TRIAL LECTURE

in partial fulfillment of the requirements for PhD degree

Assessing the Influence of EU's Upcoming Acts and Regulations on High-Risk AI Systems

An Analysis of Design, Implementation, Approval Processes and Usage

Rialda Spahić, PhD Candidate

Mary Ann Lundteigen	Supervisor
Vidar Hepsø	Co-supervisor
Eric Monteiro	Co-supervisor

August 29, 2023

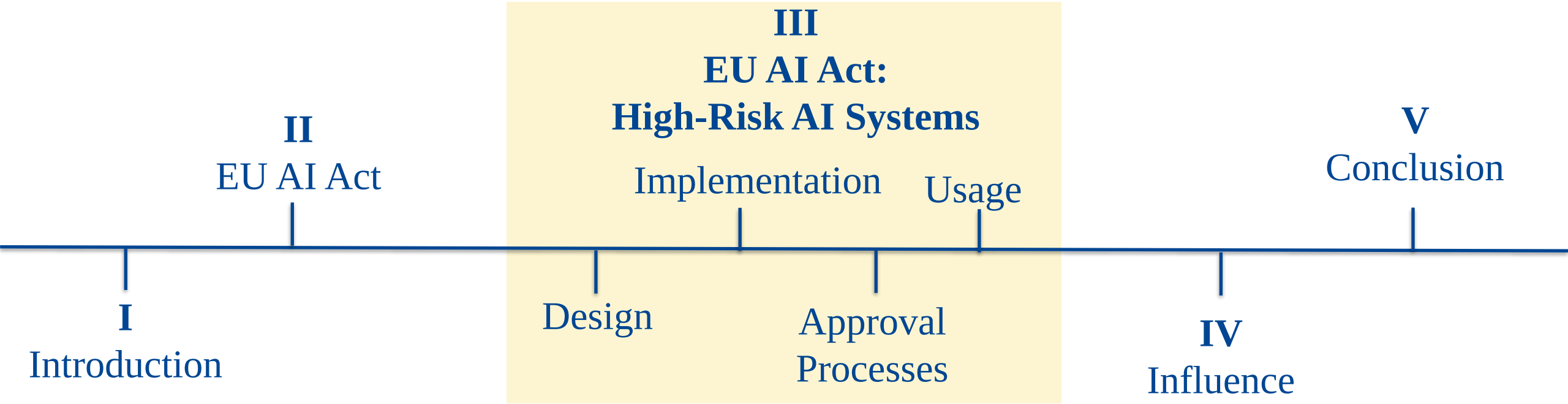
Trondheim

Norwegian University of Science and Technology

NTNU

Trial Lecture Approach

Audience: Students at Master's level



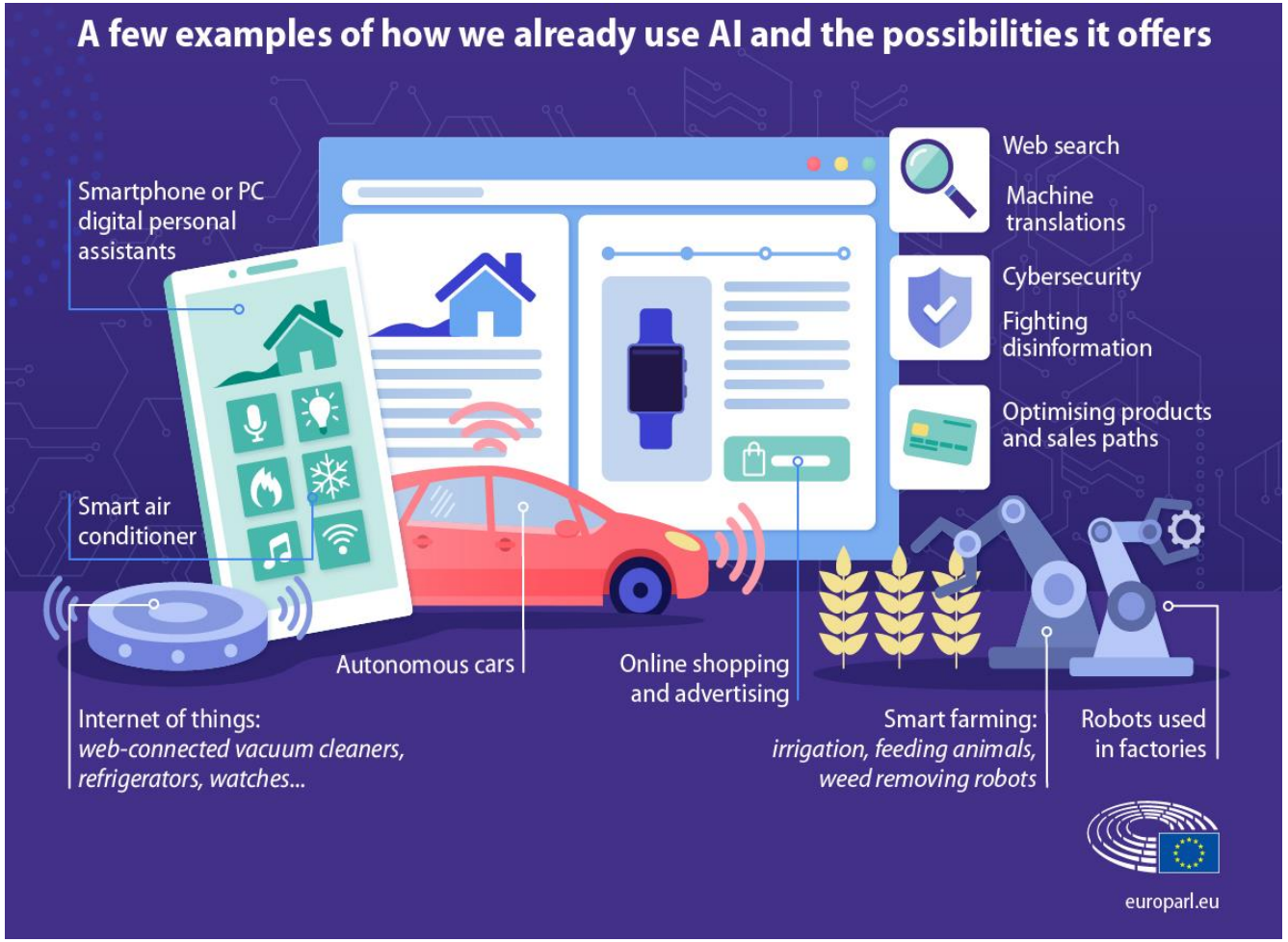
I
Introduction:

