Deloitte.

The EU AI Act

Executive Summary

February 2025

1 Overview

In force from 1st August 2024 with the first conditions effective from 2nd February 2025

Purpose
To promote human-centric and trustworthy AI – protecting health, safety, fundamental rights, democracy and the rule of law, and the environment from potential harmful effects – while supporting innovation, particularly among European SMEs.

Scope
AI (systems, models) deployed in the European Union. Extra territorial reach if AI system or its output affects individuals within the EU.

Approach
A risk-based approach, categorizing AI systems by use case into categories “unacceptable risk”, “high risk” or other which drive compliance obligations (decommissioning prohibition, declaration of conformity, transparency requirements, or voluntary standards).

General-Purpose AI models (GPAI)
GPAI models provide base capabilities and can be applied widely to many use cases. As such, GPAI models are risk-categorized using alternate criteria and generally subject to enhanced transparency obligations.

Compliance
Providers (developers) must establish Quality Management and Risk Management Systems. They must validate high-risk AI systems against trustworthy AI principles prior to issuing a Declaration of Conformity, placing onto the EU market and registering in the public EU database. Post-launch, deployers mustlog issues into the EU database and providers update conformity assessments throughout the lifecycle.

Timing
The European Parliament and the Council have approved the final text. The AI Act was translated into all EU languages, published in the official journal, and entered into force on 1st August 2024. Its several conditions become progressively effective, for example from the 2nd February 2025 specific use cases are forbidden.

Enforcement
EU-wide authorities will coordinate across member states and follow larger topics, such as GPAI[1] and their underlying models (foundation models). National supervisors, so-called market oversight authorities, will enforce compliance, appointing “notified bodies” (permitted 3rd party auditors) to assess conformity in specific cases, engaged either by providers prior to issuing Declarations of Conformity or by the supervisor for audits.

Consequences
Fines range from 35 m€/7% global turnover (prohibited cases), 15 m€/3% (other infringements) to 7,5 m€/1% (reporting errors), as well as potential non-monetary penalties, such as forced removal of the AI system from the market.

[1] GPAI = General-Purpose AI

2 Definition of an AI System

The final, agreed definition of AI aligns closely to the OECD definition

The negotiated perimeter aligns with the OECD definition

An AI system is "a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments."

AI Systems

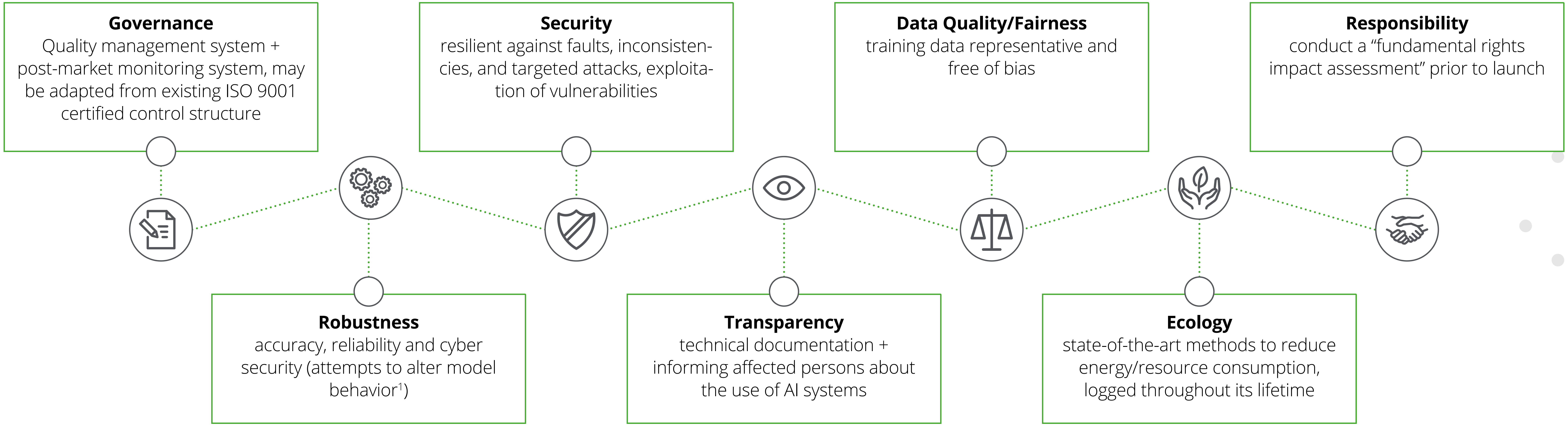
- Machine-based
- Infers outputs on input (training) data
- Adaptable after placed into operation
- Forecasts, decisions, recommendations
- Varying degrees of autonomy
- Follows explicit or implicit objectives
- Influences a physical or virtual environment

Not AI Systems

- Mathematical optimization
- Basic data processing, including descriptive analysis
- Classical heuristics
- Simple forecasts

3 Trustworthy AI at the Core

The principles of Trustworthy AI are the paradigm for AI quality and form the bedrock for emerging technical standards



