corporate governance initiatives and AI Act requirements.

Oversight

Assigns clear roles and responsibilities to close compliance gaps with strategically defined AI Act requirements in the regulatory strategy. It aligns legal, technical, and business teams and oversees the design, development, conformity, deployment, and monitoring of AI systems.

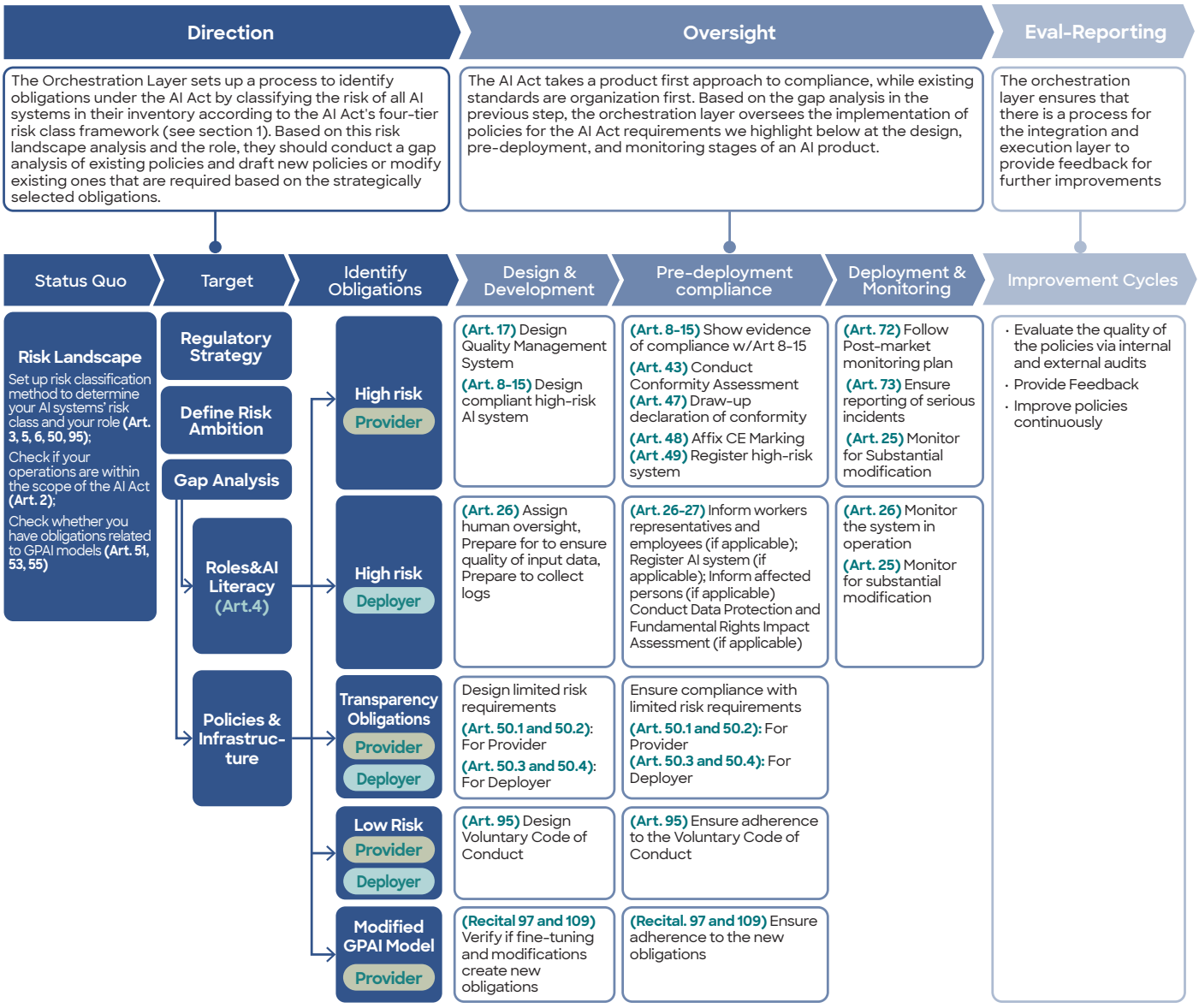
Evaluation & Reporting

Establishes mechanisms to continuously assess and report governance effectiveness across the organization.

Layer 1: Orchestration Layer

The diagram below presents the Orchestration Layer map, a navigational proposal that visualizes key phases and activities aligned with AI Act articles, helping translate regulatory requirements into actionable governance processes.

It is designed to be particularly helpful for teams starting to extend their governance processes with the AI Act requirements, enabling them to identify potential gaps, align internal processes, and build foundational knowledge.



It is essential to remember that at this stage, the orchestration layer should focus solely on legal requirements, strategy, and procedures, while the implementation of these requirements is delegated to the infrastructure and execution layer.

This map is a first proposal and we invite critical feedback for collaborative improvements. It is a living artifact based on best practices and the collaborative improvement from the current working sessions and might differ in new iterations.

Layer 2: Integration Layer

Layer Definition

Why is it necessary?

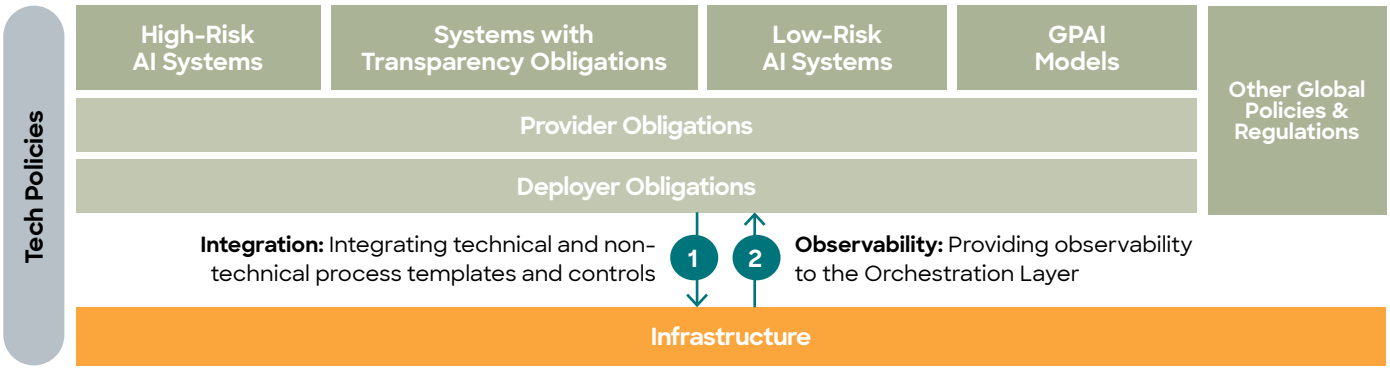
Modern enterprises rely on a variety of MLOps tools, including experiment tracking platforms like MLflow, running on cloud services such as Azure, and data platforms like Delta Lake, among others. While these tools are essential for scaling AI, companies do not necessarily use them in a way that is natively AI Act-compliant. To scale compliant AI development across projects and teams, organizations need a framework that acts as a multiplier by reusing policies, controls, and tools, reducing duplicated efforts, and avoiding the reinvention of governance processes from scratch.

What is it?

The Integration Layer acts as an enabler that integrates the legal requirements into technical infrastructure and documentation templates to be reused across use cases, and provides observability to the stakeholders in the orchestration layer. Compliance templates in the AI infrastructure enable development teams to streamline the AI Act implementation.

Layer Components and Regulatory AI Act Connection

The precise set of policies to integrate in this layer will be determined in the orchestration layer by several factors, including the risk class of AI systems, their role, and the organization's regulatory strategy (see the orchestration layer map), among others. The Integration Layer accomplishes its goals through a compliance-by-design strategy with two key components: infrastructure and policies per risk class and role.



In this layer, organizations make AI Act compliance more scalable by mapping and integrating AI Act requirements into their MLOps tool stack and workflows via compliance controls, providing observability to the orchestration layer.

	1 Integration of Technical and Non-technical Templates and Controls		2 Observability Tools
Explanations	**Integration of standardized and reusable templates** per risk class and roles into the compliance process for non-technical activities, such as use case documentation.	**Integration of AI Act-specific controls, evidence gathering, and infrastructure provisioning templates** into commonly used tools and enterprise infrastructure depending on the AI system risk class and role.	**Provisioning of observability to the Orchestration Layer** to evaluate how project teams implement the requirements and identify potential areas for improvement.
Examples	• Creating use case documentation template per AI system's risk class • Standardizing risk classification and risk management templates • Creating technical documentation templates in MLOps tools for high-risk AI systems	• Data & AI Platform: A data validation step fails if data quality thresholds (defined by governance) aren't met, blocking deployment. • ML Libraries: A model registry is configured to enable traceability and automatically centralize the metadata logs required for technical documentation. • Cloud & On-prem Infrastructure: High-risk use cases use a pre-defined Infrastructure as a Code (IaC) template to deploy necessary services on the cloud.	• Creating KPI dashboards monitoring key compliance activities including post-market monitoring • Updating technical documentation every time the ML pipeline is executed • Setting up the right access control privileges in your organization

Additional activities will depend on the complexity of the company's technology stack, the type of product or service being built, as well as domain-specific and other horizontal regulations.

Layer 3: Execution Layer

Layer Definition

Why is it necessary?

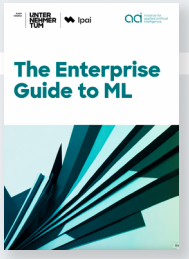
To meet the EU AI Act's requirements, organizations must not just plan or define compliance processes, but they must effectively put them into practice within their AI development workflows. Without a clear execution process, organizations risk gaps leading to non-compliance or misalignment with ethical and legal expectations.

What is it?

The execution layer is the project layer that operationalizes compliance by executing the AI Act requirements throughout the ML lifecycle. In this layer, AI systems are designed, trained, and responsibly deployed in products and services.

Info

The ML Lifecycle is the cyclical process to train, deploy and monitor AI models, using the large amount of information available in your organization. For more details, visit the "Enterprise guide to ML"[14].



Layer Components and Regulatory AI Act Connection

To meet AI Act obligations at the use case level, organizations must identify regulatory requirements, implement the necessary controls, demonstrate compliance, and maintain it over time.

1 Identifying Use Case Requirements

Developers use planning templates to document the use case, assess its risk class and role, and identify AI Act obligations in order to mitigate AI Act-related risks.

2 Implementing Requirements

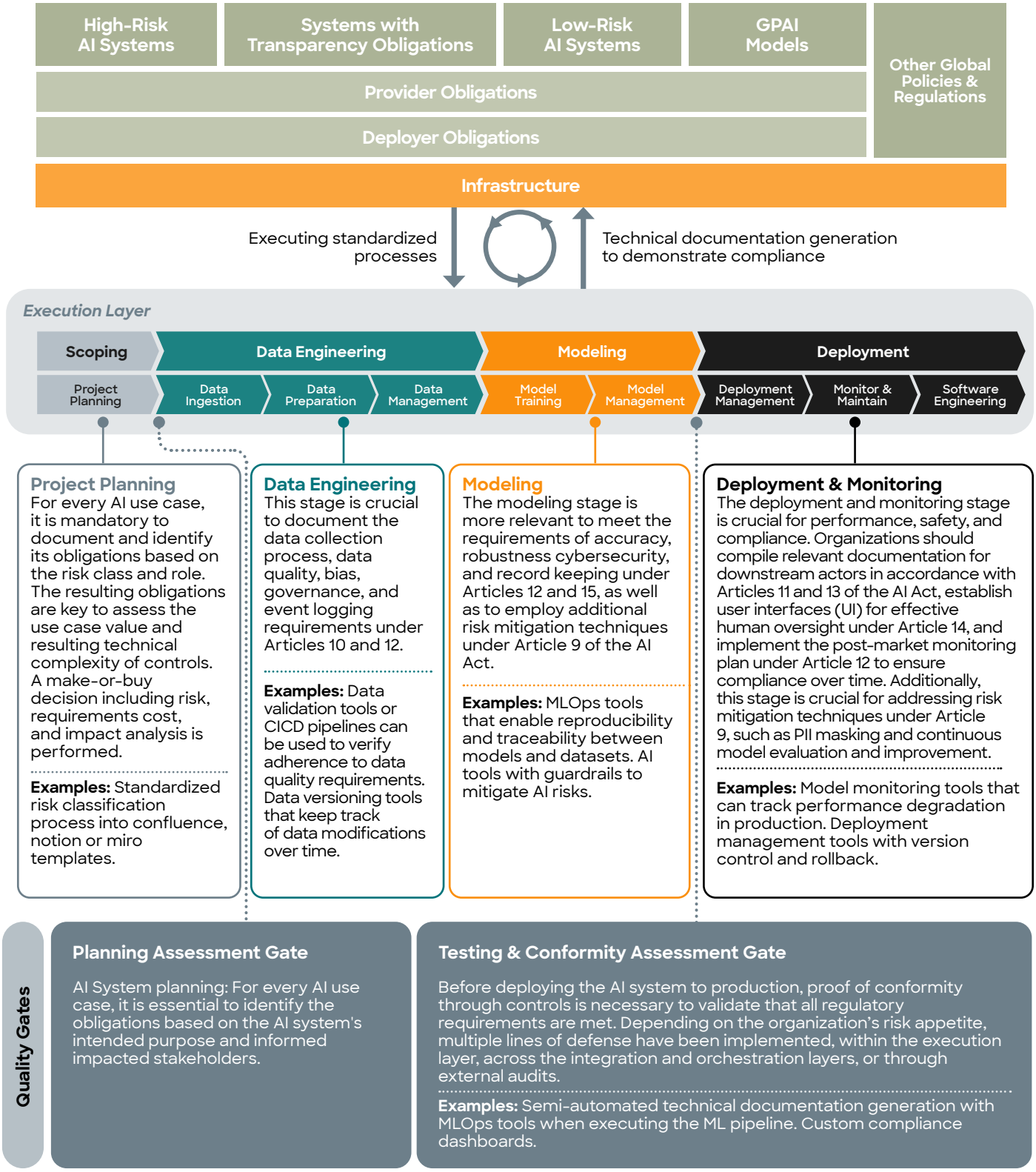
The Execution Layer layer generates compliance evidence across the ML lifecycle by using infrastructure controls and quality gates to ensure requirements are met and documented for conformity assessments and post-market monitoring (when necessary). Each stage of the ML lifecycle requires a quality gate to ensure necessary information is collected.

