Fortifying the Cybersecurity of High-Risk AI Systems

Navigating the Interplay Between Regulations and Cyber Resilience in Practice

Joe Lomako, TÜV SÜD

Add value.
Inspire trust.

Cybersecurity Trends & Major Drivers in Organisations

Building Resilient ecosystems

Evolving **cyber threats, sophistication of cybercrime** (Cybercrime-as-a Service), demand for highest levels of security principles

Targeting Harmonized policy

Emerging **local and global cybersecurity regulations**, demand for harmonization in the creation of cybersecurity policy & ensuring global market access

Accelerating with Emerging technologies

Acceleration of business and growth based on **emerging technologies such as AI and Quantum** and addressing their impact on cyber risks and trust associated

Cross-Disciplinary Education & Upskilling

Fostering cross-disciplined and cross-functional education to understand & implement latest technologies and avoiding the risks at the same time

Cybersecurity & AI Regulations

Overview of regulations on cybersecurity and AI

Cybersecurity Regulations

- NIS-2 (Critical Infrastructure Protection)
- CRA (IoT device and product cybersecurity)
- RED & Machinery Regulation (Smart device cybersecurity)
- Data Act & GDPR (Data protection and flows)

Cybersecurity Standards

- ISO 27001 (IT cybersecurity)
- IEC 62443 (OT cybersecurity)
- TISAX (automotive IT cybersecurity)
- ISO/SAE 21434 (automotive cybersecurity)
- UNECE R155 & R156 (cybersecurity & software updates)

AI Regulations

- EU AI Act
- US AI Bill of Rights
- US Executive Order on AI
- China AI Measures

AI Standards

- **ISO/IEC 42001:2023** AI Management System
- **ISO/IEC 23894:2023** Guidance on AI Risk Management
- **ISO/IEC 23053:2023** Framework for AI system using Machine Learning

Major Drivers for Cybersecurity in global organisations

Navigate through a **complex regulatory landscape**: Non-harmonized global vs. local regulations

Help shape standards, and proactively **ensure Global Market Access**

Increasing **awareness** of customers for cyber risks associated with digital technologies and demand for cybersecurity

Training and upskilling in cybersecurity (standards based today)

High costs and **reputational risks** associated with an increasing number of cyberattacks

Help increase cyber resilience through **third party risk assessment services**

Requirements for High-Risk AI Systems

Technical Criteria

Robustness
Accuracy
Cybersecurity
Human Oversight

Ethical Criteria

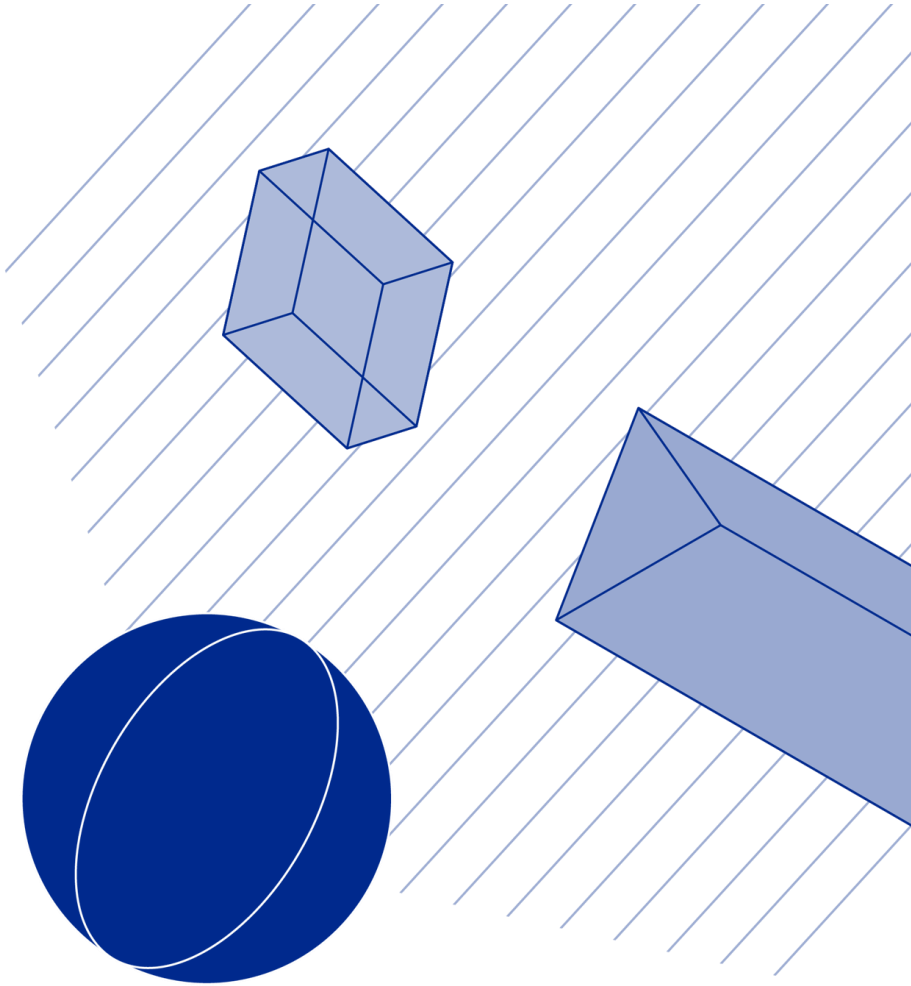
Data Governance
Bias and Non-Discrimination
Transparency