Artificial Intelligence
Risks
EU

Artificial Intelligence (AI)

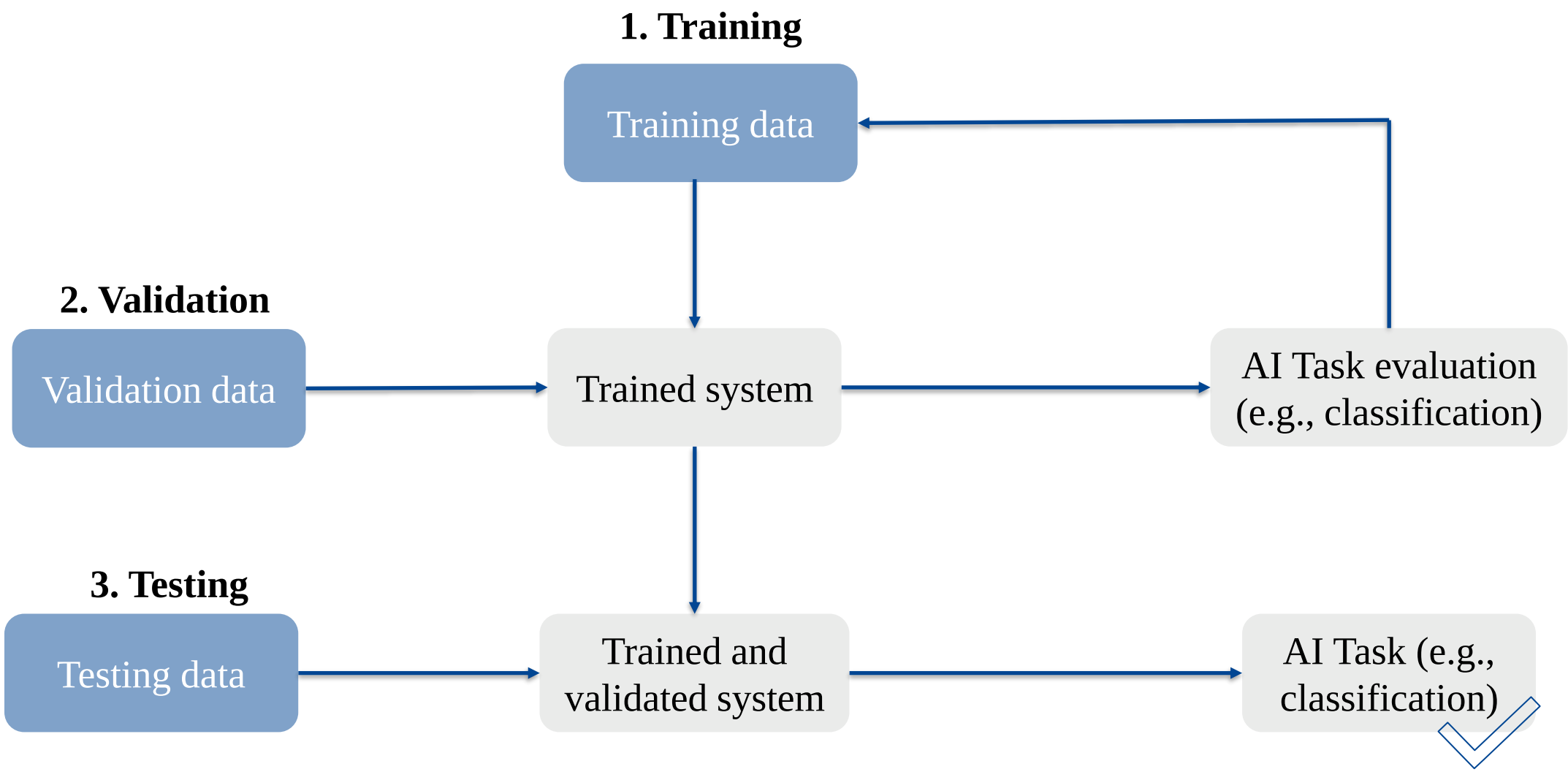
European Commission

"A software that is developed with machine learning, logic and knowledge-based or statistical approaches and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decision influencing the environments they interact with."



[Ref. European Parliament, What is AI and how is it used? 20.06.2023]

Basic AI Workflow Stages



AI creates opportunities, as well as risks for the users

- Learns through available data that may not be adequate enough.
- A network of complex algorithms, with questionable transparency.
- May result with unpredictable results.
- Operate on its own – autonomous.