Governance
Quality management system + post-market monitoring system, may be adapted from existing ISO 9001 certified control structure

Robustness
accuracy, reliability and cyber security (attempts to alter model behavior[1])

Security
resilient against faults, inconsistencies, and targeted attacks, exploitation of vulnerabilities

Transparency
technical documentation + informing affected persons about the use of AI systems

Data Quality/Fairness
training data representative and free of bias

Ecology
state-of-the-art methods to reduce energy/resource consumption, logged throughout its lifetime

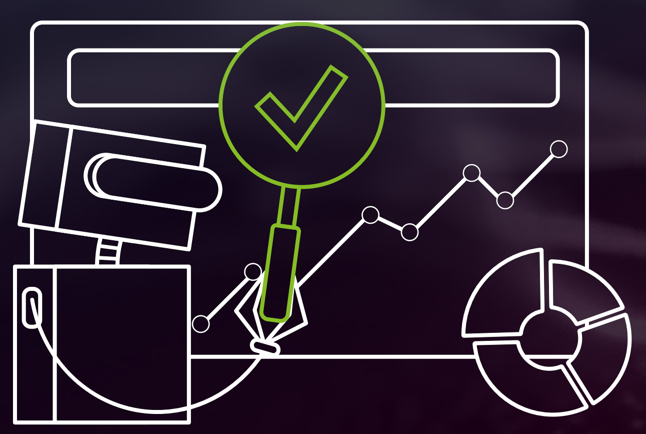
Responsibility
conduct a "fundamental rights impact assessment" prior to launch

Beyond principles, European standards setters (e.g., CEN/CENELEC) will more concretely define how each of these principles translate into technical standards, against which AI systems must demonstrate compliance by law.

[1] ... Data poisoning, model poisoning, adversarial examples, model evasion, confidentiality attacks

4 Application Scope

Any AI systems affecting European citizens[1] …

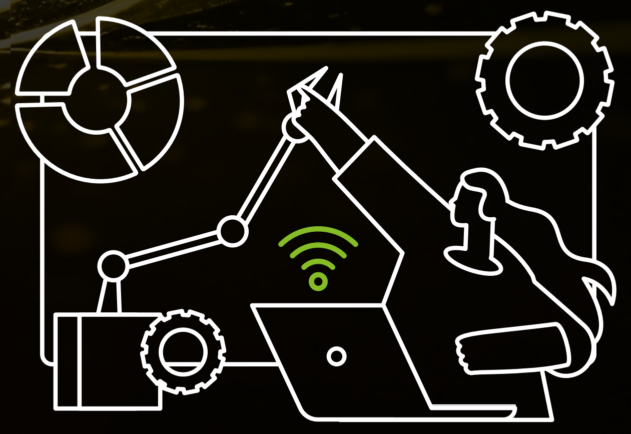


AI Systems…

- AI deployed in the EU
- Hosted outside the EU but deployed in the EU/accessed by European citizens (foundation models, very large online platforms = social media AI, or the use of their outputs)

Actors…

- Providers (Developers)
- Importers
- Distributors
- Deployers
- Authorized Representatives



Exceptions…

- Public authorities outside the EU
- R&D prior to market placement
- Academic research or personal use
- Open-source licensed[2]

HRAI systems already on the market prior to the AI Act, unless:

- they undergo "significant change" afterward
- they constitute a GPAI

GPAI already on the market prior to the AI Act are subject to an extended implementation period of 36 months.[3]

[1] not necessarily hosted or operated in Europe
[2] Unless they are (a) used in a high-risk AI system, (b) foundation models
[3] Article 111

5 Enforcement

Each Member State shall establish a national supervisor, while the EU AI Office coordinates across borders

EU-wide

European AI Office

- A new body within the European Commission
- To coordinate the implementation of the Act throughout the EU Member States
- To monitor development of foundation models & general-purpose AI