Documentation

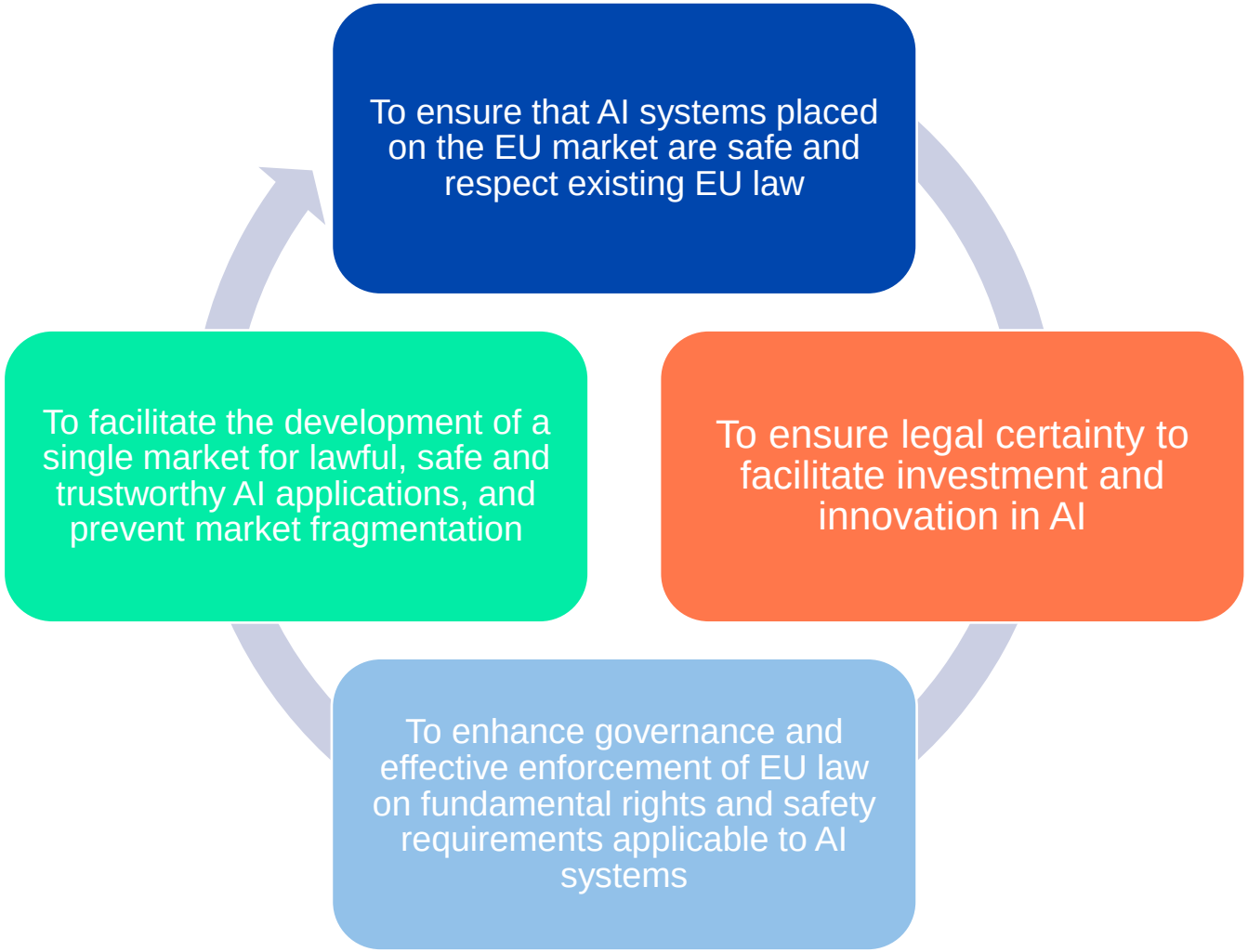
Technical Documentation
Logging
Archiving
Registration

Management Systems

Risk Management System
Quality Management System



Why EU AI Act ?


To ensure that AI systems placed on the EU market are safe and respect existing EU law
To ensure legal certainty to facilitate investment and innovation in AI
To enhance governance and effective enforcement of EU law on fundamental rights and safety requirements applicable to AI systems
To facilitate the development of a single market for lawful, safe and trustworthy AI applications, and prevent market fragmentation

EU AI Act is distinguishing different kind of AI

General Purpose AI (GPAI)

- Trained with a large amount of data using self-supervision at scale, displays generality, capable to competently perform a wide range of distinct tasks
- EU Commission can adapt definition if needed

GPAI with systemic risk

- Systemic risk: significant impact on the internal market due to its reach, actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain.

All others AI

- AI not falling in previous categories
- Downstream applications based on GPAI models

Exceptions

- Military or defense
- Developed and utilized solely for scientific research and discovery
- Research, testing, and development activities related to AI systems before their introduction to the market
- **Free and open-source** software, unless their usage would categorize them as a prohibited or high-risk AI system.