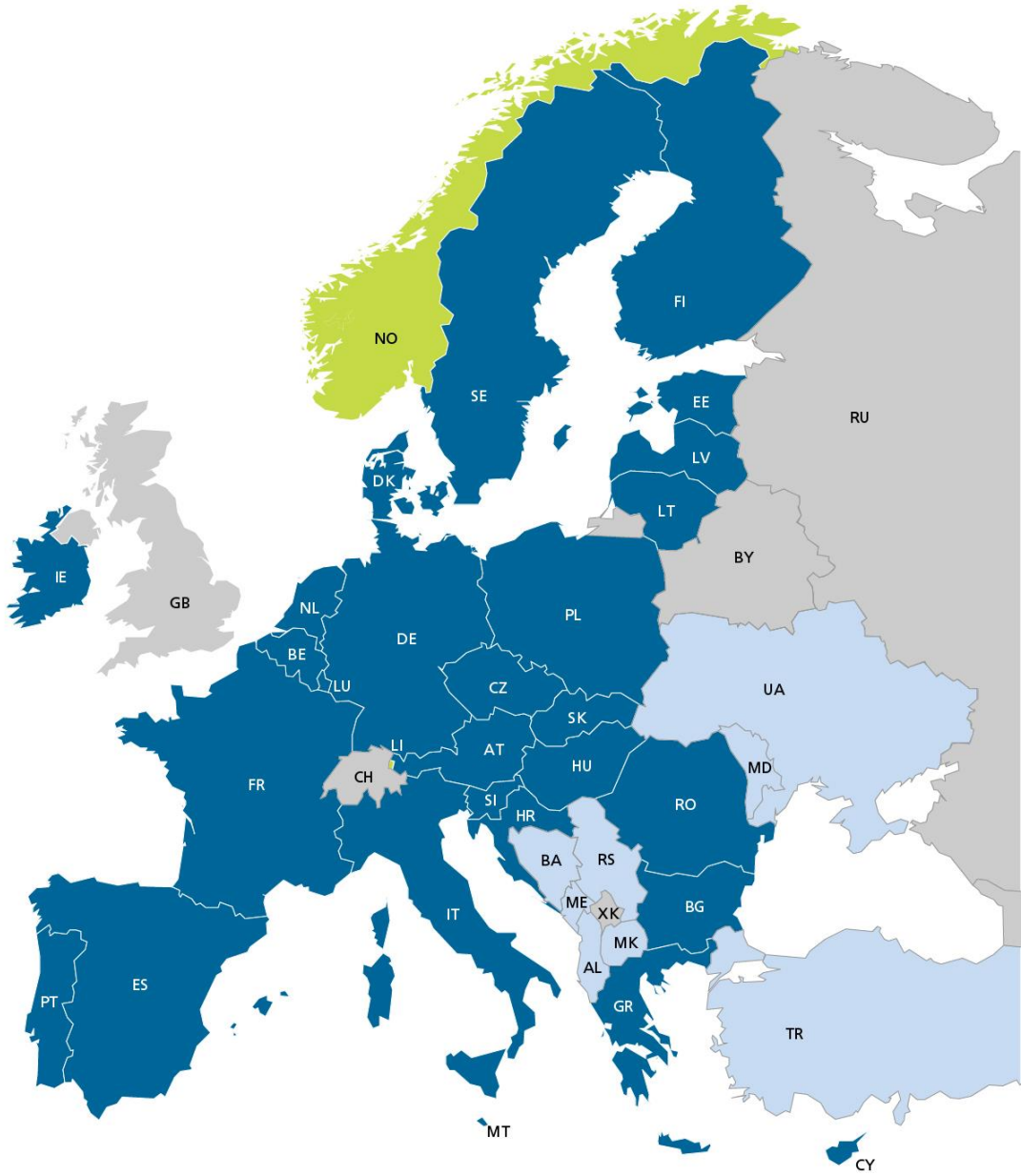
Risks

- Safety
- Legal Uncertainty
- Pervasiveness creates far-reaching implications
- Fundamental rights

European Union (EU)

Every EU citizen enjoys the same **fundamental rights*** based on the values of:

1. Equality;
2. Non-discrimination;
3. Inclusion;
4. Human dignity;
5. Freedom and democracy.



*These values are fortified and protected by the rule of law, spelled out in the EU Treaties and the Charter of Fundamental Rights.

Image Source: The Federal Council – The portal of the Swiss Government

II
EU AI ACT:

Objectives and Scope
Developments
Risk-based Approach

EU AI Act
Objectives and Scope

General objective: to ensure the proper functioning of the single market by creating the conditions for the development and use of trustworthy artificial intelligence in the Union.

- Respect existing laws and fundamentals rights of EU
- Ensure legal certainty
- Facilitate investment and innovation

- Applicable to all value chain participants:
  - Providers of AI systems to the EU market (irrespective of their place of establishment)
  - Users of AI systems within EU

EU AI Act
Developments

…

April 2021

European Commission proposed the first EU Regulatory Framework for AI.

June 14, 2023

Members of European Parliament adopted Parliament's negotiating position on the AI Act.

