

# AI Act – Getting Practical

IAB - Master your GDPR  
Maria Koskinen  
AI Policy Manager, Saidot



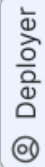

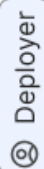
**Saidot**

# AI Act highlights for AI business and product teams

- The AI system's definition aligns with the OECD's and is **sufficiently broad to deter speculative avoidance**; however, further clarification via examples is needed
- The AI Act **does not regulate technologies or industries**; instead, it covers AI use cases across various types of organisations and sectors
- The AI Act **does not give you exact policies to implement** but puts you define a governance framework to set the limits of appropriateness in your own AI system context
- While the AI Act primarily focuses on the obligations of high-risk AI providers and deployers, it also **includes requirements for all EU AI system providers**
- The requirements apply mostly to **new systems placed on the market** – high-risk systems already on the market are mostly not covered at all

# When to get ready? Main scenarios.

1. Provider of an AI system that is prohibited in the EU market **(6 months)**
2. Provider of a general-purpose AI system **(12 months)**
3. Provider of a standalone high-risk AI system as defined in Annex III **(24 months)**
4. Provider of an AI system that interacts with natural persons or creates synthetic content, or deployer of emotion recognition, biometric categorisation, deepfake systems, or certain AI systems manipulating text **(24 months)**
5. Deployer of a high-risk AI system developed by a third-party provider **(24-36 months)**
6. Provider of a high-risk AI system subject to harmonisation legislation in Annex II **(36 months)**

PROHIBITED SYSTEMS	HIGH-RISK SYSTEMS	LIMITED RISK AND GENERAL-PURPOSE AI	GENERAL-PURPOSE AI
<ul style="list-style-type: none"> <li>Not permitted.</li> </ul>	<div data-bbox="359 211 402 376" style="float: left; margin-right: 10px;">  Provider         </div> <p><b>Requirements</b></p> <ul style="list-style-type: none"> <li>Risk management system (Art.9)</li> <li>Data and data governance (Art.10)</li> <li>Technical documentation (Art.11)</li> <li>Record-keeping (Art.12)</li> <li>Transparency and provision of information to users (Art.13)</li> <li>Human oversight (Art.14)</li> <li>Accuracy, robustness and cybersecurity (Art.15)</li> </ul> <hr/> <p><b>Obligations</b></p> <ul style="list-style-type: none"> <li>Quality management system (Art.17)</li> <li>Post-market monitoring (Art. 61)</li> <li>Conformity assessment (Art.43)</li> <li>EU declaration of conformity (Art. 48)</li> <li>CE marking (Art. 49)</li> <li>Registration to EU Database (Art. 51)</li> <li>Identification and contact information (Art.16aa)</li> <li>Demonstration of conformity (Art.16j)</li> <li>Accessibility (Art. 16ja)</li> <li>Documentation keeping (Art. 18)</li> <li>Automatically generated logs (Art. 20)</li> <li>Corrective actions and duty of information (Art. 21)</li> <li>Cooperation with authorities (Art. 23)</li> <li>Reporting of serious incidents (Art. 62)</li> <li>Access to data and documents (Art. 64)</li> </ul>	<div data-bbox="1477 211 1521 376" style="float: left; margin-right: 10px;">  Provider         </div> <div data-bbox="1477 391 1521 556" style="float: left; margin-right: 10px;">  Deployer         </div> <ul style="list-style-type: none"> <li>AI interaction transparency (Art. 52)</li> <li>Synthetic content disclosure and marking requirement (Art. 52)</li> <li>Emotion recognition and biometric categorisation AI system operation and data processing transparency (Art. 52)</li> <li>Disclosures of AI-generated deepfake content (Art. 52)</li> <li>AI-generated public interest text disclosure obligation (Art. 52)</li> </ul>	<div data-bbox="1982 211 2025 376" style="float: left; margin-right: 10px;">  Provider         </div> <ul style="list-style-type: none"> <li>Synthetic content disclosure and marking requirement (Art. 52)</li> <li>Model technical documentation, including training, testing, and evaluation results (Art. 52c)</li> <li>Provision of information and documentation to providers integrating the GPAI model in their AI system (Art. 52c)</li> <li>Compliance with copyright law (Art. 52c)</li> <li>AI system training content summary disclosure (Art. 52c)</li> <li>Cooperation with authorities (Art. 52c)</li> </ul>
	<div data-bbox="359 939 402 1105" style="float: left; margin-right: 10px;">  Deployer         </div> <ul style="list-style-type: none"> <li>Measures to comply with the instructions of use (Art. 29)</li> <li>Assign human oversight to natural persons and ensure their competence (Art. 29)</li> <li>Input data relevance and representativeness (Art. 29)</li> <li>Monitor the operation and inform providers (Art. 29)</li> <li>Record-keeping (Art. 29)</li> <li>Workplace deployment notification (Art. 29)</li> <li>Public authority registration to EU Database (Art. 51)</li> <li>Implement data protection impact assessments (Art. 29)</li> <li>Judicial authorisation for exempted use of post-remote biometric identification (Art. 29)</li> <li>AI decision transparency disclosure (Art. 29)</li> <li>Cooperation with authorities (Art. 29)</li> <li>Fundamental rights impact assessment (Art. 29a)</li> </ul>		<p><b>GPAI with systemic risk</b></p> <ul style="list-style-type: none"> <li>Standardised model evaluation and adversarial testing (Art. 52d)</li> <li>Risk assessment and mitigation (Art. 52d)</li> <li>Incident and corrective measure tracking, documenting, and reporting (Art. 52d)</li> <li>Cybersecurity protection (Art. 52d)</li> </ul>