EU AI Act is an horizontal risk-based regulation

It will impact many organisations

AI Application

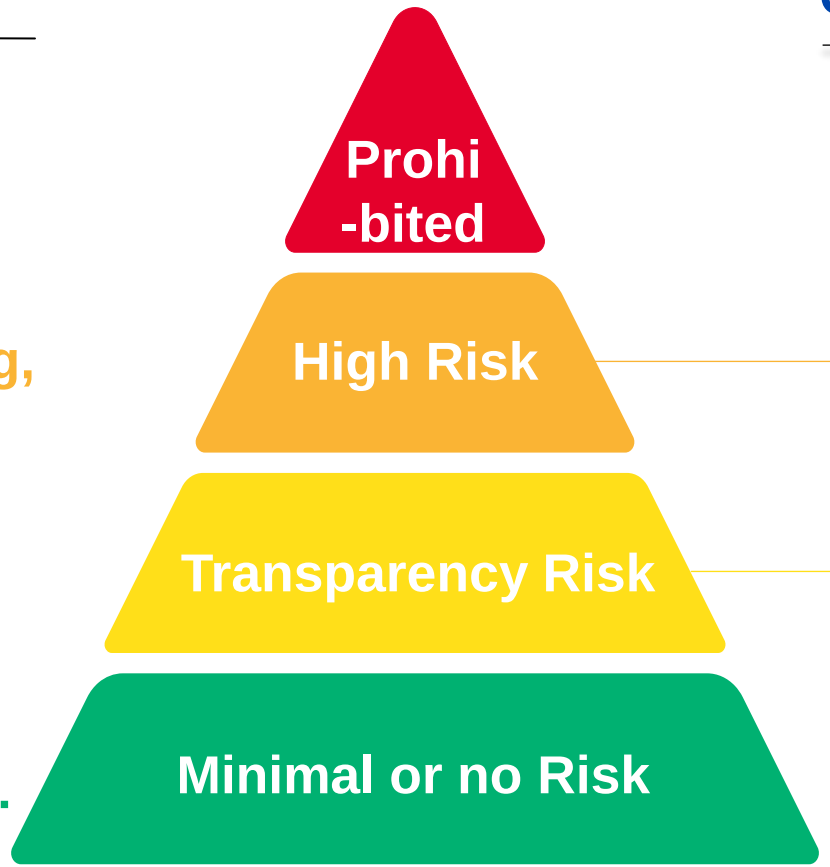
Certification needs to be compliant

Social Scoring,…

Health, Autonomous Driving, Police,…

Chat bots, Emotion recognition,…

All the rest: spams filters,…

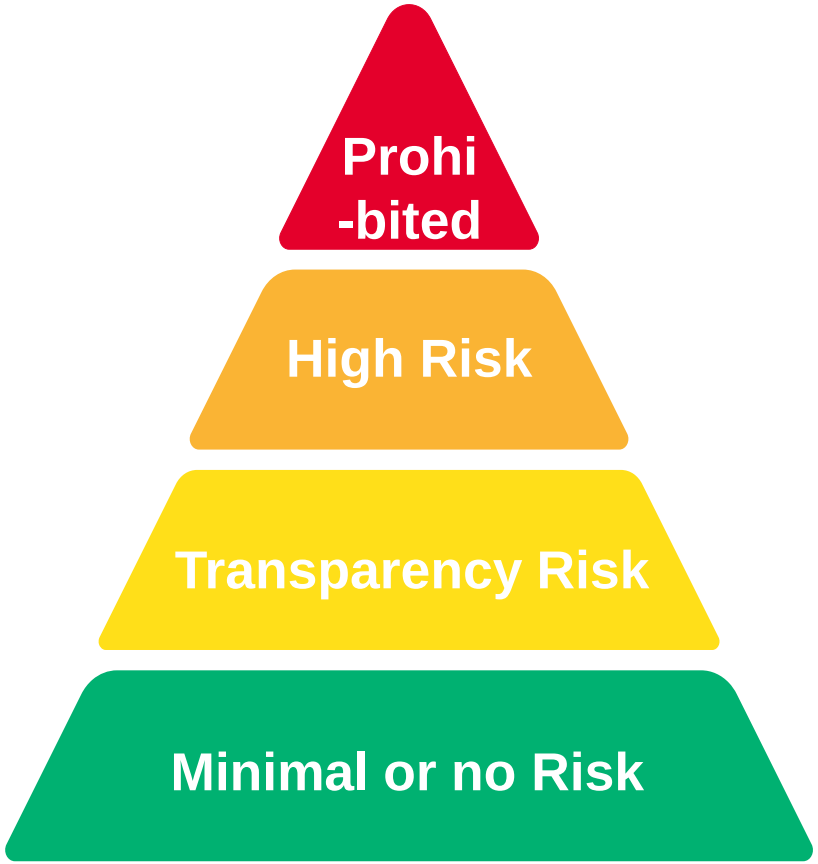


Notified Body

Self Declaration

EU AI Act categorizes AI systems

Overview of the four risk classes



Prohibited Systems | Inacceptable Risk

Putting into service prohibited (e.g. „Social Scoring“)

High-risk Systems

High Requirements (e.g. for Medical Devices)

Transparency Obligations | Limited Risk

Labelling to identify AI and its products (e.g. Chatbots)

Low Risk | Minimal Risk

No Obligations (e.g. assistance tools for text optimization)

GPAI & Foundation Models

- with Systemic Risk
  Monitoring Duties, a.o.
- without Systemic Risk
  Documentations Duties, a.o.