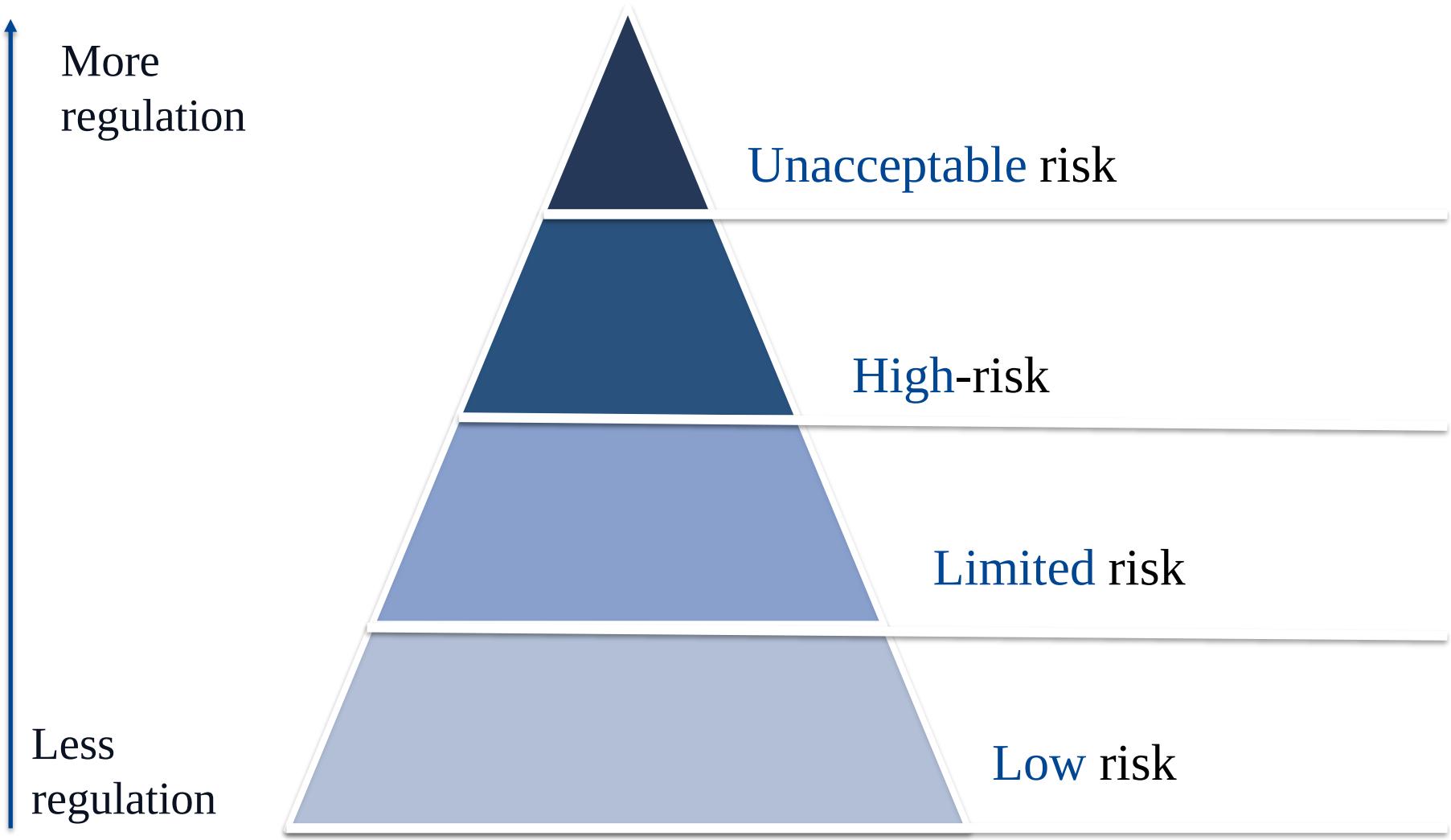
End of 2023

New updates for the final agreement.