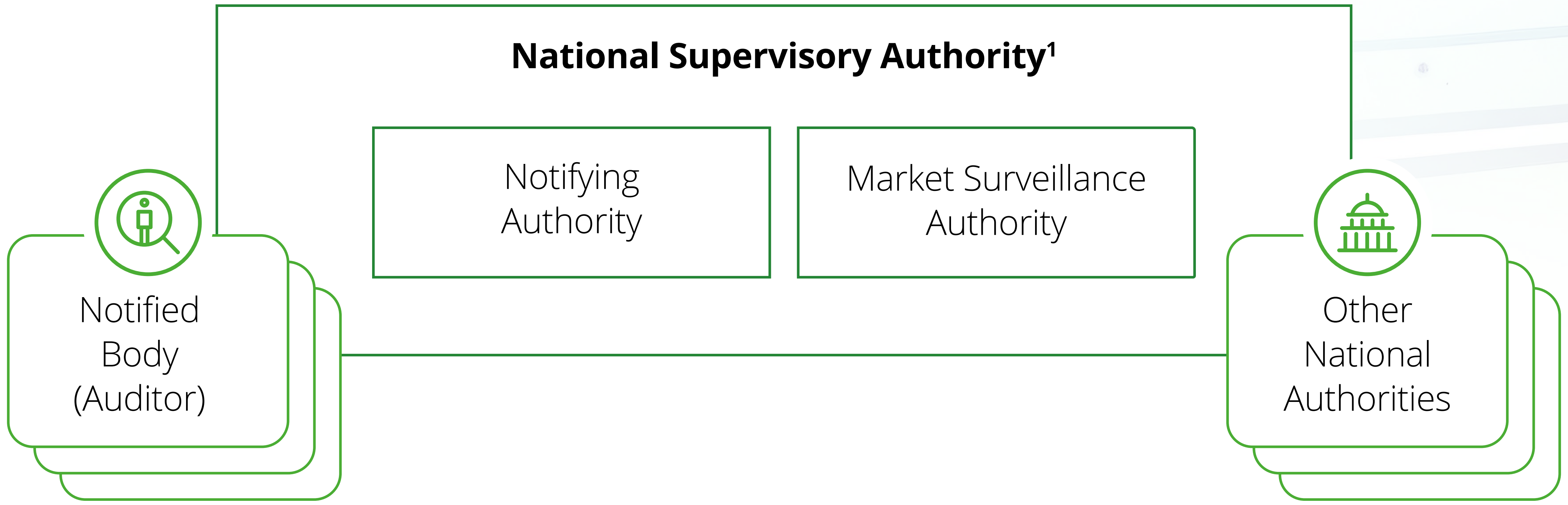
Advisory Forum

- Composed of stakeholders from the business sector and civil society
- To provide a wide spectrum of viewpoints for consideration in the implementation process

Scientific Panel

- Consisting of independent experts
- To identify systemic risks, offer guidance on model classifications, ensure enforcement based on latest scientific understanding

At a national level


National Supervisory Authority[1]
Notifying Authority
Market Surveillance Authority
Notified Body (Auditor)
Other National Authorities

- National supervisor – enforcing compliance
- Other competent authorities – i.e., to supervise other, sector-specific regulatory requirements
- Notifying authority – ensuring conformity assessments conducted properly & timely
- Notified bodies – accredited to conduct conformity assessments

[1] for example, in Germany, the role is shared between existing authorities: BaFin for Financial Services, the Bundesnetzagentur (BnetzA) for all other sectors

6 General-Purpose AI (GPAI) and Foundation Models

Contrary to SPAI[1], the risk of GPAI is measured by the power/sophistication of the foundation model

Single Purpose AI Systems

- Detecting anomalies
- Recommender systems
- Predicting trends
- Classification (good/bad)
- Computer vision
- Reading text

General-Purpose AI Models (GPAI)

- **Large Language Model**
  $< 10^{25}$ FLOP[2] training effort
  (conversation, text, code, …)
- **Diffusion Model**
  $< 10^{25}$ FLOP training effort
  (images, video, audio)
- **“High-Impact” Foundation Model**
  $> 10^{25}$ FLOP training effort
  (multi-modal, reasoning engines, …)

… capabilities alone pose a systemic risk ⚠

Responsible parties

Developers of general-purpose AI models (foundation models), deployers if core capabilities substantially altered

Risk categorization

Differentiation in between GPAI and “high-impact” GPAI posing systemic risk

GPAI

Transparency obligations, including technical documentation, a strategy to ensure training data respecting copyrights, watermarking of AI generated content (especially training data)

Systemic Risk GPAI

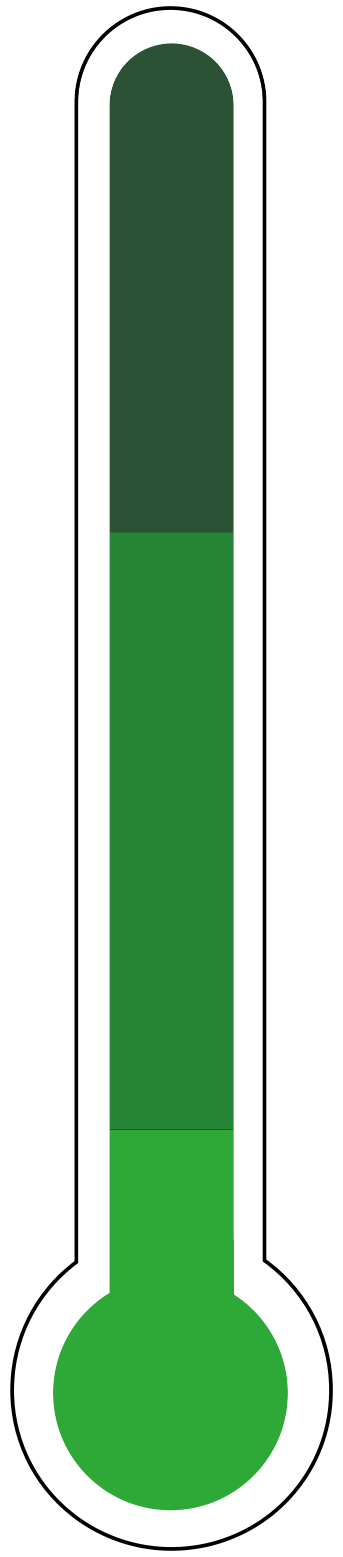
Developers of general-purpose AI models (foundation models), deployers if core capabilities substantially altered