High-risk Systems

Annex I.A —
Sectoral Regulations
e.g. Medical Devices; Machinery; Toys; Elevators

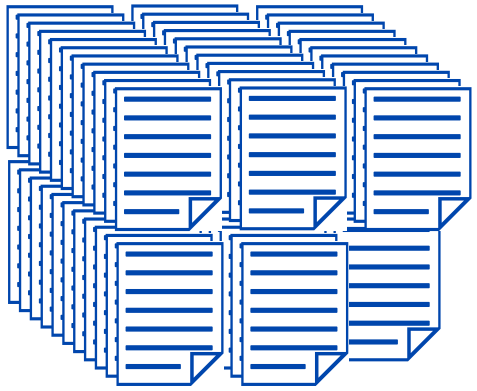
Annex III —
Area of Use
e.g. Critical infrastructure, Education, HR, Public service, etc.
**Exception: only assistance function*

Annex I.B —
(e.g. Automotive, Civil Aviation) exempted from AI Act, but: will be included via delegated acts

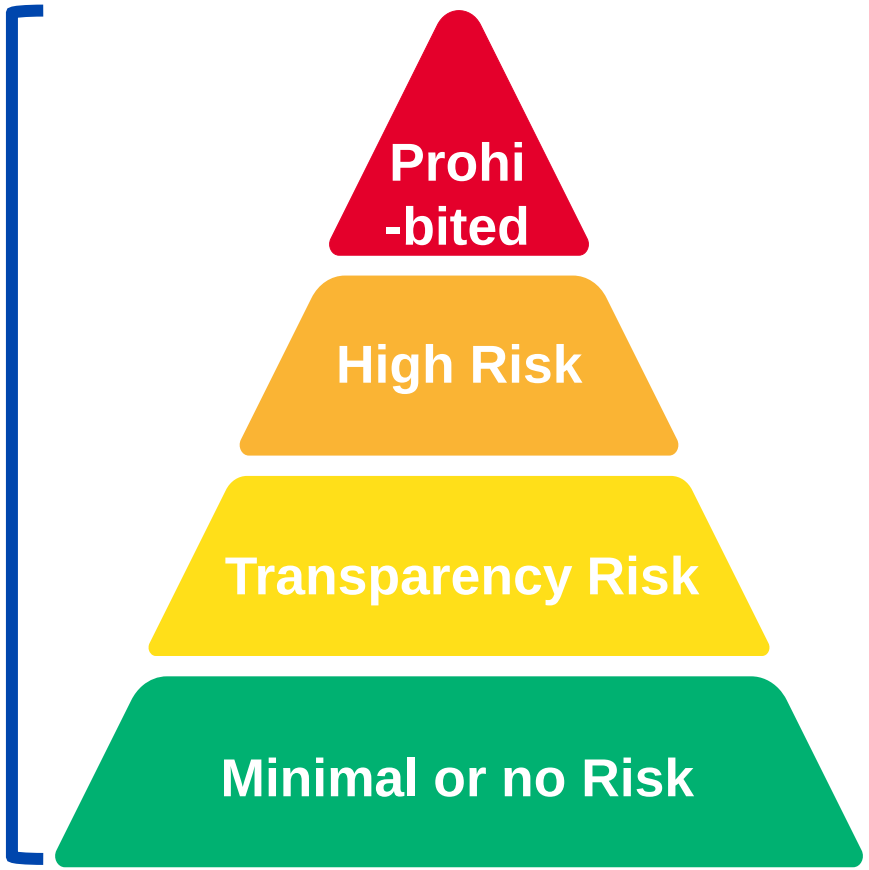
EU AI Act is a high-level regulation

Comformity with standards means being compliant

> 60 Standards



2023 - 2026

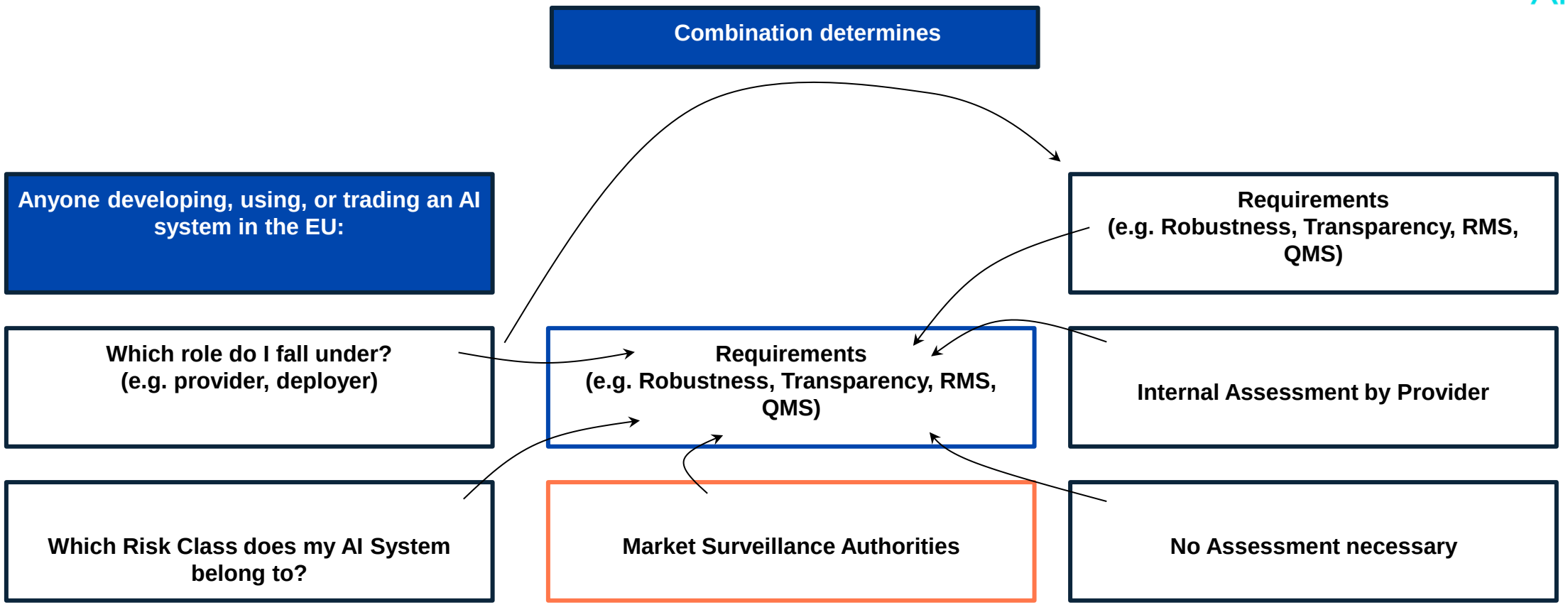


Examples of upcoming standards