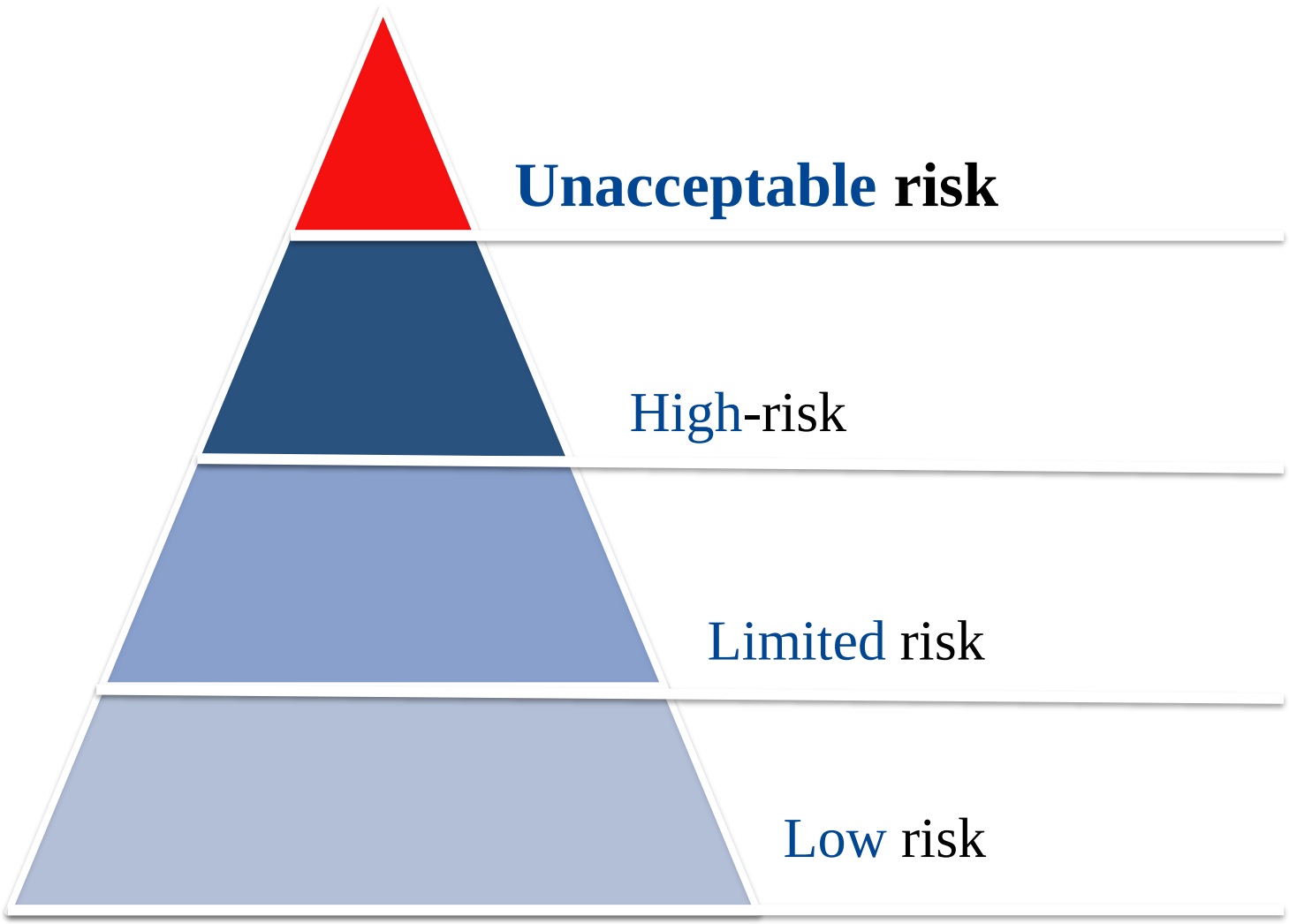
The EU AI Act is nearing final stages of adoption and will be the world's first piece of legislation governing AI.

It is projected to become a regulation in early 2024, with implementation beginning in 2027.

Risk-based approach to AI systems



Risk-based approach to AI systems

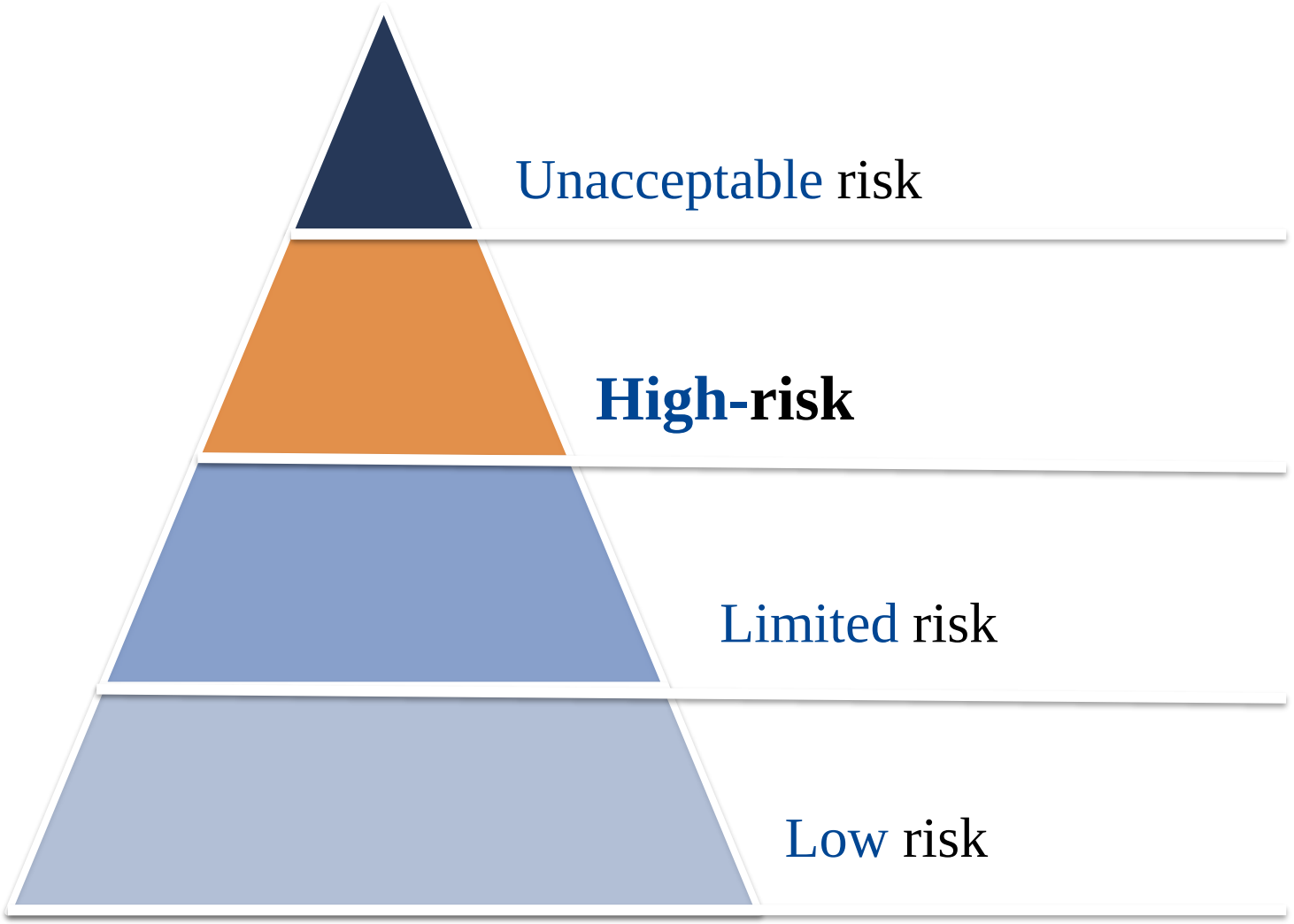


Systems with **unacceptable risk** will be prohibited in EU, as they are incompatible with EU values and fundamental rights:

1. Subliminal manipulation;
2. Biometric categorization (without explicit consent);
3. Exploitation of vulnerabilities;
4. Social scoring (general purpose);
5. Real-time remote biometric identification (in public);
6. Assessment of emotional state;
7. Predictive policing;
8. Collecting facial images (online or from surveillance videos).

Unacceptable risk
Examples

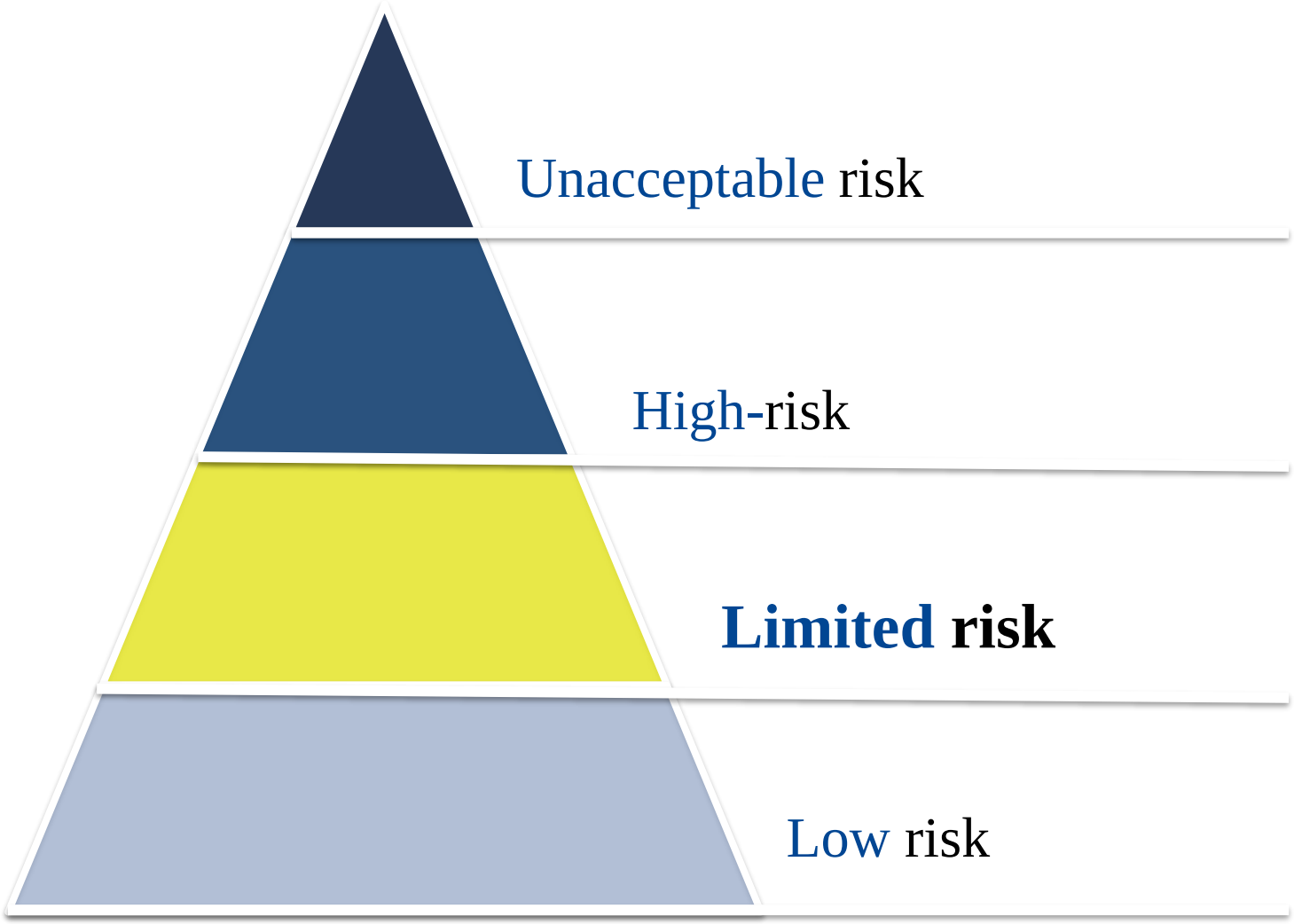
Risk-based approach to AI systems



High-risk:
AI systems that can result in significant harm to people's health, safety, fundamental rights or the environment.

- Not prohibited
- Subject to mandatory regulations

Risk-based approach to AI systems



Limited risk
AI systems with a risk of manipulation or deceit.

AI systems must be transparent:

- Users must be informed that the system in use is AI-generated or AI-based.
- Deep fakes should be denoted.