# How are companies getting ready for the coming AI Act compliance?

Good practices from working with AI Act early adopters

01

AI policy as a framework for responsible AI development and use

02

AI inventory and a way to classify AI systems based on risk-level

03

Lifecycle-based AI governance, incl. data governance, risk management

04

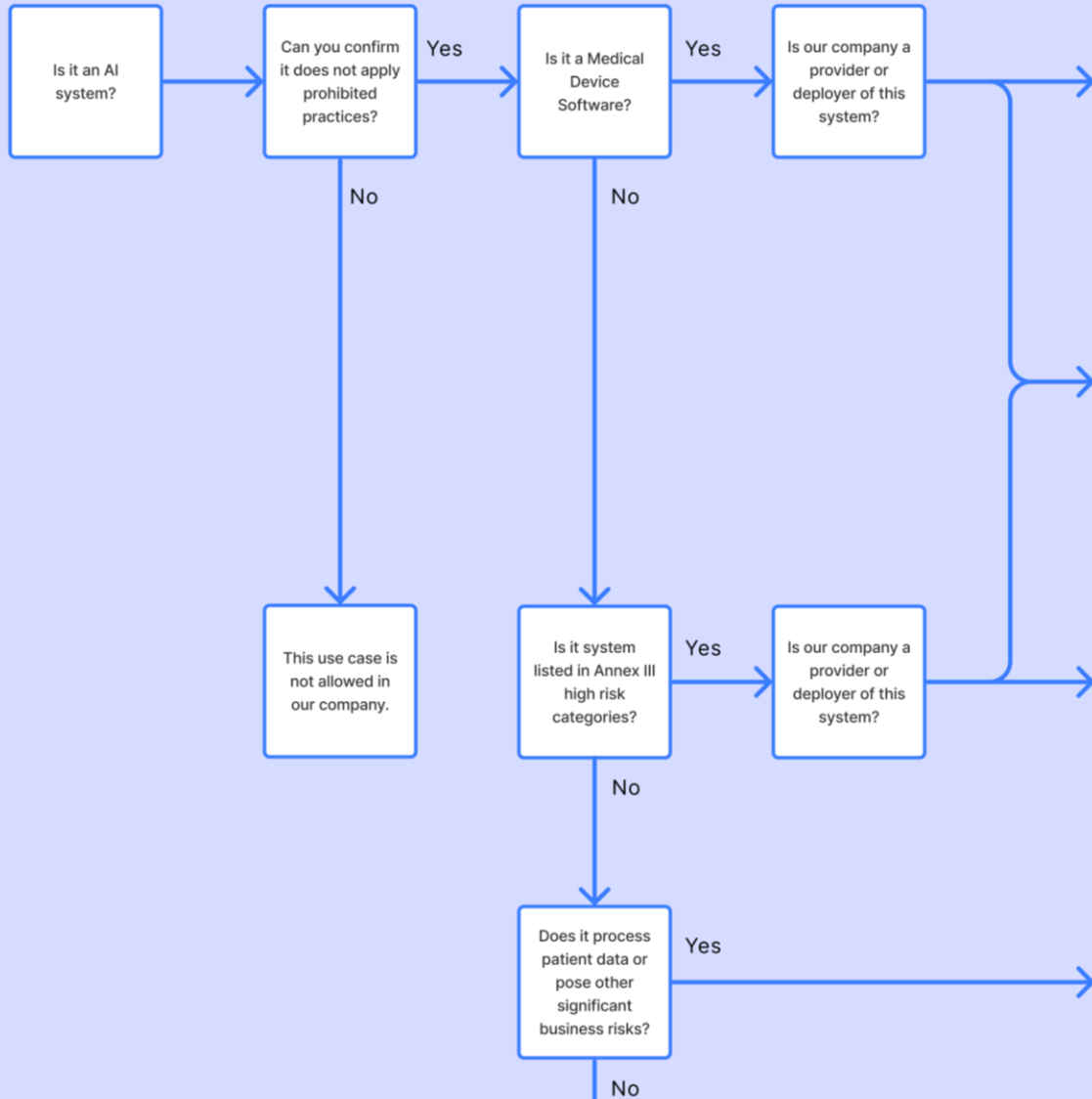
Attention to AI procurement and use of third-party components