- AI Quality Management
- AI Risk Management
- AI Life Cycle
- Data Life cycle
- ….

What is the AI Act's core logic?

Roles, Risk Classes, Requirements … and Assessments



The future of the EU AI Act

Time is running…

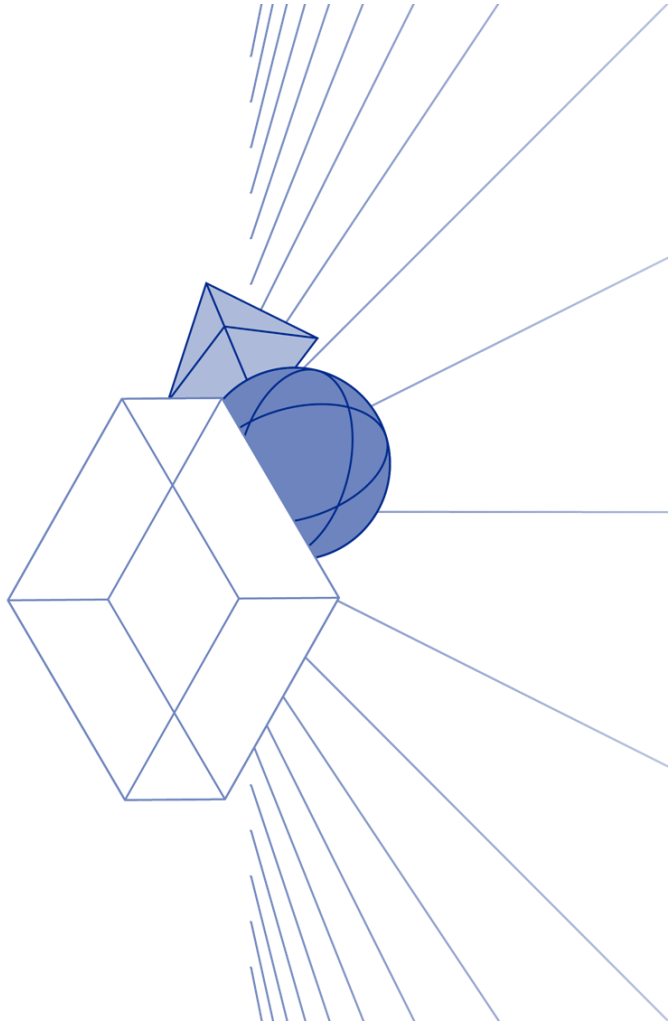
2024
Aug 1st
Entry into force,
20 days after publication

2025
Feb 2nd
Prohibition of systems
with unacceptable risk

2025
Aug 2nd
GPAI obligations apply,
Governance & Sanctions

2026
Aug 2nd
High-risk Systems of Annex III &
entire AI Act applies

2027
Aug 2nd
Art. 6 (1) / High-risk Systems
of Annex I.A



Cyber Resilience Act (CRA)

Who is affected?