3 Demonstrating and Maintaining Compliance

Before deployment, proof of conformity is required to validate regulatory compliance. Based on organization's risk appetite, companies can implement different lines of defense across the AI governance layers. Finally, the ongoing compliance of the AI system must be ensured after deployment through proper incident monitoring.

These steps must be embedded into each stage of the ML lifecycle. The diagram on the right outlines how these compliance actions are applied in practice across planning, data engineering, modeling, deployment, and monitoring.

Layer 3: Execution Layer



The diagram above provided an exemplary set of activities and not an exhaustive list of AI Act obligations. In Section 4, we dive deeper into each requirement for providers of AI systems and introduce best practices to operationalize them.

Interactions between Layers

Despite having a clear separation of concerns through layers, the implementation of AI Act Governance requires cyclical interactions between the layers. These interactions ensure that compliance is not siloed, but implemented collaboratively across governance, infrastructure, and development functions. This framework is aligned with the philosophy of a continuous improvement approach from other AI governance standards [16]. These interactions are explained below.

1 Identify Obligations to Implement based on the Risk Landscape

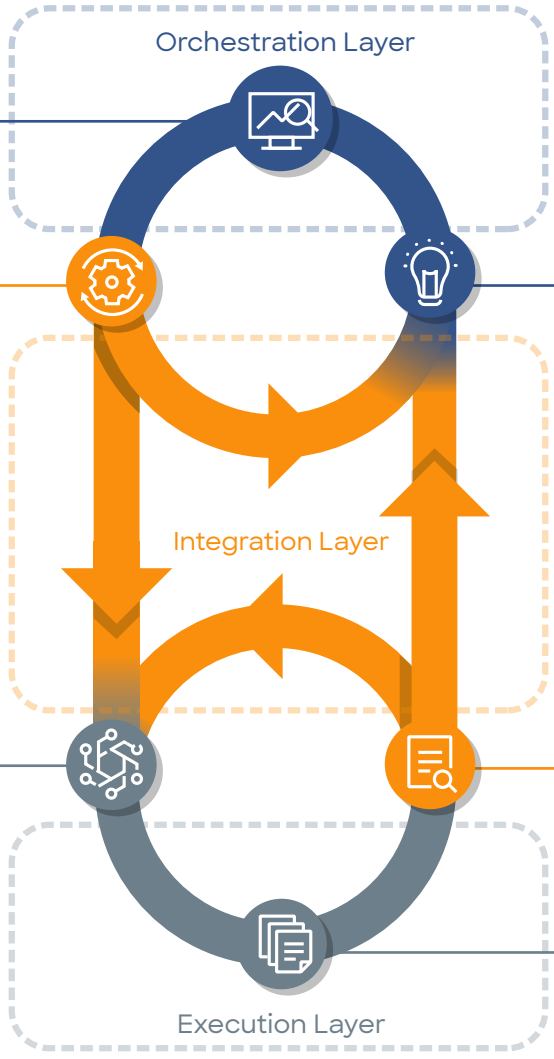
The Orchestration Layer set up a process to identify obligations under the AI Act. Based on this risk landscape analysis, the regulatory strategy outlines the requirements to be implemented. A gap analysis of existing policies and controls is performed to create an implementation plan.

2 Integrate Policies into Processes and Infrastructure

The strategically selected AI Act requirements are translated into technical controls integrated into the Integration Layer. Infrastructure acts as the scaling mechanism. The Integration Layer builds platform capabilities, such as dataset versioning, audit trace logging, or guardrail templates, that the Execution Layer can consume easily, ensuring consistency and traceability across use cases.

4 Execute Policies into AI Use Cases

Developers execute pre-configured infrastructure controls into their AI System across the ML lifecycle, enforcing requirements such as technical documentation or post-market monitoring by design. For example, applying standardized logging templates during development.



3 Evaluation & Reporting

Assess the performance of the AI Act requirements integrated into the infrastructure layer and report potential improvements.

6 Testing and Conformity Assessment

Before deploying the AI system to production, proof of conformity is necessary to validate that all regulatory requirements are met. Depending on the organization's risk appetite, companies can implement different lines of defense not only in the execution layer but also across the Integration and Orchestration Layers, as well as through external audits.

5 Collect Logs and Docs

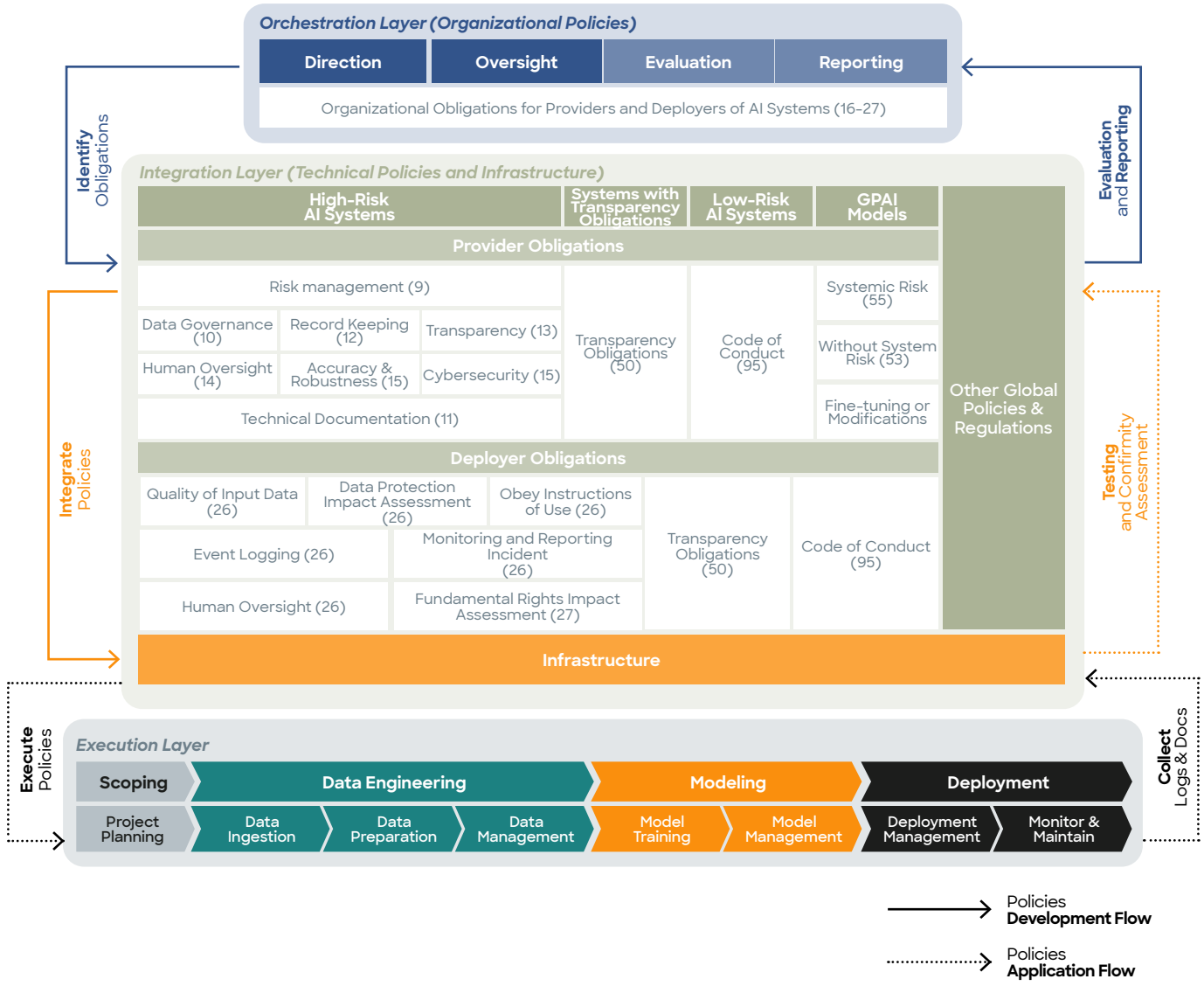
The Execution Layer generates compliance evidence across the ML lifecycle. Each stage of the ML lifecycle requires a quality gate to ensure that the necessary information for that stage is collected and documented.

Putting it All Together

To overcome the challenges in Section 2, we introduced the AI Act Governance Pyramid, a modular, flexible, scalable framework designed to clarify responsibilities, align cross-functional teams, and facilitate compliance across the organization. By promoting a clear separation of concerns and supporting extensibility to additional regulatory requirements, this framework lays the foundation for building enterprise-wide trust in AI, enabling scalability of policies.

Extended AI Act Governance Pyramid

The interaction provided in the previous section, together with the articles related to the AI Act, form this extended yet streamlined pyramid blueprint. Since the policies to be implemented depend on the company's initially assessed risk landscape, we have organized the requirements by risk class and role. This modular structure offers the flexibility to tailor policies to a company's specific governance structure and risk landscape. Additionally, global policies and regulations can be integrated as necessary, depending on the target markets for the AI solutions.



In summary, this section addressed the current gap in actionable guidelines for operationalizing the AI Act and effectively orchestrating AI Act policies among the distinct stakeholders involved across a company.

In the following section, we shift focus from governance to implementation, exploring how to design high-risk AI systems in alignment with the AI Act's technical requirements.

“Having in place a governance pyramid is not just the assurance you are working good and compliant, but integrating the quality and regulation at the very heart of your AI-engine: you can monitor and demonstrate it at any time and any point of the chain. This is what it takes to move in EU AI Act high risk context.”

Philippe Coution
Head of Digital Interaction & Lead AI Quality,
TÜV SÜD

“Throughout the appliedAI´s Working group sessions of the AI Act technical implementation, we had the opportunity to collaborate with other industry experts, exchanging insights and best practices on Trustworthy AI. We anticipate that these contributions will help us to prepare to implement the EU AI Act.”

Cecilia Carbonelli

Senior Principal-Head of Algorithm Concept & Modeling-Responsible AI Tech Lead | Infineon

Section 4: Operationalizing Requirements for Providers of High-Risk AI Systems

Operationalizing Requirements for Providers of High-Risk AI Systems

In the previous section, we focused on the first challenge, the gap in actionable guidelines for operationalizing the AI Act and effectively orchestrating AI Act policies among the distinct stakeholders involved across a company.

In this section, we will focus on addressing the second challenge that enterprises face: the need to understand and translate legal obligations into concrete technical requirements for high-risk AI systems. This challenge discourages enterprises from exploring high-risk AI systems and hinders the development of a shared understanding of compliance between technical and legal stakeholders. Companies might want to start building such systems today and struggle to identify best practices that bring them closer to compliance.

In this section, we:

- Describe our iterative approach to the challenge of unavailable harmonized.
- Provide an overview of existing international ISO/IEC standards we identified as AI-Act relevant.
- Share the working group outcomes for each high-risk requirement along with an illustrative example. For each requirement, we share:
  - Processes to implement high-risk requirements
  - MLOps best practices
  - A mapping of existing ISO/IEC standards to the requirements
  - Grey areas and trade-offs.

Methodology: An Iterative Approach

To address the challenge outlined above, the appliedAI Working Group adopted an iterative approach is designed to operationalize the AI Act high-risk requirements for providers through a structured progression of standards adoption and best practice development. This methodology recognizes that standards development and regulatory implementation occur in parallel, necessitating a flexible and progressive approach to compliance.

The process begins by identifying engineering and governance best practices and mapping the AI Act's requirements to existing ISO/IEC standards where available, providing a foundation based on established industry frameworks. As harmonized standards become available, they are systematically integrated into the framework, replacing or enhancing interim practices as appropriate.

What is our iterative approach?

1) We **develop best practices** to implement high-risk requirements from a practitioner perspective to add concrete detail to otherwise vague AI Act requirements.
2) We **identify and map relevant ISO/IEC standards** to high-risk requirements to enable companies to rely on authoritative guidance before the harmonized standards become available
3) We **integrate the harmonized standards** once they become available.

Why do we follow an iterative approach?

This iterative approach offers significant practical advantages for organizations.

1) **Timely preparation:** Companies can prepare their processes, roles, responsibilities, and checklists, keeping in mind that minor adjustments (a delta) will be needed when the final standards are available.
2) **Smaller effort later:** When the harmonized standards are available, the effort to embed them would be smaller, or is expected to be smaller. This reduces the overall implementation burden.

Overview of Identified ISO/IEC Standards

During the working group, we collectively identified certain ISO/IEC standards that we believe closely map to the requirements of the AI Act. For each high-risk requirement, we map relevant activities from the clauses and sections of these standards.

AI Act Requirement	ISO/IEC Standard
Article 10 **Data and Data Governance**	5259 1-5: Data quality for analytics and machine learning (ML) 24027: Bias in AI systems and AI aided decision making 12791: Treatment of unwanted bias in classification and regression machine learning tasks
Article 12 **Event Logging**	24970: AI system logging
Article 13 **Transparency of provision of information**	6254: Objectives and approaches for explainability and interpretability 12792: Transparency taxonomy of AI systems IEEE P2894: Architectural Framework for Explainable Artificial Intelligence
Article 14 **Human Oversight**	8200: Controllability of automated AI systems 42105: Guidance for human oversight of AI
Article 15 **Accuracy, Robustness, Cybersecurity**	4213: Performance measurement for AI tasks 24029 - 1:4: Robustness of neural networks 29119-11: Guidelines on the testing of AI-based systems 25058: Guidance for quality evaluation of artificial intelligence (AI) systems 27090: Guidance for addressing security threats and failures in AI systems

Disclaimer: This selection reflects our best effort to identify the most relevant standards, many of which map closely to the CEN-CENELEC JTC 21 Work Programme [17]. While we expect many of these standards to be a part of the final harmonized standards that will lend companies a presumption of conformity under the AI Act, the exact set of standards, and activities within them, will only be known later in 2025 or early 2026 [13].