- perform model evaluations
- assess and mitigate systemic risks
- conduct adversarial testing
- report to the Commission on serious incidents
- ensure cyber security
- report on their energy efficiency
- adhere to codes of practice until harmonized EU standards published

[1] SPAI = Single Purpose AI
[2] FLOP = Floating Point Operations

7 Risk Classification of AI Systems

A differentiated approach depending on the perceived risk to EU citizens



Risk Category	Consequence	Timeline
Unacceptable Risk, i.e. • Social scoring, profiling, manipulation • Realtime biometric identification • Emotion sensing at the workplace[1]	**Prohibited**	02.02.2025
High-Risk AI Systems, i.e. • EU Product Safety list[2] (Annex I) • Impairment of health, security orfundamental rights of EU citizens • Specific cases listed in Annex III	**i.e. Governance, Conformity Assessment, Monitoring, Logging**	02.08.2026 – in general 02.08.2027 – Annex I (EU product safety) 31.12.2030 – Annex X (specific EU systems) GPAI: 02.08.2025
Other AI Systems • Chatbots • Deep fakes • Internal models • Procedural tasks	**Freiwillige Standards**	n.a.

additionally...

Transparency Obligations, also applying to GPAI

02.08.2026

€ Penalties up to:

35 m€ or 7% global annual turnover

- Unacceptable use cases (Article 5)

15 m€ or 3% global annual turnover

- Other breaches[3]

7,5 m€ or 1% global annual turnover

- Incorrect, incomplete, misleading response to a NCA/NB[3] request

[1] Except for medical or security reasons

[2] Annex I

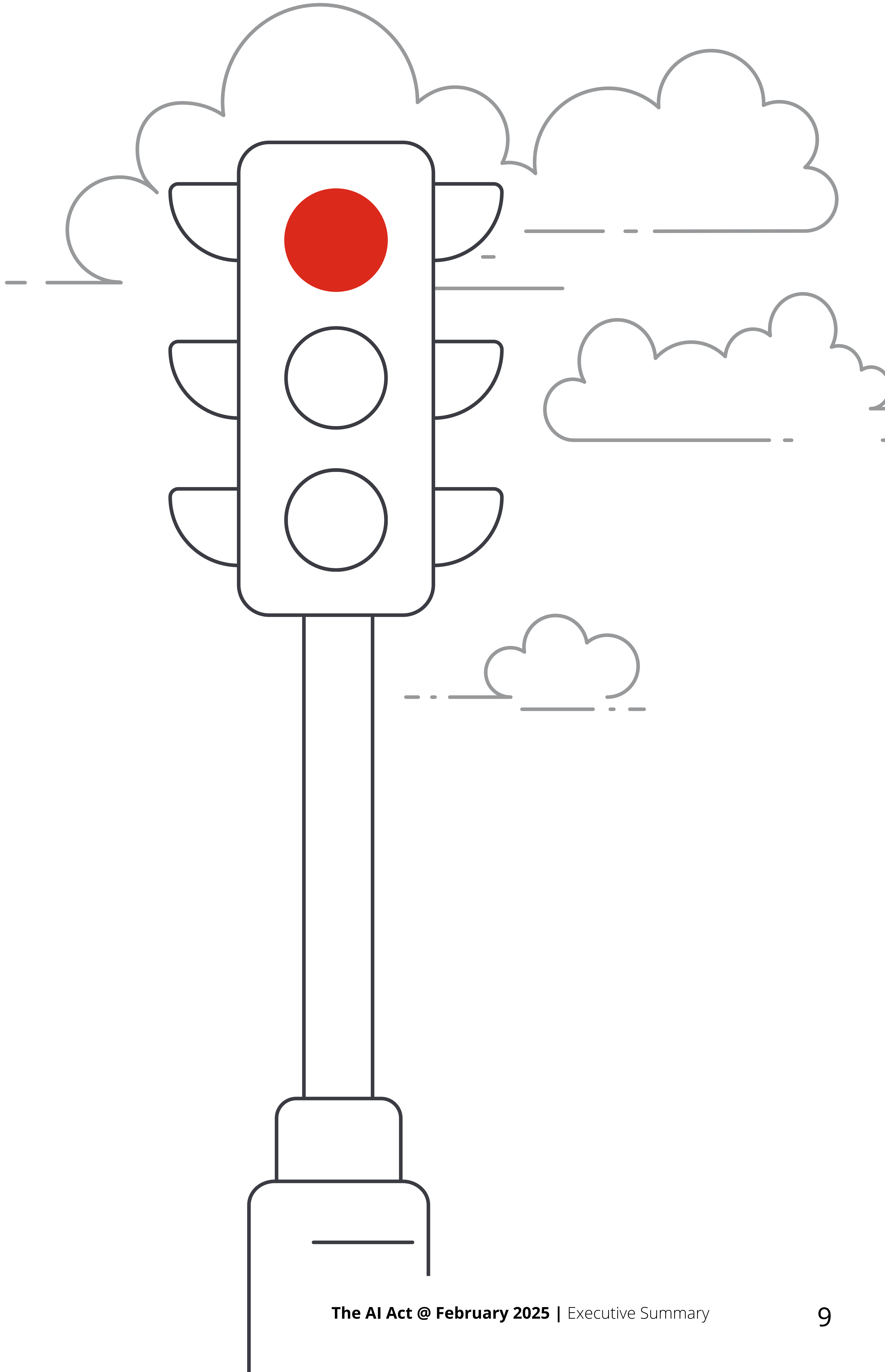
[3] e.g. failure to comply to major requirements of HRAI (Articles 16 and 26)