- Manufacturers, developers, importers, and distributors of products with digital elements within the EU, including hardware, software, and IoT devices.
- Applies to all software and hardware products as well as ‘remote’ data processing solutions, with the exception of specifically regulated products.

At a Glance

What is it about?

- Establishing cybersecurity requirements for products with digital elements, ensuring they are designed, developed, and maintained with strong security measures to protect users and prevent cyber threats.

What is its current implementation status?

- Proposed by the European Commission in September 2022.
- Currently in the legislative process, expected to be adopted towards the end of 2024.

Cyber Resilience Act (CRA)

Key facts

Requirements for products with digital elements:

- Security by design: Products must be designed with cybersecurity in mind from the start.
- Baseline cybersecurity requirements: Obligation for conformity assessment and CE marking of products.
- Product risk categorization: Classification into three risk classes: normal, critical, highly critical.

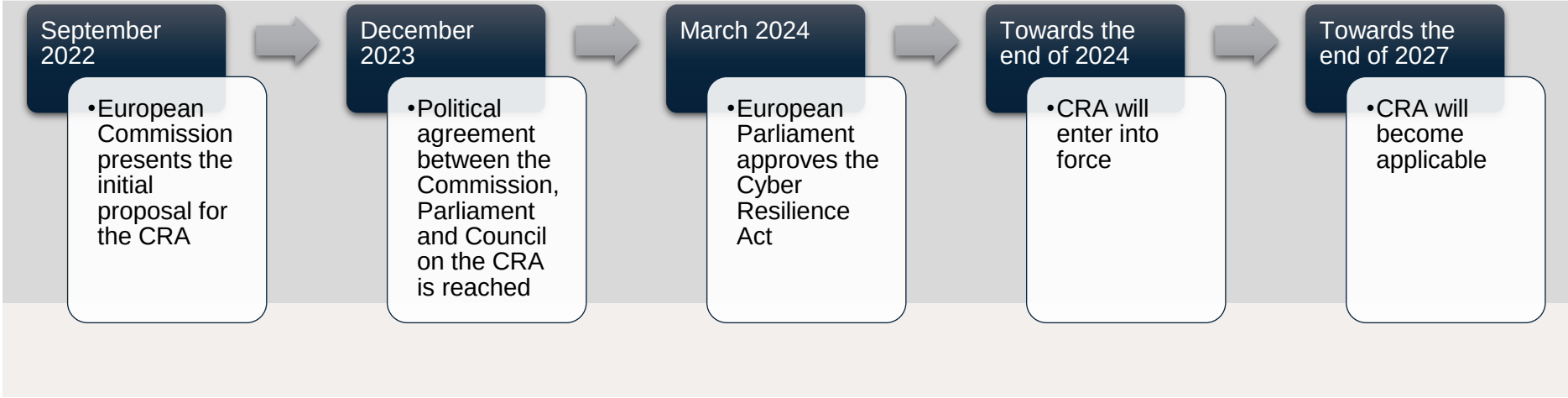
Requirements for companies active in the EU market:

- ***Manufacturers***
  - Incident reporting: Obligation to report significant cybersecurity incidents.
  - Vulnerability management: Manufacturers must provide ongoing security support (e.g. software security updates for min. 5 years), including vulnerability fixes.
  - Conformity assessment: Manufacturers must ensure their products meet certain cybersecurity standards before entering the market. This includes obligatory conformity assessments and CE marking.
  - Incident reporting: MandatoryPenalties: Non-compliance could result in significant fines.
- ***Importers***
  - Ensuring the conformity assessment, documentation and CE marking.
  - Obligation to report significant cyber security risks to the market surveillance authority.
- ***Distributors***
  - Verification of the CE marking
  - Compliance with the same obligations as importers.

Cyber Resilience Act (CRA)

Timeline

- The CRA was proposed by the European Commission in September 2022 and was approved by the European Parliament in March 2024. The CRA is set to enter into force towards the end of 2024 and will become applicable by 2027.



Challenges in Navigating Intersecting Regulations

The interplay between the EU's Cyber Resilience Act (CRA) and AI Act creates numerous challenges for compliance

1) Overlapping and Duplicative Requirements:

- Both acts require cybersecurity standards, risk management, and transparency, but with different interpretations. This can lead to redundant compliance, requiring multiple assessments and extra documentation for the same AI systems under both acts.

2) Differing Risk Classifications:

- The CRA assesses risks based on digital product security, while the AI Act focuses on AI's impact on safety and fundamental rights. Companies must navigate both frameworks, leading to potential inconsistencies and increased complexity in compliance efforts.

3) Complexity of Cybersecurity and AI-Specific Vulnerabilities:

- The CRA covers general cyber resilience, while the AI Act focuses on AI-specific vulnerabilities like data poisoning. Compliance requires addressing both broad cybersecurity and specialized AI risks, increasing complexity and requiring expert resources, which may be challenging for smaller firms.

4) Monitoring, Documentation, and Transparency Burdens:

- Both acts demand regular monitoring and documentation, but with different focuses—AI model transparency under the AI Act and lifecycle security under the CRA. This can lead to resource-heavy compliance processes, especially for high-risk AI systems that require ongoing assessments and detailed reporting.