Outcomes

Using our iterative approach, the working group produced the following outcomes per high-risk AI requirement:

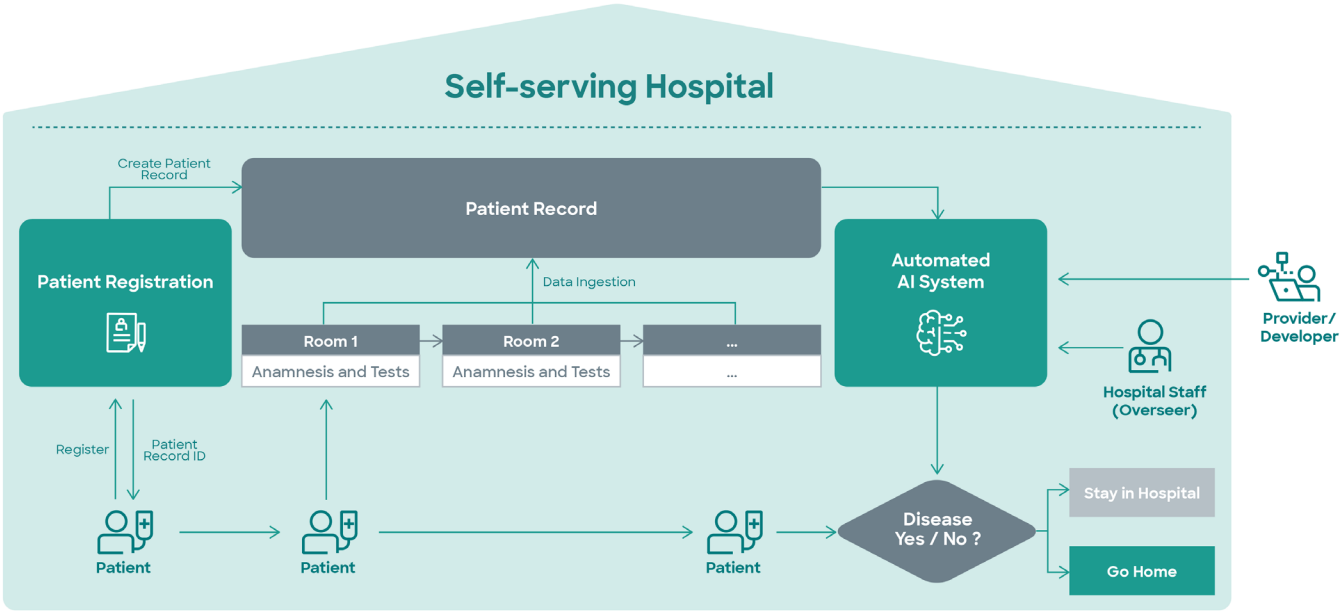
1. Processes and MLOps best practices to operationalize high-risk requirements
2. A mapping of existing ISO/IEC standards to the requirements
3. Grey areas and trade-offs

Note: We exclude Article 9 - Risk Management from this whitepaper, because it was the subject of the appliedAI Working Group in 2023. Please contact the authors if you would need support from appliedAI to implement your Risk Management Framework

Hypothetical Use Case

To make the output of the working group more tangible, we describe a hypothetical use case that will serve as an example for the outcomes for each of the high-risk requirements. We describe this use case naively without attempting to model the context of deployment with great rigour.

We are the provider and the deployer of an application that can be used to predict the presence of cardiovascular disease (CVD) [18] in patients given their examination results. The prediction will be based on the patients characteristics such as age or height, as well as measurements such as blood pressure or cholesterol. The AI system will be used in a "self-serving" hospital, where the patients can register themselves and go through various anamnesis stages independently. At the end, the AI system runs the initial assessment which then determines whether the patient needs to stay in the hospital due to the likely presence of CVD or is allowed to go home again. The medical staff supervises the whole process and should be able to intervene in case of emergency or malfunctioning.



Best Practices for Implementing Requirements of High-Risk AI Systems under the AI Act

Data and Data Governance

Article 10 of the AI Act introduces governance and quality requirements for training, testing and validation datasets used in AI systems. This provision addresses the entire data lifecycle, from collection and preparation to validation and testing, ensuring that AI systems are built on good data foundations.

Requirement Clusters

Broadly speaking, Article 10 can be broken down into 4 key clusters of activities. By way of an example, we show what type of activities from our hypothetical use case introduced at the beginning of this section are.

Article 10 Data and Data Governance		In our hypothetical use case, the following requirements might apply:
1. Data Collection Process	• Data Collection Process • Data Assessment	• We collect data via two sources: first-party collection from patients with consent, and third-party datasets from hospitals with appropriate legal licence and register it in a data catalogue
2. Data Pre-processing	• Process for transformations	• Data pre-processing operations including documenting the process to ensure data labeling consistency across data labelers. e.g. having a class to capture label uncertainty: "Stay in Hospital", "Go Home", and "Borderline" class.
3. Data Quality & Biases	• Data Quality • Examining Biases	• As a measure of representativeness, we check the distribution of samples per gender in training and production datasets and ensure that Samples are almost equally (95%) represented in both training and production dataset
4. Data Provisioning & Documentation	• Documentation via Datasheets • Data provisioning	• We generate a Data Quality Report and Datacards

Mapping the Clusters to the Requirements, WG Outcomes, and Relevant Standards

In the table below, we map the clusters above to the relevant text of Article 10, briefly outline the outcomes of the working group, and highlight international standards that may be relevant.

	Requirement Sourced from the Text of the AI Act: *Training, validation and testing data sets shall be subject to data governance and management practices appropriate for the intended purpose of the high-risk AI system. Those practices shall concern in particular...*	**AI Act Reference**	**Best Practices from the Working Group**	**Additional Reference to Standards**
Data Collection Process	...the data collection processes and the origin of data ...and in the case of personal data, the original purpose of the data collection... ...an assessment of the availability, quantity and suitability of the data sets that are needed ...the formulation of assumptions, in particular with respect to the information that the data are supposed to measure and represent;	Art. 10(2) (b) & Annex IV (2)(b), Art.10 (2)(e), Art.10(2)(d)	We identified a superset of typical metadata information relevant for companies to document data acquisition process and original purpose of collection.	ISO 5259-3
Data Pre-processing	data-preparation processing operations, including annotation, labelling, cleaning, updating enrichment and aggregation.	Art. 10(2)(c) & Annex IV(2) (b)	We identified some best practices to document the data-preprocessing process, with particular focus on data labelling consistency.	ISO 5259-3
Data Quality & Biases	examination in view of possible biases that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operations; appropriate measures to detect, prevent and mitigate possible biases identified	Art. (10)(2)(f) & (g)	We identified a superset of data quality best practices along the ML lifecycle that companies can apply to their training, validation and test datasets.	Review recommendations in ISO/IEC TR 24027: 2023 and ISO/IEC DTS 12791: 2024
	Training, validation and testing data sets should be • be relevant, • be sufficiently representative, • be, to the best extent possible, • free of errors and • complete in view of the intended purpose, • have the appropriate statistical properties the identification of relevant data gaps or shortcomings that prevent compliance with this Regulation, and how those gaps and shortcomings can be addressed.	Art. 10(3) and Art. 10(2)(e) and (h)		ISO 5259-2
Data Provisioning & Documentation	...information about their provenance, scope and main characteristics; how the data was obtained and selected... ..data requirements in terms of datasheets describing the training methodologies and techniques and the training data sets used...	A. 10 (2) (b), Annex IV (2) (b)&(d)	We ideated some typical MLOps best practices supporting provenance about how the data was obtained and documenting data quality requirements.	ISO 5259-3

Best Practices for the Implementation Steps

In the cards below, we detail the best practices developed by the working group for each cluster of activities that are relevant to Article 10.

1. Data Collection Process

The working group identified best practices for data collection processes, including: a superset of typical metadata elements to document the origin of data, recommended steps to establish a data collection process, and the original purpose of collection.

Best Practices from the Working Group

- **Register the Origin of Data**
  - Data Source
  - Data Format
  - Data Owner
  - Data Value and Purpose of collection
  - Data Variety (Structured, semi, unstructured)
  - Data Velocity: Static, batch, stream, etc
  - Data Volume: size of dataset
  - Data Veracity: Data quality requirements and expected deviations
- **Establish Data Collection Process:**
  - Establish process with roles, responsibilities and skills for data collection
  - Document data requirements early in the process through a canvas
  - Based on your requirements, assess the if availability (appropriate legal permission: document lawful basis/ contract for usage of data), quantity (data volume) and suitability (data value) are enough for your AI use case.
  - Register data sources metadata in a data card or in a central data catalog to be consumed by other teams.
  - Automate data registration process when possible
- **Original Purpose of Data Collection (in Case of Personal Data)**
  - Embed process to identify personal data for all the data products
  - Define process to treat dataset or columns with personal data: tag relevant columns, anonymization, layered-architecture with right access control to personal data layer, etc
  - Document the results and processing activities as per GDPR requirements

AI Act Reference

Art.10(2)(b) & Annex IV (2)(b), Art.10 (2)(e), Art.10(2)(d)

Additional Reference to Standards

ISO 5259-3

2. Data Pre-processing

We identified best practices to document the data preprocessing process, with particular focus on data labelling consistency.

Best Practices from the Working Group

- **Assess and document what data transformations are required, e.g., labelling, data augmentation etc.**
- **Document methodology for applied transformations**
- **Evaluate whether applied transformations satisfy data quality criteria**
- **Identify roles and responsibilities for oversight**
- **Establish a data labelling process to ensure consistency:**
  - Have multiple (at least two) labelers label same example.
  - When there is disagreement, the ML Engineer, subject matter expert and / or labelers should discuss the definition of "y" to reach agreements.
  - Document and write down the agreements.
  - Iterate until it is hard to significantly increase agreement.
  - If there is still disagreements, discuss again and update the new agreements and re-label accordingly.
  - If the input data does not contain enough information to assign a label, consider changing the input data.
  - Have a class to capture label uncertainty. E.g.: Scratch, Non-Scratch and Borderline class. This will improve labelling consistency.
  - Merge classes when they are too hard differentiate.

AI Act Reference

Art. 10(2)(c) & Annex IV(2)(b)

Additional Reference to Standards

ISO 5259-3

Best Practices from the Working Group

- We identified a superset of data quality best practices along the ML lifecycle that companies can apply to their training, validation and test datasets to ensure that datasets are:

1) relevant;
2) sufficiently representative;
3) to the best extent as possible free of errors;
4) complete in view of the intended purpose; and
5) have the appropriate statistical properties.

Lifecycle Phase	Data Engineering			Modelling		Deployment		
	Data Ingestion	Data Preparation	Data Management	Model Training	Model Management	Deployment anagement	Monitoring	Feedback Loops
Relevance	EDA and drop irrelevant records			Determine relevant features based on feature importance Ablative testing: remove certain components in a controlled setting to investigate all possible outcomes of system failure Set a threshold to remove features with less relevance Use a-posteriori local explainability methods (e.g., SHAP, LIME)		Unit testing: Automated unit tests based on set thresholds. Monitor label drift, data drift, feature drift and concept shift Dashboard with metrics to make decisions about the data		
Statistical Properties	Measure class imbalance for bias detection Stratification to equally represent demographic groups across datasets Be careful when imputing missing data as it can change the statistical properties			Calculate differentiate or disparate impact ratio: Compare model predictions across demographic groups to detect disparities. Fairness Assessment → Conduct perturbation tests by slightly modifying demographic-related inputs to see if predictions change unfairly. Apply fairness metrics (e.g., equalized odds, disparate impact), experiment with multiple training samples, and Discrimination Detection → Check if the output is skewed or disproportionately favors one group		Unit test to measure fairness Real-Time Bias Auditing → Monitor incoming data and recalculate fairness metrics (e.g., disparate impact ratio) to detect issues post-deployment and raise alerts. Drift Detection → Track whether demographic-related shifts in the data affect model performance unfairly. Guardrail: detect some biases in production and then fix it or give a message to the user Experiment with training models with multiple samples and look for model drift: → Assess whether model drift introduces new biases over time. Include user feedback (thumbs up) to validate if minority classes are well-represented (they dont get biased predictions?)		
Represen-tativeness	Data collection from diverse sources to better represent intended population in production EDA for Distribution Verification, ensuring train, test and validation sets distributions represent production data (e.g. using Great Expectations)			Measure model performance against train, test, validation set Curated Test Sets: Use curated domain-specific test sets to assess performance and identify failure points qualitatively Threshold-Based Unit Tests → Automate unit tests to flag performance deviations beyond set thresholds.		Monitoring for Data Drifts: Measure whether the training data population still represents real-world data post-deployment.		
Complete-ness	Null Value Identification Imputation method (For Numerical Data): - Mean imputation: Replace missing values with the mean of the column. - Median imputation: Replace missing values with the median – more robust to outliers. - Mode imputation: Use the most frequent value – often used for categorical but can be used for discrete numerical data. - Constant value imputation: Fill with a fixed value like 0, 1. Advanced Imputation Methods: K-Nearest Neighbors (KNN) imputation, iterative Imputer, autoencoder type architecture to predict missing values Imputation methods for categorical data: mode imputation, constant imputation. Mode imputation: Fill missing values with the most frequent category. Constant imputation: Replace with a placeholder (e.g., 'Missing', 'Unknown').			Study model uncertainty, if it is too high, then go back to the data engineering stage and take an action: get more data, etc. Performance analysis: Feature or datapoint ablation, dropout, demonstration that accuracy degrades gracefully		Unit testing: Build tests for data completeness using multiple measures on crucial features to check for null-values. Input data sanitisation and data validation Pipeline monitoring: Monitor and alert on failed pipelines and empty transfers Monitor data flow and data availability		
Free of Errors	Detect outliers and measure errors or deviation Schema ValiDation: Verify data types, ranges, if null values are possible, etc Profiling data set to automatically determine expected values Appropriate outlier treatment			Perform counterfactual and treatment analysis to identify erroneous records		Data validation based on expected values Monitor model health to detect data problems Monitor data drifts		

AI Act Reference

Art. (10)(2)(f)&(g)

Additional Reference to Standards

Review recommendations in ISO/IEC TR 24027: 2023 and ISO/IEC DTS 12791: 2024

4. Data Provisioning & Documentation

We ideated some typical MLOps best practices supporting provenance about how the data was obtained and documenting data quality requirements.