[4] NCA = National Competent Authority, NB = Notified Body

HRAI systems already on the market prior to the AI Act will be technically excluded from the scope, unless:
a) they undergo "significant change" afterward
b) they constitute a GPAI.
GPAI on the market prior to 02.08.2024 enjoy an extended implementation period until 02.08.2027

8 Unacceptable Risk = Prohibited

Specific cases are considered to violate fundamental human rights and are thus forbidden[1] applications of AI



Mass surveillance

Untargeted scraping of facial images from internet or CCTV for databases (privacy); ex-post remote biometric identification[1]

Biometric categorization

Profiling using sensitive characteristics (demographics)[2]

Emotion recognition

At the workplace[3] or in schools

Social scoring

Based on behavior or personal characteristics where all conditions of Article 5(1)(c) are cumulatively fulfilled

Behavioral manipulation

To circumvent free will of individuals – particularly from vulnerable groups[4]

Implementation timeframe

02.02.2025 – regardless whether placed on the market prior to this date

[1] Except for personal use or when open source (Article 2), unless corresponding to a use case falling under Article 5

[2] Exceptions: crowd control; law enforcement upon prior judicial authorization for the targeted search of persons convicted or suspected criminal activity, or to detect suspicious customers (at a bank, at a supermarket) to identify early-warning indicators of an impending robbery – provided employees are not tracked and sufficient safeguards are in place

[3] E.g., socio-economic status, gender, ethnicity, citizenship status, philosophical beliefs, religion, political orientation, sexual orientation

[4] Notable exceptions: client contact in a commercial context (Marketing); tone recognition as an aid to a clerk, such as a call center agent, to cope with certain angry customers; to personalize services such as online language courses or care robots; in education provided evaluation/certification of the subject is unaffected; at the workplace where strictly necessary for medical or health safety reasons

[5] Vulnerable individuals = particularly children, elderly, under-educated

9 High Risk = Conformity

Specific cases are considered to pose threats to safety or fundamental rights, depending on their implementation

1. Products listed under EU safety legislation[1]

2. Annex III - Corresponding to eight specific areas:

1. Management and operation of critical infrastructure
2. Employment, worker management, access to self-employment
3. Access to essential private & public services/benefits[2]
4. Law enforcement[3]
5. Migration, asylum and border control management
6. Education and vocational training
7. Administration of justice[4] and democratic process
8. Biometric identification/ categorization[5]

Exception: AI models supporting only procedural or narrowly defined tasks of otherwise high-risk use cases are not considered high-risk