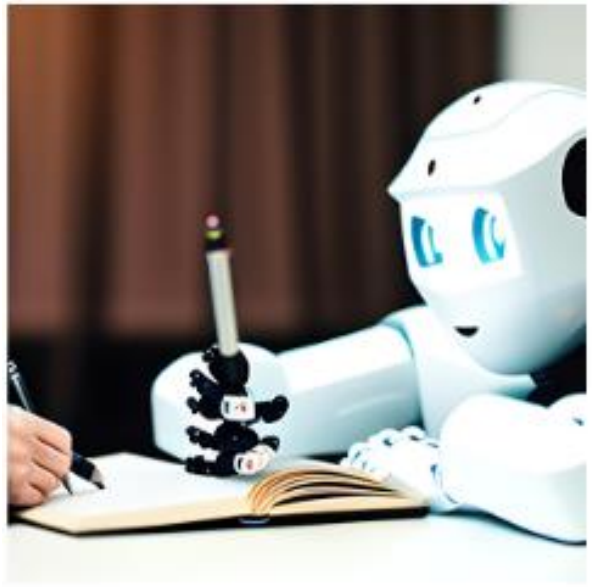
Examples: chatbots, manipulated image/video/audio content.

+ Transparency requirements
for General Purpose AI systems

+ Transparency requirements
for General Purpose AI systems

Examples: ChatGPT, OpenAI, Text to Image tools

Disclosing that the content was generated by AI



Designing the model to prevent it from generating illegal content

Publishing or disclosing summaries of copyrighted data used for training

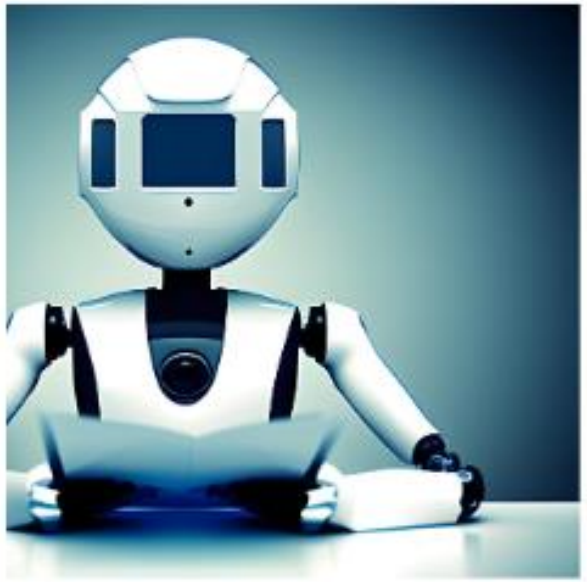
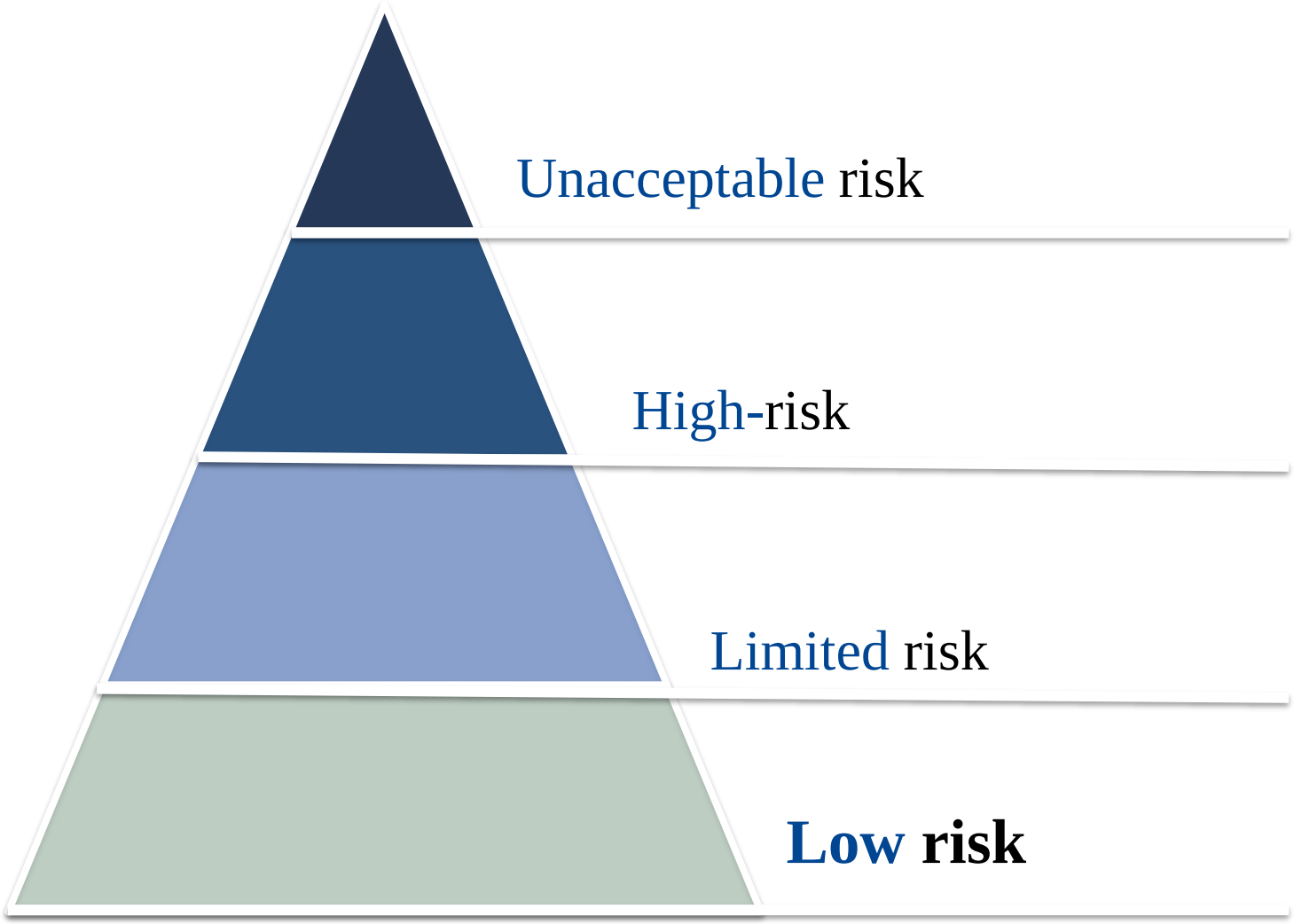