Best Practices from the Working Group

- Document which data quality measures have been selected
- Define acceptance criteria for each measure
- Set up a data quality reporting framework and document reports

AI Act Reference

Art. 10(3) and Art. 10(2)(e)and (h)

Additional Reference to Standards

ISO 5259-2, pg. 25 and 26

Best Practices from the Working Group

- Data lineage and versioning between data sources, transformations and features used to train and test ML models
- Perform and visualise EDA
- Document information about data by rely on instruments such as datasheets, data-cards, or data catalogues
- Document evidence that data supplied to AI/ML systems meets data quality criteria

AI Act Reference

Art. 10 (2) (b), Annex IV (2)(b)&(d)

Additional Reference to Standards

ISO 5259-3, pg. 7-8

Option Questions and Tradeoffs for Article 10

Open Questions

Is the data governance process per use case or across use cases?	How to justify data quality metrics?	How to integrate these requirements with the requirements of other laws, for example the GDPR or the Data Act
How should downstream actors who use general purpose AI models meet these requirements?	How much data access can you give to external auditors?	How should the responsibility for data quality be split between provider and deployer

Tradeoffs

If there are conflicting sectoral standards for data governance and quality, which standard prevails?
Better data quality can come at the cost of privacy

Record-Keeping

Article 12 of the EU AI Act establishes record-keeping (event logging) requirements for high-risk AI systems. It mandates that these systems must have logging capabilities that enable appropriate traceability of their functioning.

Requirement Clusters

Broadly speaking, Article 12 can be broken down into two key clusters of activities. By way of an example, we show what type of activities from our hypothetical use case are relevant to each cluster.

Article 12
Record-Keeping

		In our hypothetical use case, the following situations may apply:
1. Log Identification	• Log events to identify situations that may result in AI system presenting a risk • Log events that might result in a substantial modification	• If the patient is misclassified and given the wrong diagnosis, it may harm their health. How sure can we be about the AI system's output? Hence we log the confidence interval for prediction. • Replacing the current algorithm with a different machine learning approach (e.g. from traditional ML to deep learning) might result in a substantial modification. Hence we log the timestamp of model updates/replacements.
2. Post-Market Monitoring	• Monitor AI system in production by provider • Monitor AI system in production by deployer	• In the event there is a change on data statistics per cluster of patients, we log the input data from the patients to monitor data-drift over time.

Mapping the Clusters to the Requirements, WG Outcomes, and Relevant Standards

In the table below, we map the clusters above to the relevant text of Article 12, briefly outline the outcomes of the working group, and highlight international standards that may be relevant.

	Requirement Sourced from the Text of the AI Act: *...In order to ensure a level of traceability of the functioning of a high-risk AI system that is appropriate to the intended purpose of the system, logging capabilities shall enable the recording of events relevant for....*	**AI Act Reference**	**Best Practices from the Working Group**	**Standards**
Log identification	Identifying situations that may result in the high-risk AI system presenting a risk within the meaning of Art. 79(1) or in a substantial modification;	Art. 12(2)(a)	We developed a risk management framework for log identification.	ISO/IEC AWI 24970Artificial intelligence – AI system logging (Currently unavailable)
Monitoring	Facilitating the post-market monitoring referred to in Art. 72 and; Monitoring the operation of high-risk AI systems referred to in Art. 26(5).	Art. 12(2)(b) and (c)	We identified a common superset of logs that companies can use to monitor the system in production.	ISO/IEC AWI 24970Artificial intelligence – AI system logging (Currently unavailable)

Best Practices for the Implementation Steps

In the cards below, we detail the best practices developed by the working group for each cluster of activities that are relevant to Article 12.

1. Log Identification

We developed a risk management framework for log identification, designed to identify situations where an AI system may present a risk or undergo a substantial modification. It consists of the following steps:

Best Practices from the Working Group: *Risk Management Based Process for Log Identification*

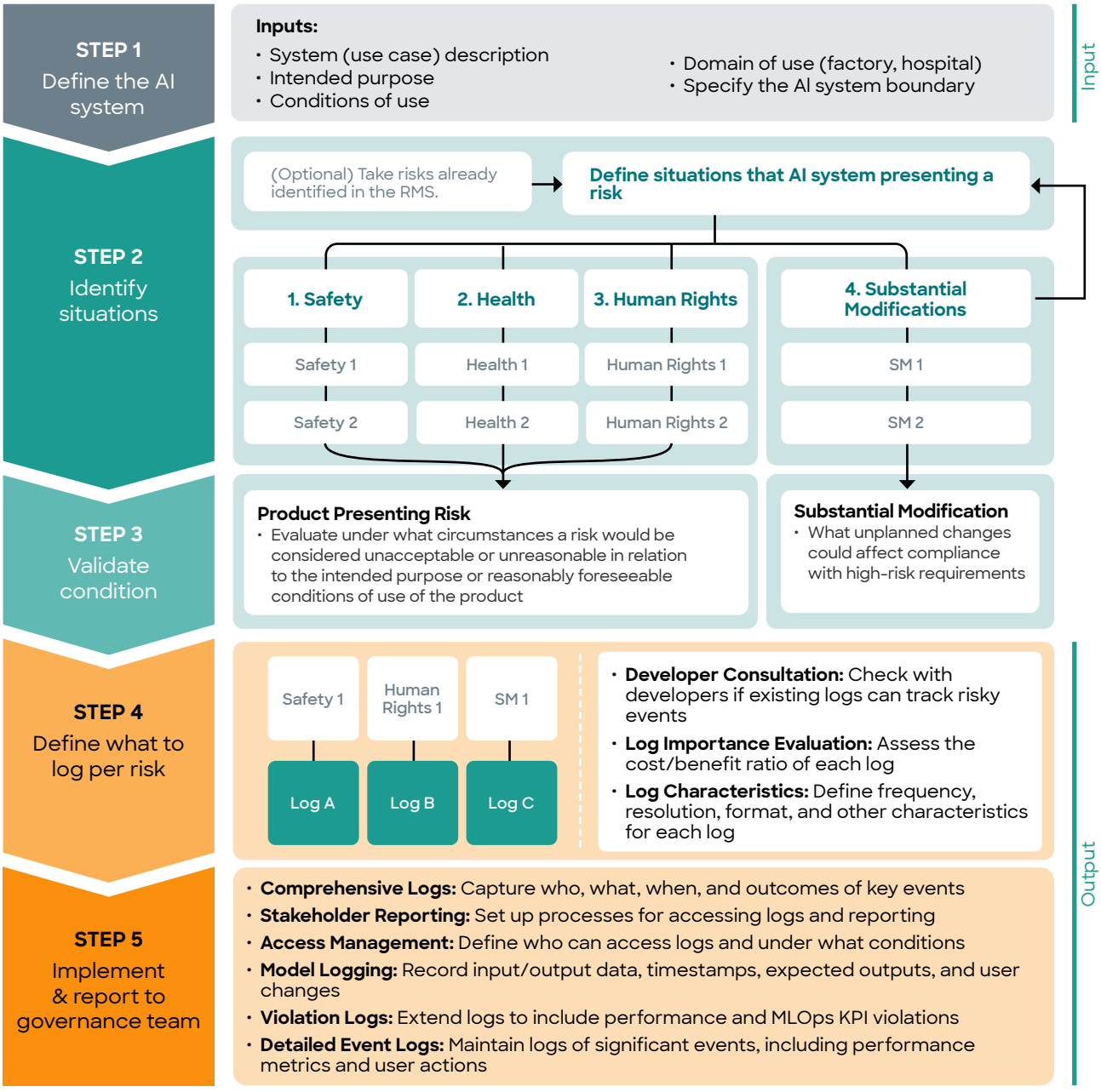
Step 1 Define the AI system.

Step 2 Identify situations where the product might pose risk to health, safety and fundamental rights or might result in a modification.

Step 3 Validate under what conditions these risks might be considered unacceptable.

Step 4 Validate under what circumstances a modification would be considered substantial.

Step 5 Document the events that should be logged to ensure these circumstances are traceable.



AI Act Reference

Art.12(2)(a)

Additional Reference to Standards

ISO/IEC AWI 24970 Artificial intelligence – AI system logging (Currently unavailable)

2. Post-Market Monitoring

Best Practices from the Working Group: *MLOps Best Practices for Post-market Monitoring*

Additionally, we identified a common set of logs that companies can use to:

a) Assess the performance of AI systems throughout their lifetime
b) Evaluate the continuous compliance of AI systems

Lifecycle Phase	Planning	Data Engineering			Modelling		Deployment		
		Data Ingestion	Data Preparation	Data Management	Model Training	Model Management	Deployment Management	Monitoring	Feedback Loops
High-Risk	Identify use case specific logs based on the Risk management based process for log identification	Log expected data statistics to validate incoming data points (e.g. bias, data usage, distribution, missing values rate) Profile datasets to be used for data validation in production Log data and code versioning Record data lineage Log data dependency graph			Register model architecture in model registry Log checkpoint files When possible vide confidence intervals Log into Model registry, experiment in production containing: model version, hyper parameters, training logs, performance (cross validation) Log expected performance per relevant category: Error analysis, counterfactuals		Log the events identified during aAI WG process Preparation of manual reports in the event of an incident Log the events identified during aAI WG process Preparation of manual reports in the event of an incident Log model explanation (when technically possible) Log data drift and any deviations of incoming data points per category Log model drift or deviations from expected performance per category Log model statistics e.g. fairness metric Log input data and corresponding predictions in evaluation store and user feedback (rejected, accepted, etc) Log when re-training takes place and update model registry with new model in production for traceability		

AI Act Reference

Art.12(2)(b) and (c)

Additional Reference to Standards

ISO/IEC AWI 24970Artificial intelligence – AI system logging (Currently unavailable)

Option Questions and Tradeoffs for Article 12

Open Questions

It is unclear how to distribute the logging responsibilities along the value chain. Example: can we access logs from upstream model providers?	How should the logs be stored for subsequent auditing if neccesary?
How will the requirement account for AI system with high frequency of logs.	What is the minimum set of MLOps best practices to log for post-market monitoring?
When would a risk be considered unaccepabtle or unreasonable in relation to the intended purpose?	Provide more guidance on what is a substantial modification.

Tradeoffs

Logging as much as possible might be an organizational and engineering overhead.	The more logs, the more log management (accessibility, visualization) necessary
Non-technical roles (such as PMs) require upskilling to formulate what needs to be logged and in what format	The higher the AI system frequence of logs, the higher the burden of this article

Transparency and Instructions of Use

Article 13 of the EU AI Act mandates transparency through instructions of use for high-risk AI systems to enable downstream actors to use an AI system appropriately. While certain elements of these instructions can be derived from the existing technical documentation for compliance under the Act, others need to be tailored to the deployment context, capabilities and role of downstream actors.

Requirement Clusters

Article 13 can be broken down into three key clusters of activities. By way of an example, we show what type of activities from our hypothetical use case are relevant to each cluster.

Article 13
Transparency & Instructions of Use

		In our hypothetical use case the following practices might apply :
1. Performance, Capabilities and Limitations	• Instructions of use elements related to the AI system performance, capabilities, and limitations, that would allow the deployer or human overseer to understand the system	• Explaining to hospital staff the underlying logic of the model and when to ignore classifications
2. Compliance Information from Other Articles	• Information elements that will allow a downstream actor to better understand and use the system that can be drawn from existing technical documentation under Art. 11 and Annex IV	• Explaining to hospital staff how the UI for human oversight functions
3. Interpretability (When required)	• Applying interpretability or explainability engineering practices	• We identify feature relevance using tools like Layer-wise Relevance Propagation (LRP) and SHAP and explain these results in an accessible manner to hospital staff

Mapping the Clusters to the Requirements, WG Outcomes, and Relevant Standards

In the table below, we map the clusters above to the relevant text of Article 13, briefly outline the outcomes of the working group, and highlight international standards that may be relevant.