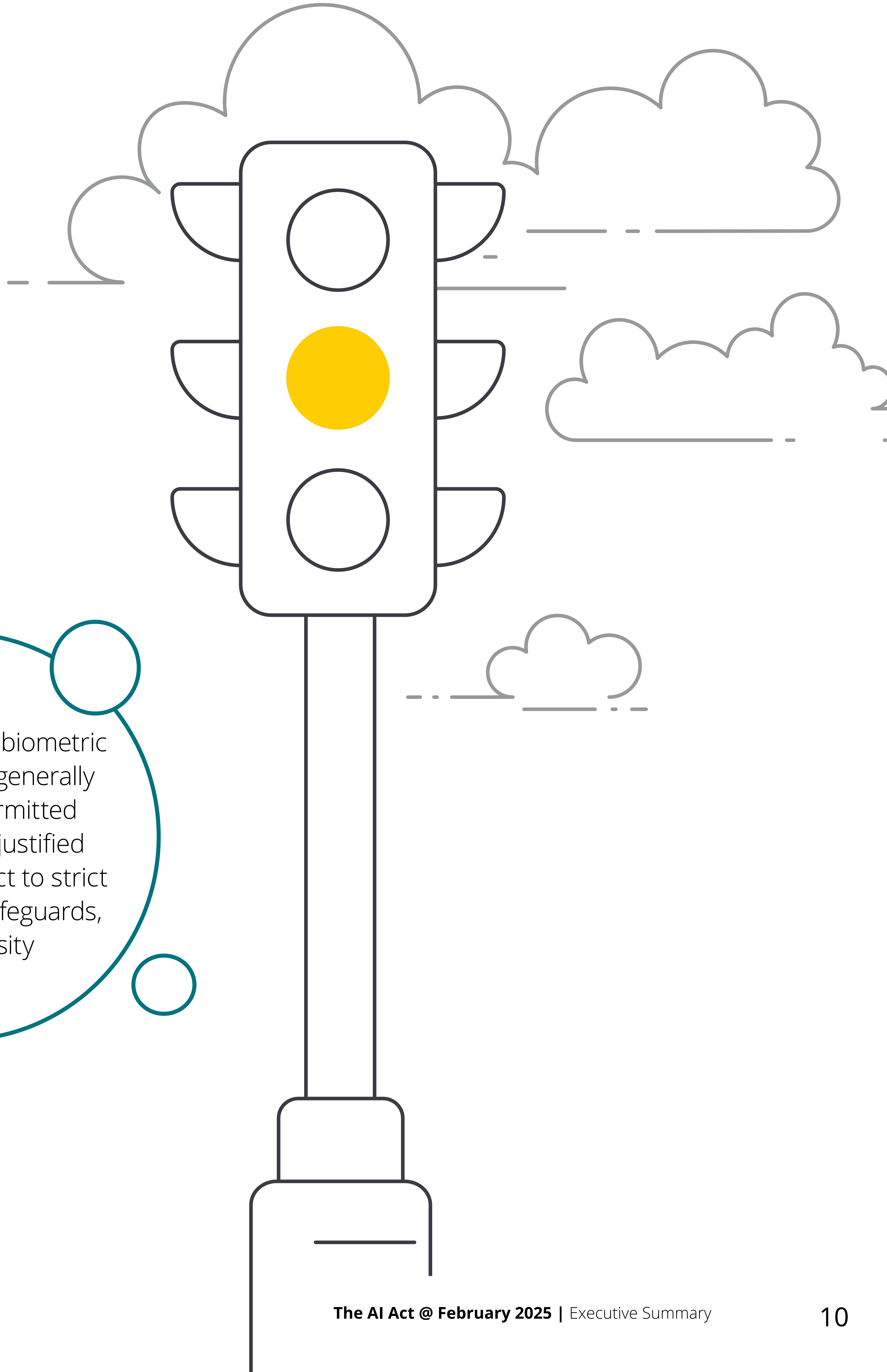
Implementation timeframe

02.08.2026 – in general

02.08.2027 – for systems on the EU Product Safety list (Annex I)

31.12.2030 – for integration of AI into specific (complex) EU-wide systems in operation prior to 02.08.2027 (Annex X)

Real-timeremote biometric identification is generally prohibited, permitted only in limited, justified exceptions subject to strict authorization, safeguards, and necessity



[1] E.g., machinery, toys, aviation, cars, medical devices and lifts.

[2] Such as insurance, credit, housing, utilities, health care, internet access, ... Exception: detecting fraud in application to such services not considered high-risk

[3] Except for administrative proceedings to detect, prevent, prosecute criminal activity

[4] AI may support, but not replace human decision-making for interpretation of law. Exception to AI used for administrative support processes without directly affecting the outcome of justice.

[5] Targeted search of victims, prevention of specific & present terrorism threat, localization or identification of a person convicted or suspected of specific, serious crimes

10 Transparency Risk, No Risk/Other Risk

AI systems which do not negatively impact natural persons, differentiated directly interacting with them or not