5) Evolving Standards and Uncertainty:

- Compliance standards are evolving, with varying interpretations between cybersecurity and AI-specific regulations. This fluidity makes it hard for companies to maintain long-term compliance strategies, increasing uncertainty and compliance costs.

Commonality between CRA & AI Act

1. **Cybersecurity Requirements**: Both the CRA and the AI Act emphasize the need for robust cybersecurity standards, especially for high-risk AI systems. The AI Act specifically requires AI systems to be designed with cybersecurity in mind to ensure resilience against threats such as unauthorized alterations and malicious attacks. This alignment is reflected in the CRA's focus on overall cyber resilience.
2. **Risk-Based Approach**: The AI Act identifies high-risk AI systems that must adhere to stricter requirements. Similarly, the CRA emphasizes the importance of a risk-based approach to cybersecurity, ensuring that systems identified as high-risk are subject to rigorous controls and standards to mitigate vulnerabilities and potential threats.
3. **Accuracy, Robustness & Reliability**: The AI Act stresses that high-risk AI systems should be accurate, robust, and reliable throughout their lifecycle, maintaining an appropriate level of performance. The CRA complements this by ensuring that digital elements, including AI, uphold these characteristics to maintain cyber resilience.
4. **Secure Development Lifecycle**: Both acts require a secure design and development lifecycle for AI systems, with mechanisms in place to prevent or minimize undesirable behaviors and to manage vulnerabilities effectively. This includes technical and organizational measures that align with the requirements of both the CRA and the AI Act.
5. **Regular Monitoring & Updates**: Both acts emphasize the need for continuous monitoring, timely security updates, and patch management to ensure ongoing cybersecurity. The AI Act requires that high-risk AI systems are monitored regularly for performance and security vulnerabilities, in line with the requirements of the CRA for maintaining up-to-date and secure systems.
6. **Transparency & Compliance**: Both acts stress the importance of transparency, requiring that documentation and compliance processes be clear and well-communicated. The AI Act promotes the need for proper conformity assessments and quality management systems, ensuring that high-risk AI systems are compliant with cybersecurity standards as required by both regulations.

Quick note on Standards creation within legal initiatives and industry participation