	Requirement Sourced from the Text of the AI Act	AI Act Reference	Best Practices from the Working Group	Standards
Performance, Capabilities and Limitations	High-risk AI systems shall be accompanied by instructions for use...shall contain at least the following information..... (b)the characteristics, capabilities and limitations of performance of the high-risk AI system, including: (iv)where applicable, the technical capabilities and characteristics of the high-risk AI system to provide information that is relevant to explain its output; (v)when appropriate, its performance regarding specific vpersons or groups of persons on which the system is intended to be used; (vii)where applicable, information to enable deployers to interpret the output of the high-risk AI system and use it appropriately; ..... (e) the computational and hardware resources needed, the expected lifetime of the high-risk AI system and any necessary maintenance and care measures, including their frequency, to ensure the proper functioning of that AI system, including as regards software updates; ..... [Note: By and large, the documentation required for Article 13 overlaps with the technical documentation required for high-risk AI systems. For the sake of brevity, we have only included the novel requirements in this cell]	Art. 13(3)	We developed an enterprise process to help prepare instructions of use required by the AI Act for a given use case.	ISO/IEC 12792
Compliance Information from Other Articles	..... (a) the identity and the contact details of the provider and, where applicable, of its authorised representative; (b)the characteristics, capabilities and limitations of performance of the high-risk AI system, including: (i) its intended purpose; (ii) the level of accuracy, including its metrics, robustness and cybersecurity referred to in Article 15 against which the high-risk AI system has been tested and validated and which can be expected, and any known and foreseeable circumstances that may have an impact on that expected level of accuracy, robustness and cybersecurity; (iii) any known or foreseeable circumstance, related to the use of the high-risk AI system in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, which may lead to risks to the health and safety or fundamental rights referred to in Article 9(2); (vi) when appropriate, specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets used, taking into account the intended purpose of the high-risk AI system; ..... (c) the changes to the high-risk AI system and its performance which have been predetermined by the provider at the moment of the initial conformity assessment, if any; (d) the human oversight measures referred to in Article 14, including the technical measures put in place to facilitate the interpretation of the outputs of the high-risk AI systems by the deployers; ..... (f) where relevant, a description of the mechanisms included within the high-risk AI system that allows deployers to properly collect, store and interpret the logs in accordance with Article 12.	Art.10(2)(c) & Annex IV(2)(b)	After applying all high-risk requiremnts, ensure that resulting information is re-formatted and included in the instructions of use.	ISO 5259-3, pg. 10-11
Interpretability	High-risk AI systems shall be designed and developed in such a way as to ensure that their operation is sufficiently transparent to enable deployers to interpret a system's output and use it appropriately. An appropriate type and degree of transparency shall be ensured with a view to achieving compliance with the relevant obligations of the provider and deployer set out in Section 3.	Art. 13(1)	• Identify if the model inherently interpretable or is it a black box model; • Identify what tools and methods for interpretability/ explainability can be used during development and deployment based on the state-of-the-art.	ISO/IEC 6254 (Standard unavailable)

Best Practices for the Implementation Steps

In the cards below, we detail the best practices developed by the working group for each cluster of activities that are relevant to Article 13.

1. Performance, Capabilities and Limitations

Best Practices from the Working Group : *Process to Prepare Instructions of Use*

ISO/IEC 12792 contains a list of elements that are useful for enhancing transparency. appliedAI identified the sections that are relevant to the AI Act. We created an enterprise process to help prepare instructions of use for a given use case consisting of the falling steps: [Note that Steps 2.1 and 4 are optional, and depend on the use case and enterprise needs]

Step 1 Stakeholder needs have been identified and documented

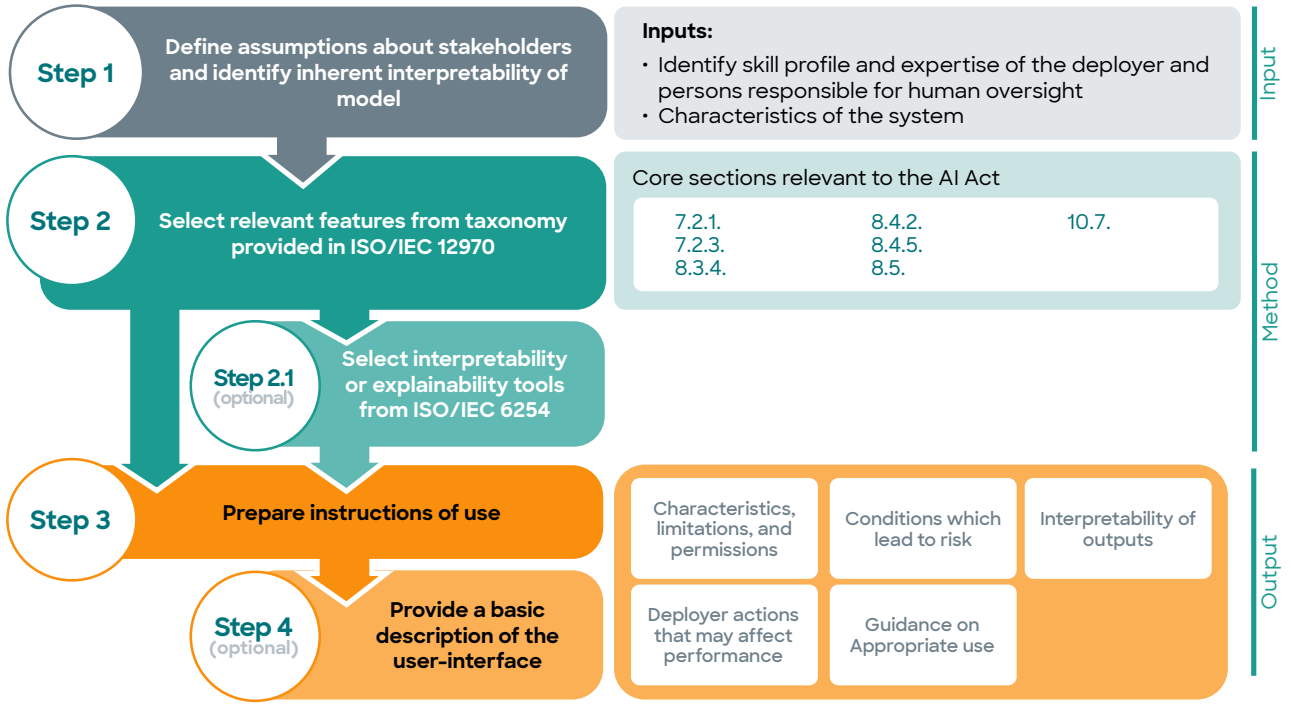
Step 2 The characteristics, capabilities and limitations of performance have been described according to the information elements contained in the standard (The list below is indicative):

- The characteristics of the system (See Section 8.5, pg. 20-22)
- The capabilities and limitations of the system (See Section 8.4.3 and 8.4.4, pg 19)
- The intended and precluded uses (See Section 8.4.2, 8.4.5, 8.4.6)
- The conditions that may lead AI system to present a risk (See Section 7.2.1, 7.2.3, 8.3.4, 10.7)
- The measures to facilitate interpretation of outputs (See Section 8.5.2, 9.4.6, 7.2.3)
- How to make correct choices of the system and use it appropriately (See Section 8.5.7)

Step 2.1 (optional) Consider if interpretability or explainability tools will be used

Step 3 Prepare instructions of use

Step 4 (optional) Consider how these instructions will be delivered.



AI Act Reference

Art. 13(3)

Additional Reference to Standards

ISO/IEC 12792

2. Compliance Information from other articles

Best Practices from the Working Group

Ensure that information is re-formatted to provide information rather than as a compliation of evidence

AI Act Reference

Art. 10(2)(c) & Annex IV(2)(b)

Additional Reference to Standards

ISO 5259-3, pg. 10-11

3. Interpretability

Best Practices from the Working Group

- Identify if the model inherently interpretable or is it a black box model
- Identify what tools and methods for interpretability/explainability can be used during development and deployment based on the state-of-the-art.

  Note: Importantly, the AI Act does not mandate the use of explainability or interpretability techniques. However, the use of these techniques would certainly support the creation of effective instructions of use.

AI Act Reference

Art. 13(1)

Additional Reference to Standards

ISO/IEC 6254 (Standard unavailable)

Option Questions and Tradeoffs for Article 13

Open Questions

Extent to which explainability techniques need to be applied is unclear	How to justify the "appropriate level" of transparency
Template for instructions of use	

Tradeoffs

The less technically adept the user, the more effort is required to fulfil the transparency obligations	Transparency vs. security

Human Oversight

Article 14 of the EU AI Act outlines essential human oversight requirements for high-risk AI systems. It specifies that human operators must be able to properly understand system capabilities and limitations, avoid automation bias, monitor operations for anomalies, and, crucially, have the ability to intervene in operations to address the anomalies. This includes powers to not use, override, or reverse AI outputs, and to interrupt the system through mechanisms like 'stop' buttons that bring the system to a safe state.

Requirement Clusters

Article 14 can be broken down into three key clusters of activities. By way of an example, we show what type of activities from our hypothetical use case are relevant to each cluster.

Article 14 Human Oversight		In our hypothetical use case the following practices might apply :
1. Observability	• Enable human operator to monitor operations	• A UI allows the hospital staff to monitor the confidence interval of the models prediction in an accessible manner
2. Upskilling Human Overseer	• Avoid automation bias • Provide Instructions of Use	• A UI allows the hospital staff to review the models classification and the edit or otherwise change it.
3. Controllability	• Enable human operator to control operations	• All staff that rely on model results are trained to identify situations where the model might not perform consistently

Mapping the Clusters to the Requirements, WG Outcomes, and Relevant Standards

In the table below, we map the clusters above to the relevant text of Article 14, briefly outline the outcomes of the working group, and highlight international standards that may be relevant.

	Requirement Sourced from the Text of the AI Act The oversight measures should allow/enable the human operator ...	**AI Act Reference**	**Best Practices from the Working Group**	**Standards**
Observability	(a) to properly understand the relevant capacities and limitations of the high-risk AI system and be able to duly monitor its operation, including in view of detecting and addressing anomalies, dysfunctions and unexpected performance;	Art. 14(4)(a)	We identified a superset of typical activities that the human overseer could use to monitor AI system's operation in order to detect anomalies, disfunctions and unexpected performance. Depending of the use case they can select a subset of them and enrich them use case specific observability operations.	ISO/IEC 8200, Section 5, 6 and 7 ISO/IEC CD 42105 (Currently unavilable)
Upskilling Human Overseer	(b) to remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system (automation bias), in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons;	Art. 14(4)(b)	We identified some recommendations to upkskill the human overseer.	ISO 5259-3, pg. 10-11
	(c) to correctly interpret the high-risk AI system's output, taking into account, for example, the interpretation tools and methods available;	Art. 14(4)(c)		ISO 12792
Controllability	(d) to decide, in any particular situation, not to use the high-risk AI system or to otherwise disregard, override or reverse the output of the high-risk AI system; (e) to intervene in the operation of the high-risk AI system or interrupt the system through a 'stop' button or a similar procedure that allows the system to come to a halt in a safe state.	Art. 14(4)(d) and (e)	We identifed a superset of typical potential activities that the human overseer could use to control an AI system's operation in order to address anomalies, disfunctions and unexpected performance. Depending of the use case they can select a subset of them and enrich them use case specific controllability operations.	ISO IEC 8200, Section 7 and 8 ISO/IEC CD 42105 (Currently unavilable)

Best Practices for the Implementation Steps

In the cards below, we detail the best practices developed by the working group for each cluster of activities that are relevant to Article 14.

1. Observability

Best Practices from the Working Group

Certain monitoring and control operations are common to most ML models, while others depend on the specific use case. We identified those observability elements that are common and primarily relevant to a human overseer with technical skills and mapped them to the ML lifecycle.

Lifecycle Phase	Data Ingestion	Data Preparation	Data Management	Model Training	Model Management	Deployment Management	Monitoring	Feedback Loops
	Data Engineering			Modelling		Deployment		
Observability Catalogue: Superset of potential activities that the human overseer could use to monitor AI system's operation in order to detect anomalies, disfunctions and unexpected performance.	Ingesting data: acquiring new data/labels Preprocessing input data			Training model Training logs, performance Registering model		Model registered Model deployed Running inference Model stopped Low confidence Drift detected Imbalanced performance Retraining triggered Viewing model accuracy & feature importance		

AI Act Reference

Art. 14(4)(a)

Additional Reference to Standards

ISO/IEC 8200, Section 5, 6 and 7
ISO/IE CD 42105 (Currently unavilable)

2. Upskilling Human Overseer

Best Practices from the Working Group

- Identify if it is neccesary to provide training to human overseer
- Identify what UI based features can prevent the overseer from over-relying on the AI system, including, e.g.
  - Warning messages to human overseer
  - UI Design to not opt-in to the AI system result by default
- Identify what organizational measures can prevent the overseer from over-relying on the AI system, including, e.g., rotating human overseer at appropriate times

AI Act Reference

Art. 14(4)(b)

Best Practices from the Working Group

Ensure that the instructions of use prepared under Article 13 also account for the knowledge and skill of the human overseer.

AI Act Reference

Art. 14(4)(c)

Additional Reference to Standards

ISO 12792

3. Controlability

Best Practices from the Working Group

Certain controllability operations are common to most ML models, while others depend on the specific use case. We identified those controllability elements that are common and primarily relevant to a human overseer with technical skills and mapped them to the ML lifecycle.

Lifecycle Phase	Data Ingestion	Data Preparation	Data Management	Model Training	Model Management	Deployment Management	Monitoring	Feedback Loops
	Data Engineering			Modelling		Deployment		
Controllability Catalogue: Superset of potential activities that the human overseer could use to monitor AI system's operation in order to address anomalies, disfunctions and unexpected performance.	Acquire new data Integrate production data Choose different data version			Select training parameters (Experiment tracker e.g. MLFlow) Train model: perform hyperparameter optimization Choose different model version Use different feature set		- Rollback - Disable endpoint (Stop button) - Retrain - Manually correct data predictions - Stop Button - Disregard output - Override output - Reverse the output		

AI Act Reference

Art. 14(4)(d) and (e)

Additional Reference to Standards

ISO IEC 8200, Section 7 and 8
ISO/IEC CD 42105 (Currently unavilable)

Option Questions and Tradeoffs for Article 14

Open Questions

How frequently should the operator monitor the AI Systems?

To what extent should the oversight measures correspond to the skill of the human operator?

How should the responsibilities be split between the provider and the deployer?

Does Article 14(4)(c) introduce an additional interpretability requirement or is it downstream of the transparency requirement in Article 13?

Oversight might be required by more than one stakeholder. How should the User Interface address all needs?

Tradeoffs

The performance or speed of a system often makes it

The more observable and controllable, the less speed a system operates with. Based on the risk, context and level of autonomy you need to select your trade off

The more oversight is required, the higher the costs to design and operate.