Transparency Risk – transparency obligations if affecting EU citizens

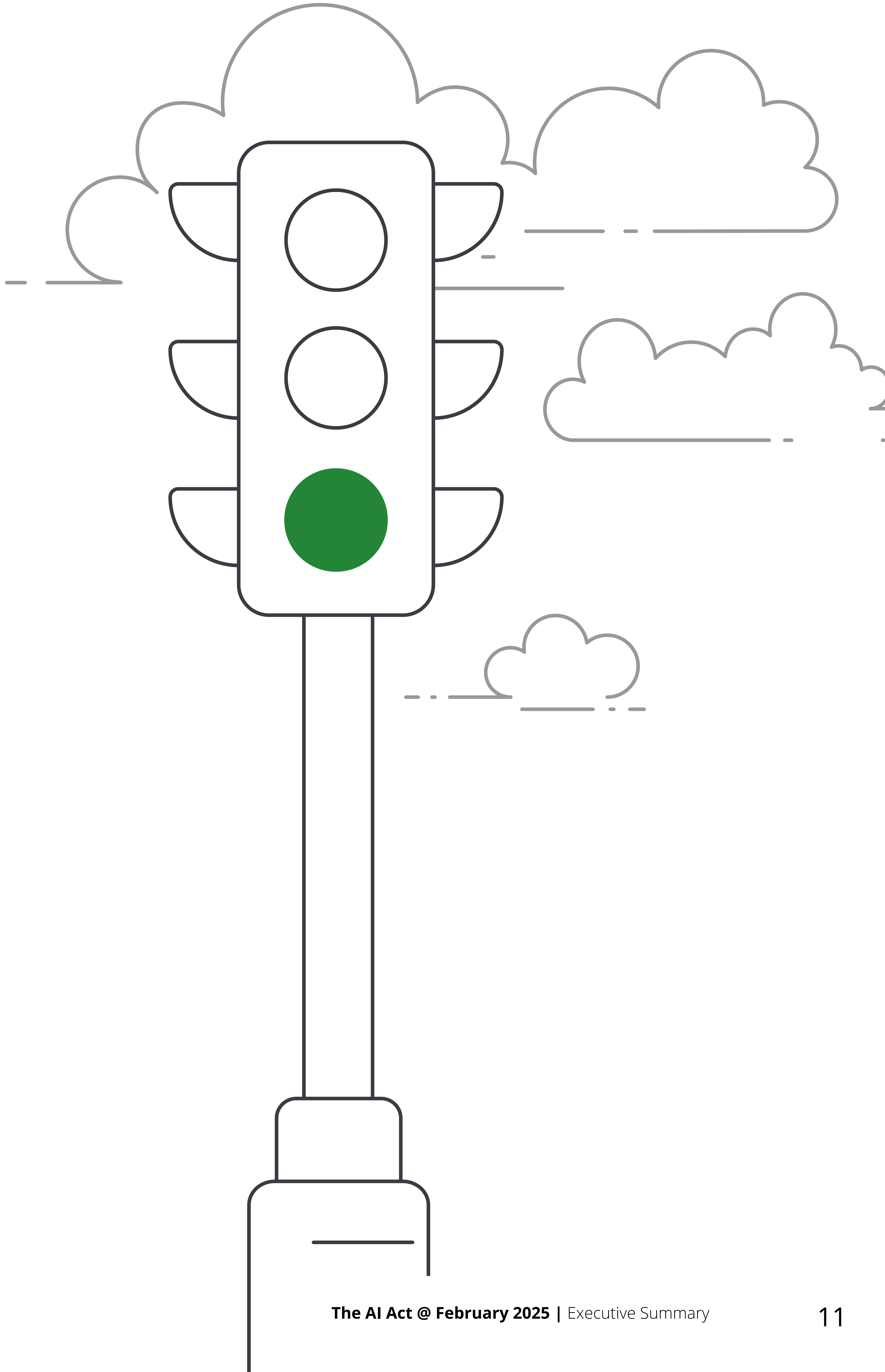
Subject to transparency obligations, namely informing the user of interaction with an AI. Examples include...

- chatbots
- deep fakes (manipulation of image, audio, video)

No or other Risk – only voluntary standards if internal models or limited to procedural tasks

Only subject to voluntary quality standards. Examples include...

- internal rating models
- recommender systems helping internal staff



11 Obligations

Providers[1] of high-risk AI must demonstrate conformity prior to market placement, maintain throughout the lifecycle

1. Quality Management System

- Technical specifications & standards
- Quality[2] design & development
- Validation and testing
- Data management & engineering
- Post-deployment monitoring & logging
- Documentation[3] & record keeping[4]
- Human oversight, roles & responsibilities
- Compliance strategy
- Establishing a risk management system