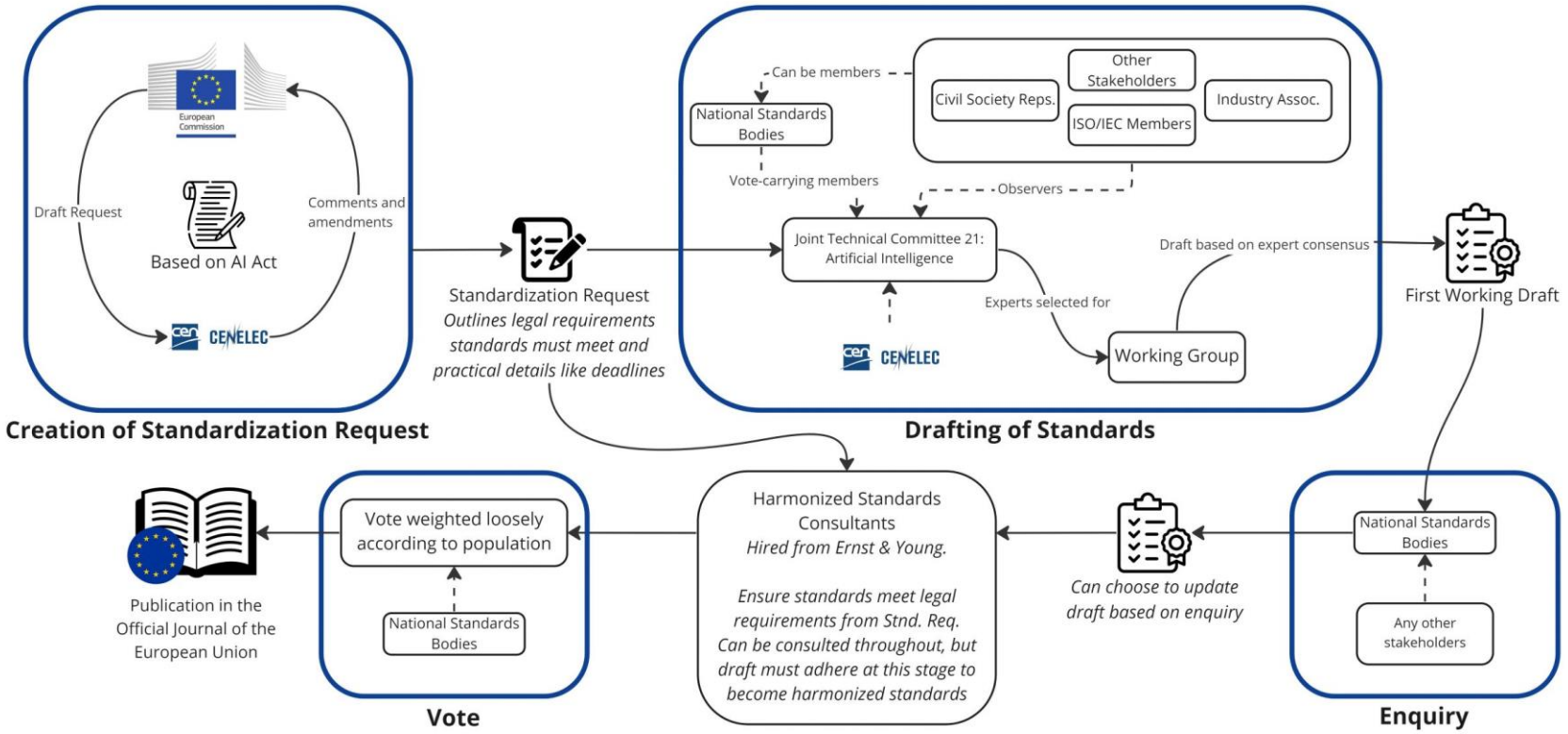
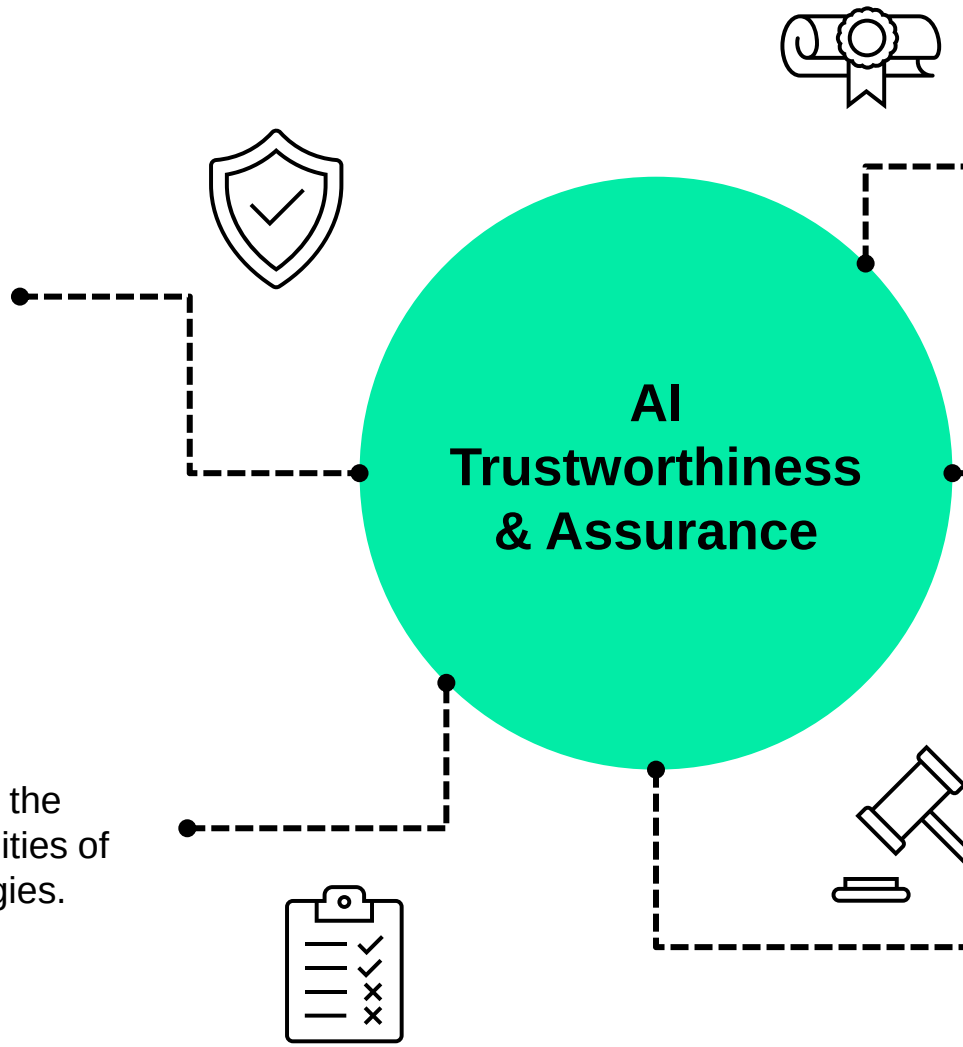


Image Source: The EU Artificial Intelligence Act. A simplified view of the creation of harmonised standards

- The **European Commission** issues a standardization **request** for the **European Standards Organisations** (ESOs), which draft the standards
- **National Standards Bodies** (NSBs) provide technical experts who work within a **technical committee** to run the drafting process for ESOs
- The technical committee forms a **Working Group** of experts to draft the document
- **TÜV SÜD** aims to be an **active** participant in such Working Groups

Regulatory Support Expectation



Risk Assessment

Identify and mitigate potential risks associated with AI security, privacy, fairness, and accountability.

Certification

Assess the conformity of management systems and organizations according to standards (e.g. ISO 42001).

Training

Train the workforce to increase employee's competence level and overall awareness for AI and related risks.

Testing

Pentesting to evaluate the security and vulnerabilities of advanced AI technologies.

Regulatory Compliance

Help to navigate through and understand the AI regulatory landscape (e.g. EU AI Act, Data Act).

Thank you!

Come speak to me!