Oversight might be required by more than one stakeholder. Define your sweet spot with enough controllabilty and observability to serve different personas

The less technically adept the user, the simpler the observability and controllability UI must be

Proportional to risk, context and autonomy.

Accuracy, Robustness and Cybersecurity

Article 15 of the AI Act requires providers to meet certain performance, robustness and cybersecurity obligations. Namely, it requires that providers select and meet accuracy targets for the AI system in production. It requires the AI system to be resilient to changes in the input, including through redundancy measures. And finally, it requires cybersecurity measures at the data, model and infrastructure level..

Requirement Clusters

Article 15 can be broken down into three key clusters of activities. By way of an example, we show what type of activities from our hypothetical use case are relevant to each cluster.

Article 15 Accuracy, Robustness, Cybersecurity		In our hypothetical use case the following practices might apply :
1. Accuracy	• Ensure level of accuracy is consistent • Declare metrics	• We select recall as performance metric. And the minimum performance threshold accepted by the domain expert (doctor) is recall 90%.
2. Robustness	• Resilience • Redundancy • Avoiding Biased Feedback Loops	• We ensured that our training data was representative of the production data and use anomaly detection to identify out of distribution cases.
3. Cybersecurity	• Cyber Resiliance Act	• We employ Role Based Access Control to manage activities and actions

Mapping the Clusters to the Requirements, WG Outcomes, and Relevant Standards

In the table below, we map the clusters above to the relevant text of Article 15, briefly outline the outcomes of the working group, and highlight international standards that may be relevant.

	Requirement Sourced from the Text of the AI Act	AI Act Reference	Best Practices from the Working Group	Standards
Accuracy	High-risk AI systems shall be designed and developed in such a way that they achieve an appropriate level of accuracy, robustness, and cybersecurity, and that they perform consistently in those respects throughout their lifecycle. The levels of accuracy and the relevant accuracy metrics of high-risk AI systems shall be declared in the accompanying instructions of use. ..... - Open communication channels with Product Owner, Domain Experts & include AI Governance team	Art. 13(3)	We develop a superset of best practices to ensure accuracy and robustness	Some commonly used performance metrics: ISO/IEC 29119-11 Section Art.8 ISO/IEC 24029-1 Section 5.2
Robustness	High-risk AI systems shall be as resilient as possible regarding errors, faults or inconsistencies that may occur within the system or the environment in which the system operates, in particular due to their interaction with natural persons or other systems. Technical and organisational measures shall be taken in this regard. The robustness of high-risk AI systems may be achieved through technical redundancy solutions, which may include backup or fail-safe plans. High-risk AI systems that continue to learn after being placed on the market or put into service shall be developed in such a way as to eliminate or reduce as far as possible **the risk of possibly biased outputs** influencing input for future operations (feedback loops), and as to ensure that any such **feedback loops** are duly addressed with appropriate mitigation measures.	Art. 15 (4)		ISO/IEC TS 25058, Section 10.5, pg. 9-10
Cybersecurity	High-risk AI systems shall be resilient against attempts by unauthorised third parties to alter their use, outputs or performance by exploiting system vulnerabilities. The technical solutions aiming to ensure the cybersecurity of high-risk AI systems shall be appropriate to the relevant circumstances and the risks. The technical solutions to address AI specific vulnerabilities shall include, where appropriate, measures to prevent, detect, respond to, resolve and control for attacks trying to manipulate the training data set (data poisoning), or pre-trained components used in training (model poisoning), inputs designed to cause the AI model to make a mistake (adversarial examples or model evasion), confidentiality attacks or model flaws.	Art. 15 (5)	Please refer to the Cyber Resilience Act	

Best Practices for the Implementation Steps

In the cards below, we detail the best practices developed by the working group for each cluster of activities that are relevant to Article 15.

1. Accuracy | **2. Robustness**

Best Practices from the Working Group

1. **Setup collaboration mode**
   - Open communication channels with Product Owner, Domain Experts & include AI Governance team
2. **Select performance metrics**
   - Discuss with Domain Experts & AI Governance team
   - Identify specific use case requirements and constraints
   - Match metrics to ML problem type (classification/ regression/etc.)
3. **Set appropriate thresholds**
   - Discuss with Domain Experts & AI Governance team
   - Get input from technical roles (advisory only)
   - Benchmark against market/competitor standards & review similar use cases for reference points
   - Incorporate trade-off considerations due to deployment context (e.g., limited processing power with edge computing)
   - Evaluate risk factors (Internal vs external use, reputational impact, acceptable bias levels)
   - Specify and document the target thresholds and quality gates to continue with the AI use case
   - Revisit target performance metrics defined in step 2 regularly during development
4. **Perform consistently throughout the lifecycle (see overview on the right)**
   - Select measures to perform consistently over the lifecycle
   - Implement measures to perform consistently over the lifecycle
   - Consider adaption to changes of the operational domain. Clarify if drift adaption is a feature instead of a problem (e.g. inflation)
5. **Document in instructions for use**
   - The levels of accuracy and the relevant accuracy metrics are declared in the instructions of use.

Learning Type	ML Technique	Metric	
Supervised	Regression[19]	• Root Mean Squared Error (RMSE) • Max Error	• Actual/Predicted Correlation
Supervised	Classification [19]	• Accuracy • Confusion Matrix • Precision and Recall • F1-score • AU-ROC	
Unsupervised	Clustering [20]	• Silhouette Score • Rand Index • Mutual Information	• Calinski-Harabasz Index • Davies-Bouldin Index
Self-supervised	LLM [21]	**RAG** • Faithfulness • Answer Relevancy • Contextual Precision • Contextual Recall • Contextual Relevancy	**Fine-tuning** • Hallucination • Toxicity • Bias • Summarization

	Data Engineering			Modelling		Deployment		
	Data Ingestion	Data Preparation	Data Management	Model Training	Model Management	Deployment Management	Monitoring	Feedback Loops
Accuracy	Data profiling Data augmentation Create diverse dataset Detect outliers Clean datasets Data validation			Model fit: no over-/ underfitting Regularization Error analysis per class		Prediction monitoring: Track model predictions for accuracy Drift detection: Identify shifts in data distribution or model performance Model updates: Refine and improve model systematically Post-market monitoring: Assess real-world model performance Continuous retraining: Maintain performance through periodic optimization		
Robustness	Redundancy in Data Platform Collect field data and understand the application Data distribution analysis Identify unexpected situations and try to collect data for them Test Representativeness in your train, test and validation datasets Test for unexpected situation in your dataset			Model versioning: Manage iterations and enable systematic retraining Multi-model approach: Deploy parallel AI models for enhanced resilience Fail-safe plan: Switch to rule-based systems during model failures		Performance monitoring: Continuously track system performance and environmental conditions Anomaly detection: Use unsupervised techniques for out-of-distribution analysis Human-in-the-Loop: Integrate human oversight Stress testing: Regularly simulate real-world scenarios to validate system robustness Cybersecurity standards: Follow the current standards and implement comprehensive measures Fault tolerance: Utilize redundant components and distributed load-balancing systems Deployment strategy: Develop comprehensive plan with fail-safe mechanisms and rollback options Operations: Audits and incident response plans Fallback plan in case of unexpected situation or bad performance: e.g. Backup procedures and manual override options		

1. Accuracy

AI Act Reference

Art. 13(3)

Additional Reference to Standards

Some commonly used performance metrics:
ISO/IEC 29119-11 Section Art.8
ISO/IEC 24029-1 Section 5.2

2. Robustness

AI Act Reference

Art. 15(4)

Additional Reference to Standards

ISO/IEC TS 25058, Section 10.5, pg. 9-10

3. Cybersecurity

Best Practices from the Working Group

Data Poisoning	Model Poisoning	Adversarial Examples or Model Evasion	Confidentiality Attacks
Data Integrity Checks: Use cryptographic techniques like hashing	Audit models periodically during and after training	Red teaming	Use metrics to evaluate LLMs (hallucination, relevance)
Data validation to filter out suspicious samples, duplicates, etc	control access (rights) to model	Adversarial training	Prompt engineering to defend against sharing private information
Quality (human / automated) Gates prior to adding data	Outlier analysis	Traing with adversarial examples	IP detection models
Filters	"Traditional" cybersecurity measures	Input Sanitization: Preprocess inputs to remove noise	Limit queries rate
			Differential privacy
			Pre-processing (Intent)
			Post-processing (Filter)

AI Act Reference

Art. 15 (5)

Option Questions and Tradeoffs for Article 15

Open Questions

There are many sectorial, but also AI capability-specific standards for accuracy. Which ones should prevail?

How should an enterprise justify that they have met an "appropriate level" of accuracy, robustness and cybersecurity?

Tradeoffs

Incremental improvements can impose significant costs. At what level should enterprises stop?

Required level of cybersecurity influences model choices and model hosting (e.g. on prem vs. cloud provider)

A Brief Overview on Transparency Obligations and Low Risk AI Systems

While the working group did not focus on this topic this year, we provide an overview of the transparency obligations for providers and deployers below.

Transparency Obligations

Under **Article 50** of the AI Act, the transparency obligations for providers and deployers of AI systems are defined.

	Definition	Obligation
Provider	1. The AI system **directly interacts** with natural persons	**Disclosure: Inform natural persons** they are interacting with an AI system, **unless this is obvious**
Provider	2. The AI system **generates synthetic** audio, image, video or text content	**Watermarking:** Outputs marked in a **machine-readable format** and **detectable as artificially generated** Exception: assistive function for standard editing or not substantially alter the input
Deployer	3. The AI system is intended to be used for **biometric categorization** or emotion recognition.	**Consent: Inform the natural persons** exposed thereto of the operation of the system, and shall process the personal data in accordance
Deployer	4. The AI system is generating image, audio or video content constituting **deep fake** or generating text.	**Disclosure:** Shall **disclose** that the content has been artificially generated or manipulated.

Common exception: AI systems authorized by law to detect, prevent, investigate or prosecute criminal offence.

Low Risk

Under **Article 95** of the AI Act, operators of low-risk AI systems can voluntarily choose to follow a Code of Conduct. The Commission and Member States will encourage companies to apply some or all of the requirements for high-risk AI systems even while developing and building low-risk AI systems. In the table below, we identify potential high-risk requirements, or parts therof, that operators of low-risk systems could consider voluntarily applying.

High-risk Requirement	Optional Elements that are Potentially More Relevant to the Voluntary Code of Conduct
Article 9	• Assessing and preventing the negative impact of AI systems on vulnerable persons or groups of vulnerable persons • Reputational risks or others beyond AI Act • Assessing and minimising the impact of AI systems on environmental sustainability
Article 10	• Implementing all of Article 10's data governance requirements • Data quality considerations relevant to business use case
Article 12	• Monitoring the performance of the AI system in production
Article 13	• Documentation best practices like model or system cards
Article 14	• Humans-in-the-loop based on context of deployment • AI literacy trainings (note: mandatory under Article 4)
Article 15	• Meeting and monitoring basic performance criteria • Enterprise cybersecurity standards

Limitations

- This working group focused on developing best practices to operationalize high-risk requirements of AI systems. Nevertheless, we expect the harmonized standards to look different. The target of this approach is to help companies to already start preparing their processes, skills, infrastructure before the harmonized standards are released.
- The working group did not discuss best practices to implement limited risk and low risk AI systems. Neverthless, we provide an overview of the relevant requirements from the AI Act.

Section 5: ML Skills Profiles under the AI Act and Getting Started Today

ML Skill Profiles under the AI Act

In the previous section, we addressed the second challenge: the lack of harmonized standards. In this section, we focus the third main challenge identified in our analysis in Section 2: the uncertainty surrounding which skill profiles are needed to operationalize the AI Act effectively and and what companies can start doing today to prepare to operationalize the AI Act.

Why Do We Need Skill Profiles?

The Challenge

We have established in the previous sections that implementing the AI Act requires tight coordination and collaboration across diverse stakeholders across all the layers of governance. From data scientists and MLOps engineers to AI governance and legal professionals.

In addition, organizations vary widely in structure and maturity. Hence, there is no one-size-fits-all approach to implementing the EU AI Act. Without a clear understanding of the required expertise, organizations risk both under-engineering and over-engineering their governance structures, which can slow down compliance and innovation.

Method

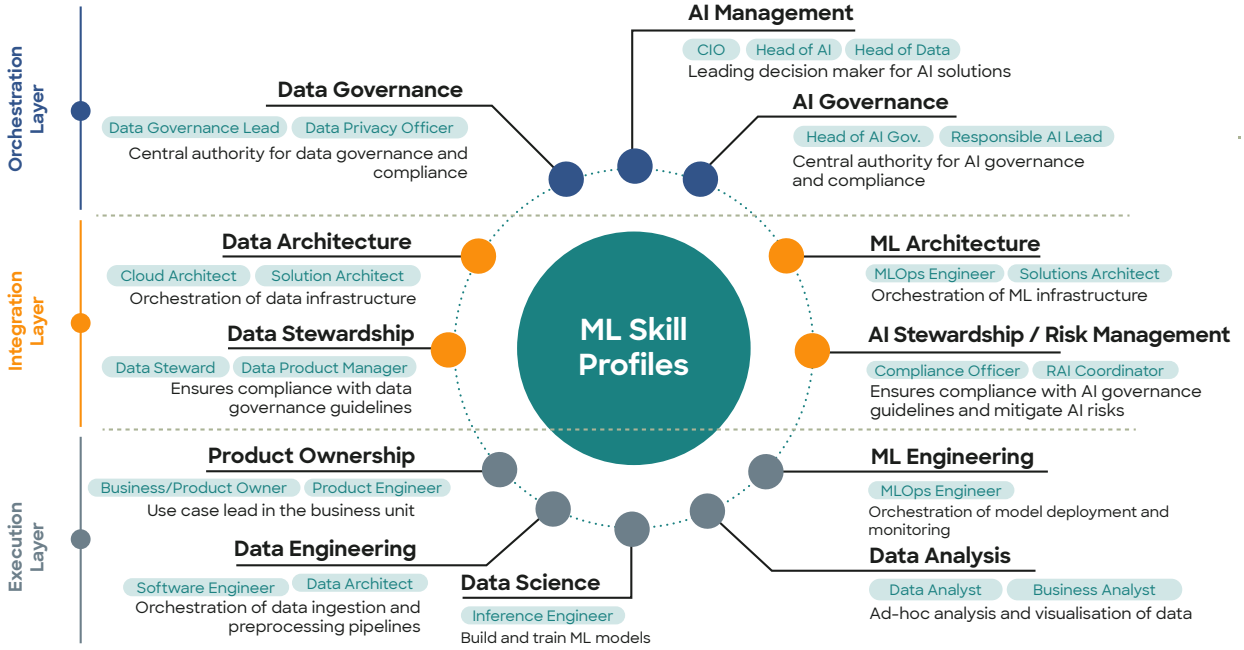
To address this challenge, we build upon the initial ML Skill Profiles framework developed by appliedAI in 2024, as it is a flexible framework based on functional roles rather than rigid job titles. Although this earlier framework was not specifically tailored to meet regulatory demands, its structure provides a strong foundation for identifying skill gaps that may emerge under the AI Act.