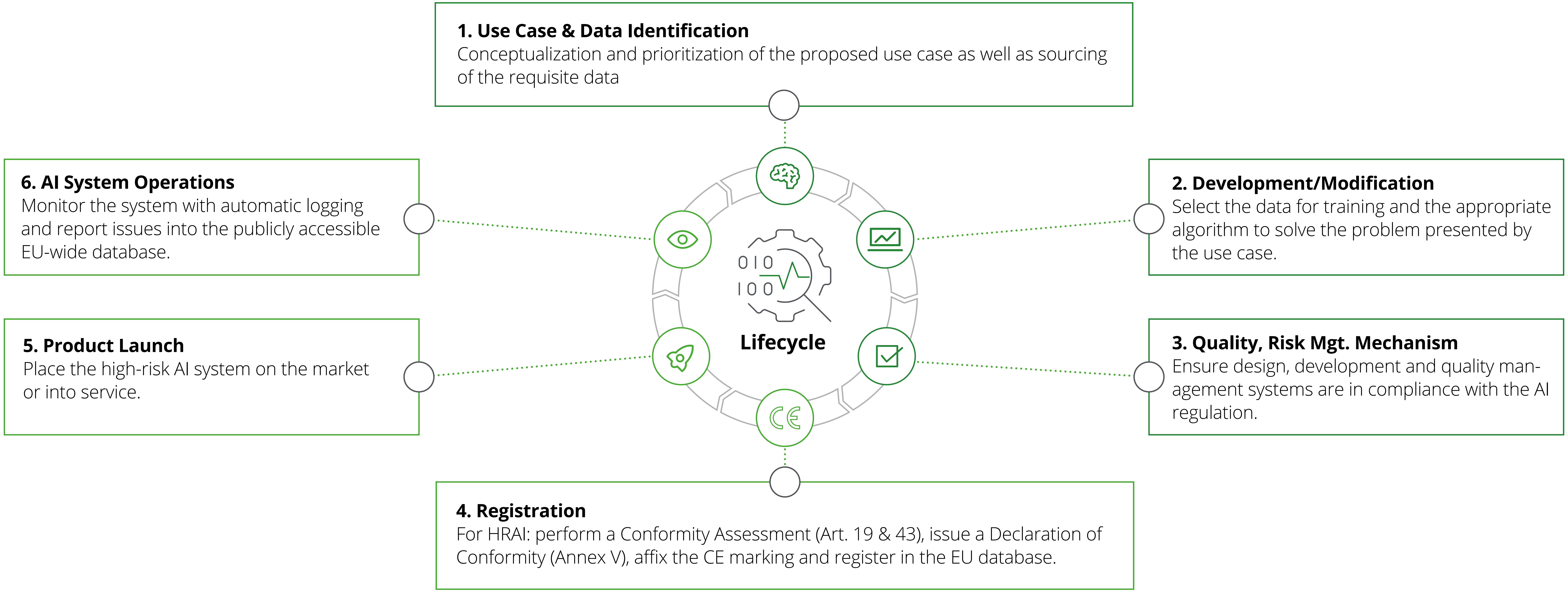
2. Risk Management System

- Anticipatory risk identification
- Risk evaluation, impact assessment
- Risk management measures to a level of acceptable residual risk
- Technically feasible risk elimination
- Technical level mitigation & control
- Continuous review, iterative process
- Suitable use via training & instructions
- Testing & documentation of the RMS

[1] Developers or those commissioning development by IT experts on their behalf
[2] Defined by principles of Trustworthy AI: data quality, transparency, robustness, accuracy, cyber security
[3] Technical specifications, Declaration of Conformity, Fundamental Rights Impact Assessment, CE-Marking …
[4] Registration in the EU database, reporting of issues, …

12 Lifecycle

Conformity to the quality, governance, and documentation standards of the AI Act is a continuous process to be maintained throughout the product lifecycle



1. Use Case & Data Identification
Conceptualization and prioritization of the proposed use case as well as sourcing of the requisite data

2. Development/Modification
Select the data for training and the appropriate algorithm to solve the problem presented by the use case.

3. Quality, Risk Mgt. Mechanism
Ensure design, development and quality management systems are in compliance with the AI regulation.

4. Registration
For HRAI: perform a Conformity Assessment (Art. 19 & 43), issue a Declaration of Conformity (Annex V), affix the CE marking and register in the EU database.

5. Product Launch
Place the high-risk AI system on the market or into service.

6. AI System Operations
Monitor the system with automatic logging and report issues into the publicly accessible EU-wide database.

13 Contacts

David Thogmartin
Partner
Risk Advisory
Tel: +49 211 8772 2336
dthogmartin@deloitte.de

Dr. Till Contzen
Partner
Digital Law
Tel: +49 69 71918 8439
tcontzen@deloitte.de

Torsten Berge
Director
Business Assurance
Tel: +49 151 58072499
tberge@deloitte.de

Atrak Yadegari
Director
Strategy, Brand & Reputation
Tel: +49 221 97324 521
ayadegari@deloitte.com

Deloitte.

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited (DTTL), its global network of member firms, and their related entities (collectively, the “Deloitte organization”). DTTL (also referred to as “Deloitte Global”) and each of its member firms and related entities are legally separate and independent entities, which cannot obligate or bind each other in respect of third parties. DTTL and each DTTL member firm and related entity is liable only for its own acts and omissions, and not those of each other. DTTL does not provide services to clients. Please see www.deloitte.com/de/UeberUns to learn more.

Deloitte provides industry-leading audit and assurance, tax and legal, consulting, financial advisory, and risk advisory services to nearly 90% of the Fortune Global 500® and thousands of private companies. Legal advisory services in Germany are provided by Deloitte Legal. Our people deliver measurable and lasting results that help reinforce public trust in capital markets, enable clients to transform and thrive, and lead the way toward a stronger economy, a more equitable society and a sustainable world. Building on its 175-plus year history, Deloitte spans more than 150 countries and territories. Learn how Deloitte’s approximately 457,000 people worldwide make an impact that matters at www.deloitte.com/de.

This communication contains general information only, and none of Deloitte GmbH Wirtschaftsprüfungsgesellschaft or Deloitte Touche Tohmatsu Limited (DTTL), its global network of member firms or their related entities (collectively, the “Deloitte organization”) is, by means of this communication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

No representations, warranties or undertakings (express or implied) are given as to the accuracy or completeness of the information in this communication, and none of DTTL, its member firms, related entities, employees or agents shall be liable or responsible for any loss or damage whatsoever arising directly or indirectly in connection with any person relying on this communication. DTTL and each of its member firms, and their related entities, are legally separate and independent entities.

Issue 03/2025