# 01. AI Policy

Company's formally expressed commitment and direction to act responsibly in its role as a developer, provider and deployer of AI systems.

Content topic	What do we want to determine?
Framework for setting <b>AI policy objectives</b>	What do we want to achieve in the responsible use of AI systems in our company?
<b>Principles</b> that guide all activities of the organisation related to AI	Which principles guide our AI activities?
Commitment to meet <b>applicable requirements</b>	What measures do we have to meet the objectives of responsible use of AI in our company?
Commitment to <b>continual improvement</b>	How is our AI policy going to be reviewed and improved over time?
Processes for <b>handling deviations</b> and <b>exceptions</b> to policy	Who does this policy apply to and how are potential deviations handled?

# 02. AI inventory and risk-based classification system



The screenshot shows the 'Mental Health Coach' system interface. The breadcrumb trail is 'Govern / Systems / Mental Health Coach'. The page title is 'Mental Health Coach'. A search bar for templates is visible. The left sidebar contains navigation options: LIBRARY (Feed, Policies, Risks, Models, Evaluations), GOVERN (Systems, Templates). The main content area is titled 'Recommended Policy Templates' and lists three templates:

- AI Act Requirements for High-Risks...** (Provider, 78 Requirements): Ensure your AI system compliance with the EU AI Act requirements for high-risk AI-based medical device software systems. This template combines the requiremen...
- AI Act Requirements for Standalone...** (Provider, 78 Requirements): Ensure your AI system compliance with the EU AI Act requirements for high risk systems. The template contains requirements applicable to providers of high risk AI systems.
- Records of Processing Activities** (Controller, 7 Requirements): Records of Processing Activities is a template based on the...

# 03. Lifecycle-based AI governance

**PoWind Forecast** | Tasks

Overview | Policies | Risks | Review | Monitor

## Overview

**Intended purpose**

PoWind Forecasts aims to solve the issue of ensuring the balance of demand and supply on an electricity grid. In the grid, the balance must be maintained all the time between consumption and generation of electricity. If the balance is not ens...