Info

The ML Skills Profile framework is a comprehensive organizational blueprint to scale machine learning in enterprises. It consists of a set of skill profiles with well-defined responsibilities and skill sets along the ML lifecycle. Skill profiles do not map directly to full-time employees (FTEs); a single individual may take on multiple roles or "hats".

The ML Skill Profiles under the AI Act

In this whitepaper, we present an extension of the ML Skill Profiles framework, explicitly aligned with the requirements of the EU AI Act. For organizations seeking to scale up the number of projects deployed into production systems, we identified the skill profiles that contribute to a machine learning project throughout its lifecycle. These skill profiles are the following:

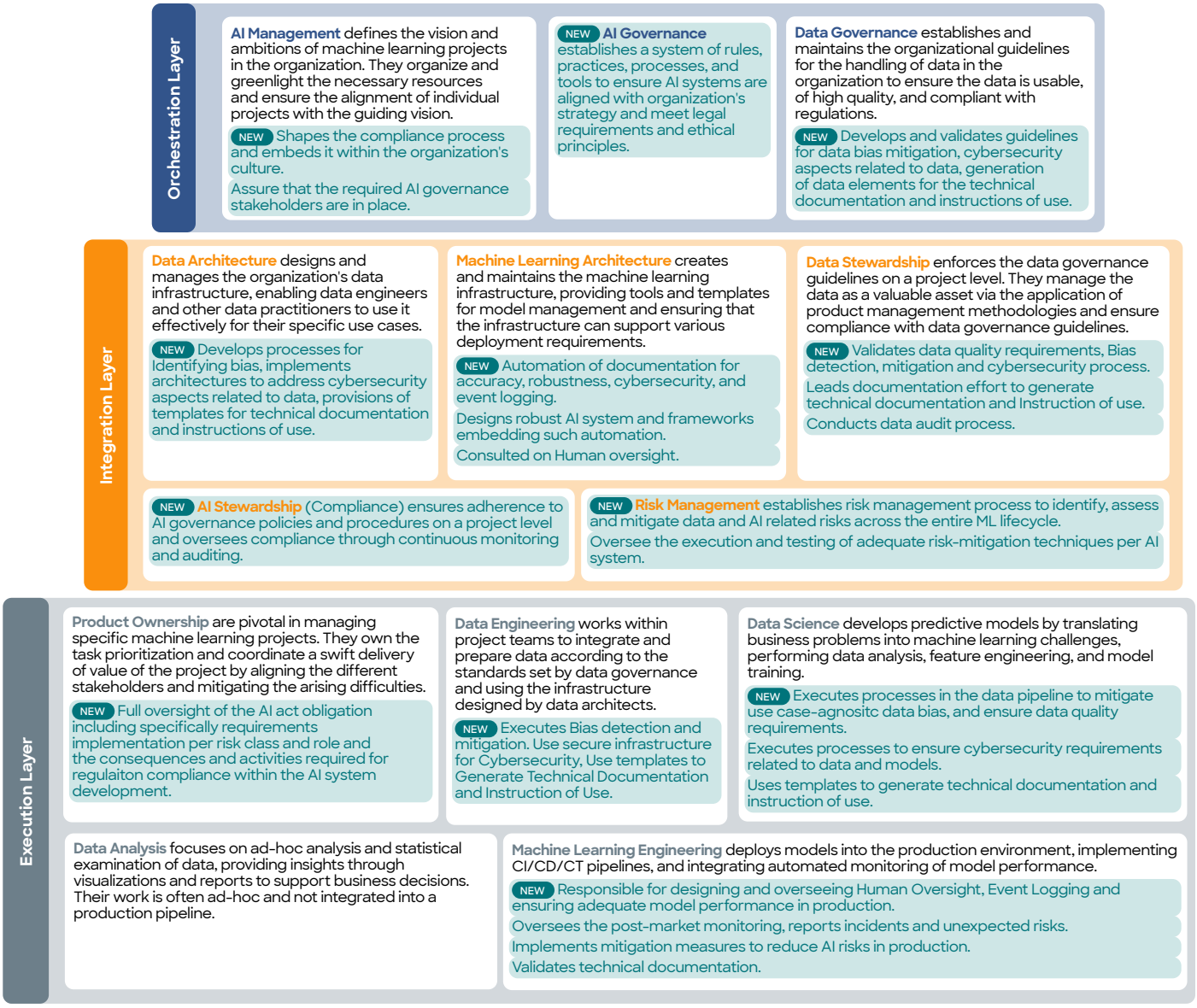


ML Skill Profiles under the AI Act

ML Skill Profiles Mapped to the AI Act Governance Layers

This enhanced model introduces three key new skill profiles usually executed across the orchestration, integration, and execution layers to support scalable, cross-functional compliance: AI Governance, AI Stewardship and risk management.

We also map these skill profiles to the AI Governance pyramid layers, and describe them below.



NEW All activities highlighted in the green boxes are the extended skills related to the delta that the AI Act responsibilities might require or reinforced in comparison to those previously defined in the ML skill profiles whitepaper.

Note: While most of the Skill Profiles are mapped to a specific layer for clear accountability, the Skill Profiles in the integration layer typically set up processes that are applied in the Execution Layer.

How to Apply the ML Skill Profiles to Your Company

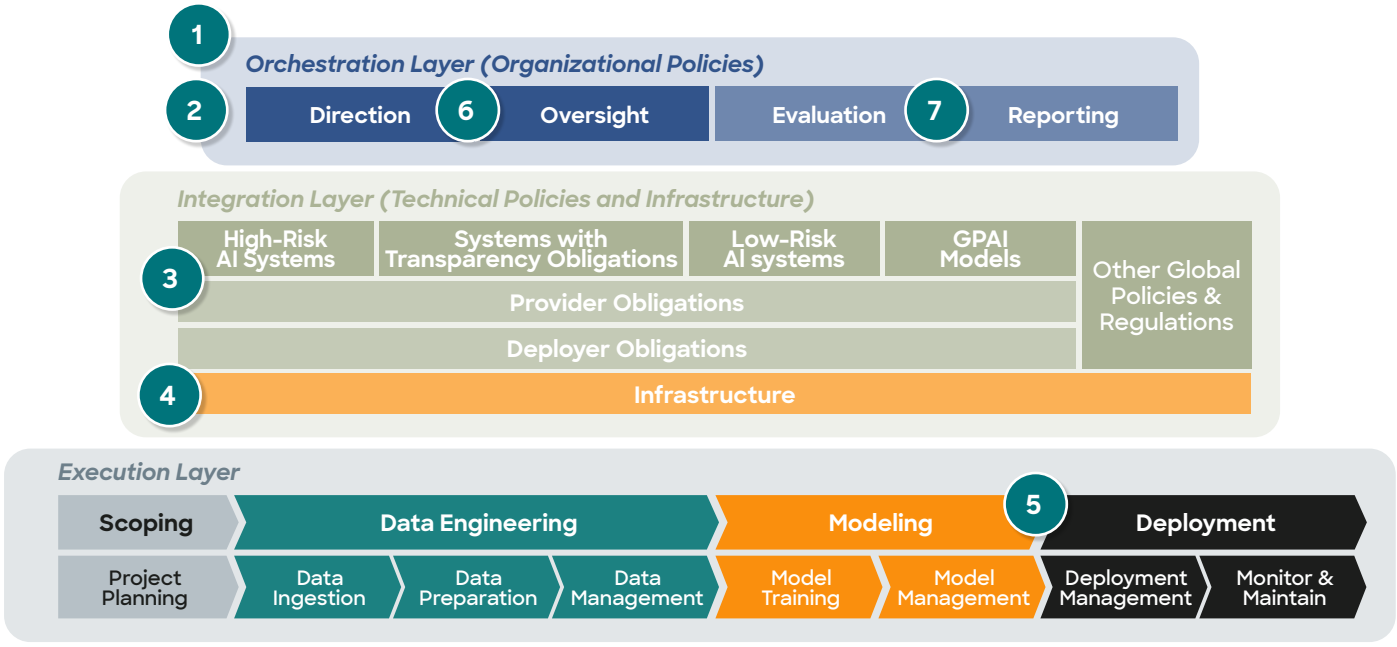
While the framework is based on common patterns identified across various companies, it's important to note that it may not capture the unique aspects of every organization. Tailoring the framework to fit the specific context and requirements of your organization is key. The way of application depends on the organization and machine learning team size, the company's AI maturity level, and the desired degree of centralization of certain activities. More details on how to apply the framework to your company's specifics, can be found on the last section of the original ML Skill Profiles whitepaper.

Getting Started Today: *Bridging the Gap*

As AI adoption accelerates, regulatory compliance is evolving from a legal obligation into a **strategic differentiator.** The EU AI Act demands **rigorous processes, clear accountability,** and **comprehensive documentation.** To succeed, organizations must take a structured approach and choose the right tools, processes and partners to enable operationalization at scale.

Here are seven actionable steps to help your organization begin its journey toward EU AI Act compliance:

1. **Identify your landscape of obligations** (risk landscape) under the AI Act by creating an inventory of AI systems in your company and classifying their risk class and role.
2. **Set up an EU AI Act governance process** to define your regulatory strategy based on your risk landscape. Decide which risk classes to implement and define clear roles & responsibilities on policies oversight, and governance evaluation.
3. **Assess gaps** in your processes and MLOps tools depending on the risk classes your company strategically decided to implement.
4. **Streamline AI Act technical implementation** through standardized and reusable infrastructure components that automate as much as possible the compliance evidence gathering and speed up time-to-compliance.
5. **Execute and test policies** for your respective obligations under the EU AI Act with tools that support collaboration between technical and non-technical stakeholders
6. **Upskill your workforce with AI literacy**, depending on your risk classes and roles. Define and execute proper AI literacy training for the AI user, AI system developers and deployers, and new proposed skill profiles.
7. **Establish a robust evaluation and reporting mechanism** for your governance process for further improvements. **Joining communities or working groups** can help to learn from others by exchanging AI Act best practices..



Outlook

This whitepaper is the baseline for operationalizing specific AI system requirements of the AI Act by addressing the current challenges in corporations. We map this work with available international standards and provide best practices and the necessary processes to implement those requirements.

This document is instrumental at a stage where regulatory simplification is required as one of the main pillars of the new EU AI continent action plan [22] and could contribute to Commission initiatives that gather examples for implementing the AI Act [23].

With this work, we demonstrate that regulation, corporate strategy, and AI implementation can be unified and optimized, through efficient AI governance, potentially reducing the time to compliance..

Legal Disclaimer

appliedAI does not offer legal advice: We are not a law firm, but we aid organizations to navigate through the AI Act from a practical perspective.

AI Transparency Content Voluntary Statement

All frameworks, tables, and blueprints presented in this whitepaper have been created by their respective authors based on their expertise, information extracted from working groups, interactions, and feedback from collaborators. Therefore, for transparency purposes, it is stated here that none of the commented contents above has been generated using AI technology. The image on the cover of the first page of this paper was generated using AI, and its transparency note includes the prompt and GPAI platform that provided the outcome.