UID: 232432

**Lifecycle:** Operation and monitoring [Edit lifecycle](#)

**Org. Role:** Deployer

**Tags:** Wind

**Region:** EU

**Industry:** Energy

**System Type:** Forecasting system

**Capabilities:** Question answering

**Risk level:** High-Risk

**Risk justification:** The PoWind Forecast is used as an input to critical energy infrastructure operations decisions. [Edit](#)

**Business Impact:** Medium impact [+ Add justification](#)

## Tasks

- Initiation** (100%)
  - Register system
  - Determine context
  - Determine suppliers
  - Determine intended use
  - Determine business impact
  - Determine risk level
  - Determine applicable policies
- Design and development** (100%)
  - Determine roles and responsibilities
  - Determine system components
  - Evaluate capabilities and limitations
  - Perform risk management
  - Implement policies
  - Manage third parties
- Verification and validation** (100%)
  - Conduct reviews
  - Adjust based on review feedback
  - Approve system
- Deployment** (100%)

**PoWind Forecast**

Search risks

**Recommended Risks**

Click on a risk card to add and start

- Prompt injection attack** (2 Mitigations)

Prompt injection attacks feed generative models with tailored text prompts that cause them to overlook original or previous instructions prompt to produce unexpected and potentially harmful text or to
- Bias amplification** (10 Mitigations)

All systems can amplify biases in the training data, which means the model makes certain predictions a lot more often for some groups than is expected based on t
- Slow response times** (1 Mitigation)

All systems may encounter performance issues such as slow response times during system operation. This



# 04. Attention to the use of 3rd party products

The screenshot shows the Hugging Face Model Library interface. At the top, there's a navigation bar with the user's name 'Kuba Martins, Pineapple Inc' and a search bar. Below the search bar, there are several filter buttons: Category, AI SystemType, Industry, Tasks, Data Location, Source Code, and Architecture. The main content area displays a grid of model cards. Each card includes the model name, provider, update date, number of evaluations, and a 'Follow' button. The models shown are GPT-4 (Open AI), PaLM 2 (Google), DALL-E 3 (Open AI), Command (Cohere), Llama-2 (Meta), and BLOOMZ (BigScience). The interface also features a left sidebar with navigation options like LIBRARY, GOVERN, and various tool icons.

This is a detailed view of the PaLM 2 model page. The header shows the model name 'PaLM 2' and the provider 'Google'. There are buttons for 'Follow' and 'Product page'. The page is updated as of Nov 24, 2023. Below the header, there are tabs for 'Overview', 'Training', and 'Evaluations'. The 'Overview' tab is active, showing a list of metadata: Type (Large Language Model), Modality (Language), Regions (US), Industry (General purpose), and Tasks (Language generation, Commonsense reasoning, Multitask language understanding). The 'Source code' is listed as 'Closed source'. There is a 'Versions' section with links to 'chat-bison' and 'text-bison'. The 'Technical Details' section lists Languages (Multilingual), Architecture (Transformer), Data Location (US, Canada, Germany, Belgium), and Parameters (340 billion).

# Where to start?

## Start immediately

- **Assign responsibilities** for AI governance and the AI Act preparedness
- Create a **policy** as a framework for responsible AI development and use
- Create a **classification system** to identify systems with different levels of business and regulative risks
- Make an **inventory** of your AI systems, own and third parties
- Manage potential **prohibited use cases**

## Do over next 12 months

- Establish **core AI governance processes** in the development of **new AI systems**: data governance, risk management, evaluations, technical documentation
- **Manage contracts** with third parties to avoid contractual infringements and align with regulation's requirements
- Assign roles and responsibilities for upcoming **conformity processes**
- Initiate **quality management system** preparedness
- Ensure **AI transparency** across all AI systems

## Wait for further clarity

- Identify and adopt relevant **standards, code of conducts, and supervisory guidance**
- Adopt **detailed requirements, quality management system, post-market monitoring and conformity processes**
- Registration to **EU Database** and processes for **communication with relevant authorities**
- Governance and compliance of systems falling under **high-risk based on sectoral regulations** based on Commission guidelines on the relationship of the AI Act with Annex II sectoral legislation
- Establish mechanisms for **post-market feedback collection**
- Adapt **AI systems already on the market** to requirements based on Commission Guidelines on **substantial modifications**

# Thank you! Kiitos!

## Contact us

Maria Koskinen  
AI Policy Manager  
maria@saidot.ai

## Follow us



[@ai\\_saidot](https://twitter.com/ai_saidot)



[company/saidot](https://www.linkedin.com/company/saidot)



[saidot.ai](https://www.saidot.ai)