References

[1] From responsible AI governance to competitive performance: The mediating role of knowledge management capabilities, 21st IFIP Conference I3E2022 (2022)

[2] Organizations Face Challenges in Timely Compliance With the EU AI Act, MIT Sloan Management Review (2024)

[3] The EU AI Act: A Double-Edged Sword For Europe's AI Innovation Future, Forbes (2025) The Impact of the EU AI Act's Transparency Requirements on AI Innovation, 19th International Conference on Wirtschaftsinformatik (2024)

[4] ML Skill Profiles: An Organizational Blueprint for Scaling Enterprise ML, appliedAI initiative GmbH (2024)

[5] Corporate Partners and references ("What our customers say" about appliedAI AI Governance Working Group), appliedAI initiative GmbH (2025)

[6] AI Act Official Text. Regulation (EU) 2024/1689, The European Parliament and the Council of the European Union (2024)

[7] General-Purpose AI Models in the AI Act - Questions & Answers, AI Office (2025)

[8] Guidelines on the definition of an artificial intelligence system established by AI Act, AI Office (2025)

[9] General-Purpose AI Code of Practice (CoP), AI Office (By the time of this publication the final CoP version was not published, 2025)

[10] Harmonised Standards for the European AI Act & AI Standards, The Joint Research Centre, EU Commission (2024)

[11] CEN CENELEC & JTC21 Artificial Intelligence (N/A)

[12] Narayanan, M., & Schoeberl, C. (2023). A Matrix for Selecting Responsible AI Frameworks. Center for Security and Emerging Technology. CSET (2023) & Implementing AI Governance: from Framework to Practice EU AI alliance iunder EU commission (2023) WIP to complete more frameworks links, add source and year

[13] European AI Standards - Technical Standardization and Implementation Challenges under the EU AI Act, KI Bundesverband (2025)

[14] MLLC is the Machine Learning Lifecycle is based on the initially proposed in The Enterprise Guide to Machine Learning, appliedAI initiative GmbH (202X) and aligns with the first OECD MLLC proposal, OECD (2019)

[15] Other available framework used in AI Governance:
NIST AI Risk Management Framework (RMF) 1.0, NIST, 2023
AIGA, Artificial Intelligence Governance and auditing Framework (also Hourglass model), Türku University, 2022

[16] ISO/ IEC Standards related with high-level of governance, AI Management System respectively:
- ISO/IEC 38507:2022 Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations, ISO/IEC (Edition 1, 2022), Published
- ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system, (OBP) ISO/IEC, (Edition 1, 2023), Published
- ISO/IEC 42006 Information technology — Artificial intelligence — Requirements for bodies providing audit and certification of artificial intelligence management systems, (Edition 1, 2025) Under development at the time of this publication and not included in the whitepaper's outcome

[17] CEN/CLC/JTC 21, Artificial Intelligence, Work Programme, CEN-CENELEC, (N/A)

[18] Cardiovascular disease (CVD), WHO (2021)

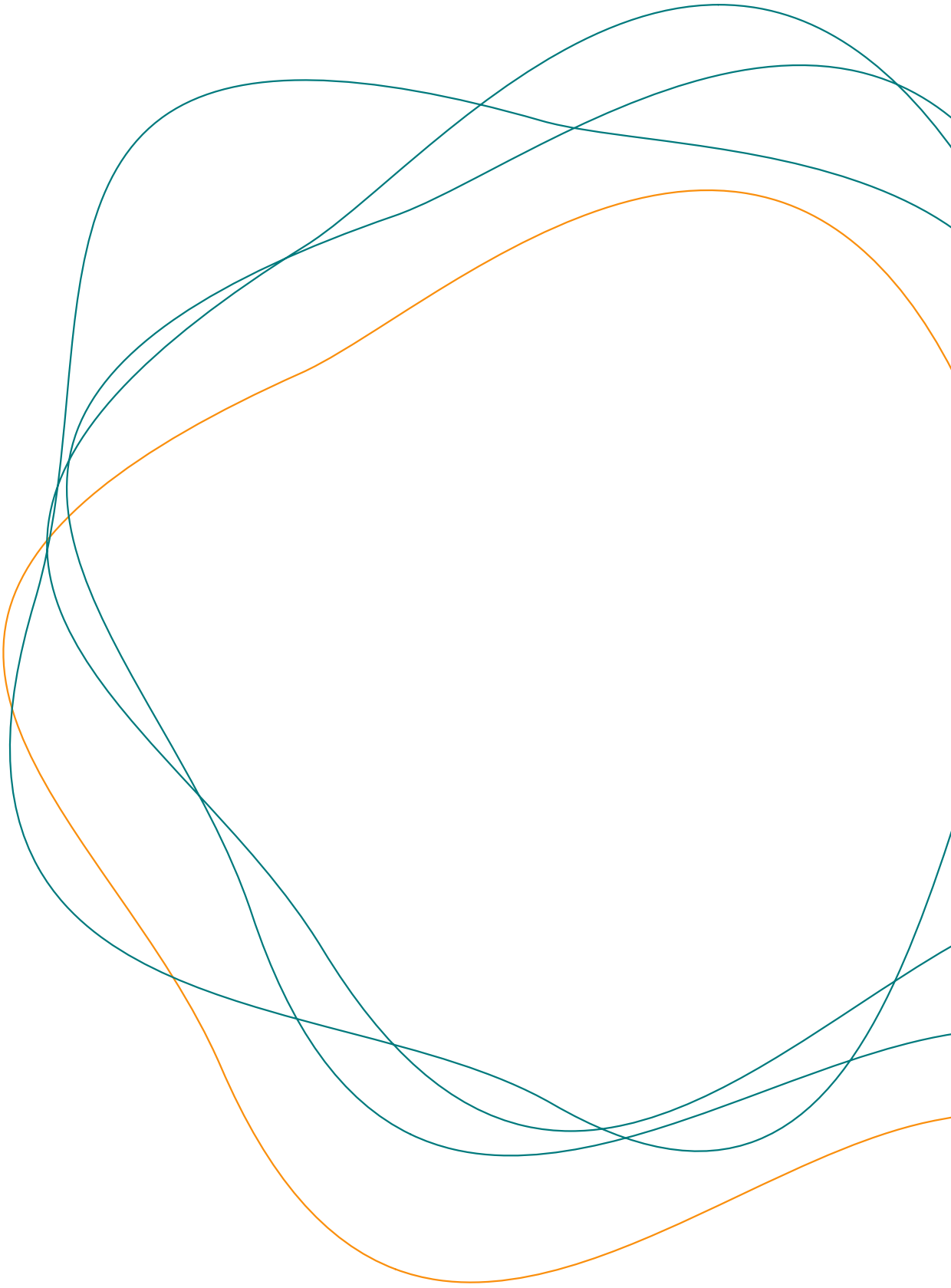
[19] Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow, O'Reilly Media, Inc. (2022)

[20] Exploring Unsupervised Metrics, KD Nuggets, (2023)

[21] LLM Evaluation Metrics: The Ultimate LLM Evaluation Guide, Confident AI (2025)

[22] AI Continent Action Plan Fact pages, EU Commission (2025)

[23] Public consultation on high-risk AI systems, EU Commission (2025)



Authors

Alexander Machado is the Head of Trustworthy AI CoE and former Head of MLOps Processes at appliedAI Initiative. He has a decade of experience in Data Science, Artificial Intelligence, and Data Engineering at appliedAI, the Max Planck Society, and BMW. His work focused on leading, planning, and developing AI solutions from experimentation to production. He has led multiple AI Governance-MLOps working groups, published an MLOps online course, and developed practical frameworks that address the inherent challenges of production systems and compliance with the EU AI Act.

Manuel Jiménez Mérida is a Senior AI Governance Strategist and Trustworthy AI Expert. His prior work ranged from advancing AI strategies for major corporations, to developing new AI governance frameworks, and he brings over a decade of experience in Corporate Technology in Japan and EMEA. He holds a Master's degree from the Technical University of Munich in collaboration with UC Berkeley, where his thesis centered on operationalizing Responsible AI Governance. His current focus is on simplifying and scaling the implementation of Trustworthy AI Governance.

Akhil Deo is a senior AI regulatory expert. He has 6 years of experience working on technology and public policy. At appliedAI, he led the 2024 AI Governance working group with leading European companies on implementing the requirements for high-risk AI systems.

Anish Pathak is an ML Engineer with broad experience in AI engineering, data engineering, and data mining at appliedAI, Memodo, and Novartis. Having worked across the entire data lifecycle, he is passionate about MLOps and implementing its best practices. At appliedAI, he contributed to translating the EU AI Act's requirements into practical technical guidelines.

Contributors

The frameworks in this whitepaper were developed with the support of many companies in Germany that shape the domain as part of the AI Act Governance and MLOps under the AI Act working groups at appliedAI. All partners of appliedAI have contributed in one way or another, and we wish to thank them. Additionally, several individuals from these companies have made an extra effort to make this whitepaper available to the public. We want to thank these individuals and their respective companies.

Simone Oldekop
Former Head of Responsible AI Office
Carl Zeiss AG

Steffen Herterich
Lead Principal Engineer - Data Protection and Privacy
Infineon Technologies AG

Cecilia Carbonelli
Senior Principal - Head of Algorithm Concept & Modeling - Responsible AI Tech Lead
Infineon Technologies AG

Eljalill Tauschinsky
Consultant data protection and data law
EnBW Energie Baden-Württemberg AG

Alexandra Wander
Program Manager - Responsible AI
Carl Zeiss AG

Asad Preuss-Dodhy
Sr. Principal - Data Anonymisation and Privacy Technologies Expert
Roche Diagnostics GmbH (Information Solutions)

Araceli Alcala
RA Manager | RA SME for AI
Carl Zeiss Meditec AG

Sebastian Hallensleben
Chair of Joint Technical Committee (JTC) 21, AI
CEN and CENELEC

Dirk Wacker
AI Lead
Giesecke+Devrient GmbH

Geoffroy Pavillet
Data Protection Counsel
Linde GmbH

Christiane Miethge
Senior Manager AI Communication and Policy
Infineon Technologies AG

Heinrich Dold
Senior Transformation Manager
EnBW Energie Baden-Württemberg AG

Simone Heitzer
AI Strategist
MTU Aero Engines AG

Sona Jose
Responsible AI Consultant
Carl Zeiss AG

Philippe Coution
Head of Digital Interaction & Lead AI Quality
TÜV SÜD AG

Acknowledgements

The authors want to thank also Arash Heidarian, Hendrik Scherner, Marika Horne, Doris Melanie Fonseca Lima, Susanne Klausing, Ronnit Wilmersdörffer, Pawan Kumar Goyal, and Julia Pfeiffer for their feedback and contributions in this paper. Additionally, to the family members who encouraged the authors to keep working on Trustworthy AI and left us in 2024.

About appliedAI Initiative GmbH

appliedAI is your partner in AI transformation. Empowering businesses to lead in the age of AI.

Our goal: Advancing businesses to compete in the age of AI, shaping a future we desire to live in.

We guide our partners from first steps to full maturity unlocking sustainable value and reshaping their markets - individually or in programs.

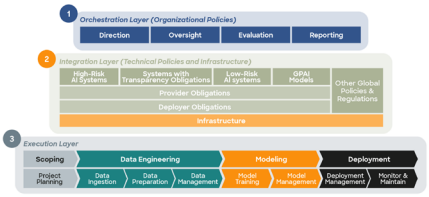
- ✓ **We support strategically how to do value creation with AI.**
- ✓ **We develop and implement AI use cases.**
- ✓ **We upskill your teams.**

For more information, please visit https://www.appliedai.de/en/

The EU AI Act could put compliance hurdles in the way of business.
BUT also gives us an opportunity to create high-quality AI products and services.

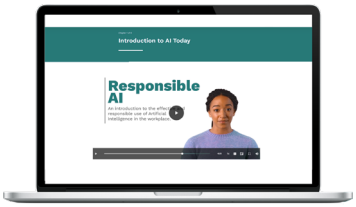
Setup AI Act Governance in your organization

Get clarity on your organisation's risk ambition and the required organIsational setup, processes and responsibilities to achieve AI Act readiness.



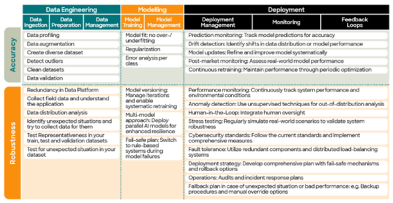
A compact, company wide AI literacy training

Get AI Act ready in 45 minutes, scaling AI literacy to upskill thousands of AI users in your organization



Implementation of automated technical AI Act policies

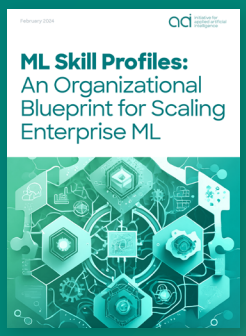
Build the right infrastructure to develop, deploy, and monitor AI systems in line with the EU AI Act requirements. Reach out for our demo.



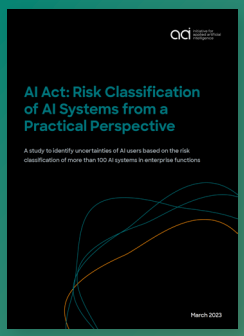
appliedAI is your trusted partner to help you achieve these goals.

Do you want to dive deeper into implementing the AI Act?
Start your journey with our white papers.

This white paper outlines how MLOps platforms can support compliance by simplifying collaboration, standardizing processes, and automating reporting.



This whitepaper presents the ML Skill Profiles framework, an organizational blueprint for scaling machine learning in production.



This study examines how the risk classification criteria of the EU AI regulation affect AI innovation within companies.

Join the appliedAI Partnership

The frameworks in this whitepaper were developed in appliedAI's working groups.

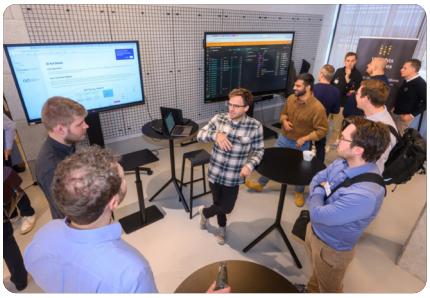
Join the appliedAI partnership program to participate in the working groups and access these benefits:

Co-creation, collaboration, learning from and with others

- Roundtables: "Engineering Design in the Age of Generative AI", "AI Agents Ops", "AI Agents under the AI Act".
- Working groups: "AI Act Governance", "MLOps under the AI Act".
- Agentic AI / GenAI Delegation Visits, e.g. London.
- Quarterly Senior AI&Data Executive event for leading Applied AI partner companies and selected ecosystem participants, invite only, no vendors.

Access to state-of-the-art knowledge and unique ecosystem

- Meetups & Conferences, e.g. "AI Agents and the New Era of Business", "MLOps Day: AI Agent Ops", "AI Act Governance", "Creating and Capturing Value with AI Business Model Innovation".
- Learn @ Lunch Sessions
- Masterclassess e.g. Intro to AI Project Management, AI Project Management under the AI Act, AI Strategy, Innovating Business Models with AI.



Individual guidance

- Direct sparring
- Annual strategy development

Partner with us!

initiative for
applied artificial
intelligence

AI Act Governance:
Best Practices for
Implementing the EU AI Act

appliedAI Initiative GmbH

August-Everding-Straße 25
81671 München
Germany
www.appliedai.de