Image Sources: Generated by AI with *Canva Text to Image*

Risk-based approach to AI systems



Other AI Systems with low risk:

- No mandatory obligations;
- Follow general principles of non-discrimination, fairness, and human oversight – code of conduct.

Examples: email spam filter.

III
EU AI ACT: High-Risk AI Systems

Categories

Design
Implementation
Approval Processes
Usage

What makes an AI system a high-risk system?

Criteria used to assess whether an AI system poses a risk of adverse impact on fundamental rights:

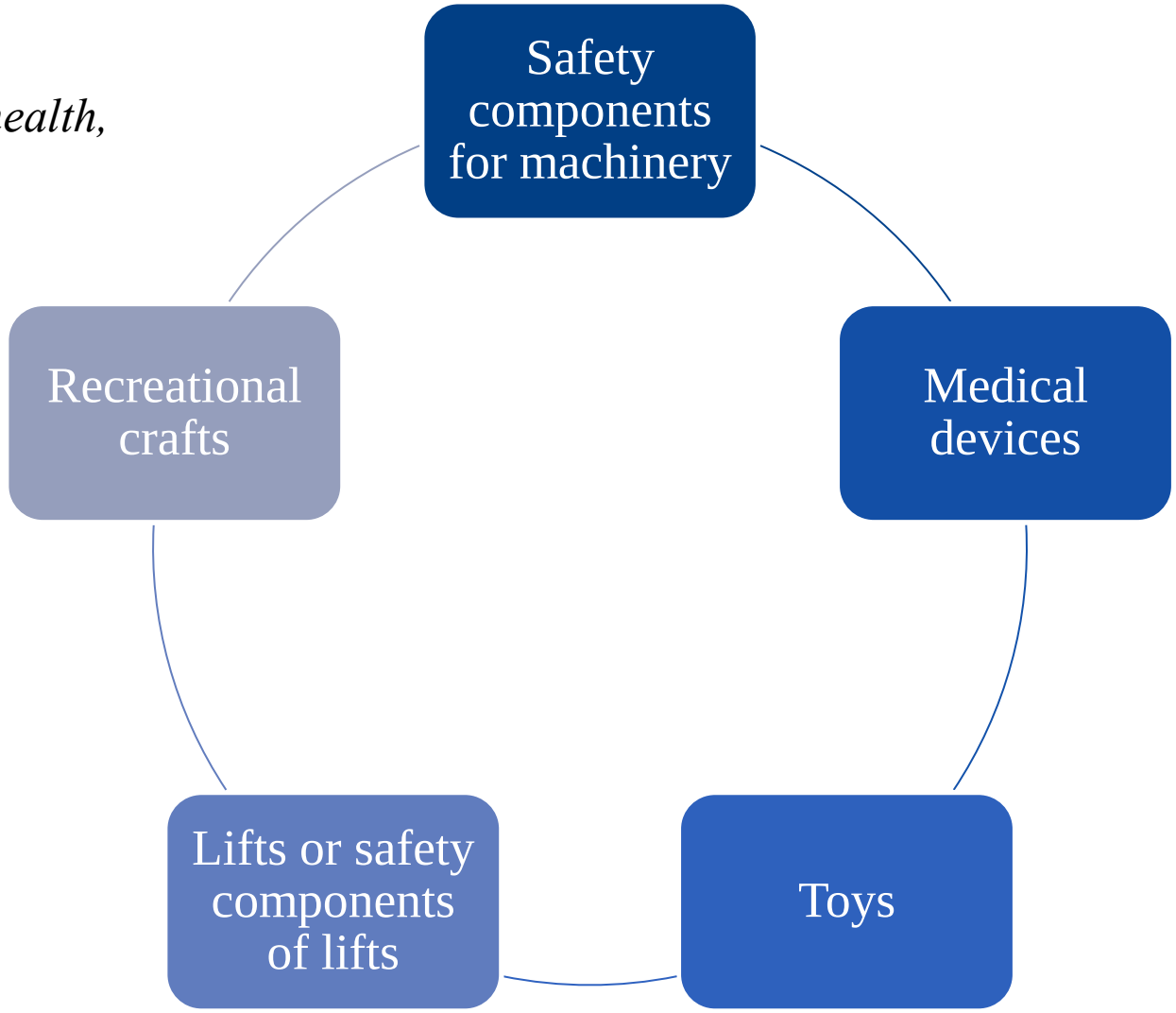
- Intended purpose of AI system
- Potential extent of the harm
- The extent to which harmed persons are in a vulnerable position
- The extent in which the outcome of the system is reversible
- The extent in which existing legislation provides for effective measures to address and minimize the risks

High-risk AI Systems

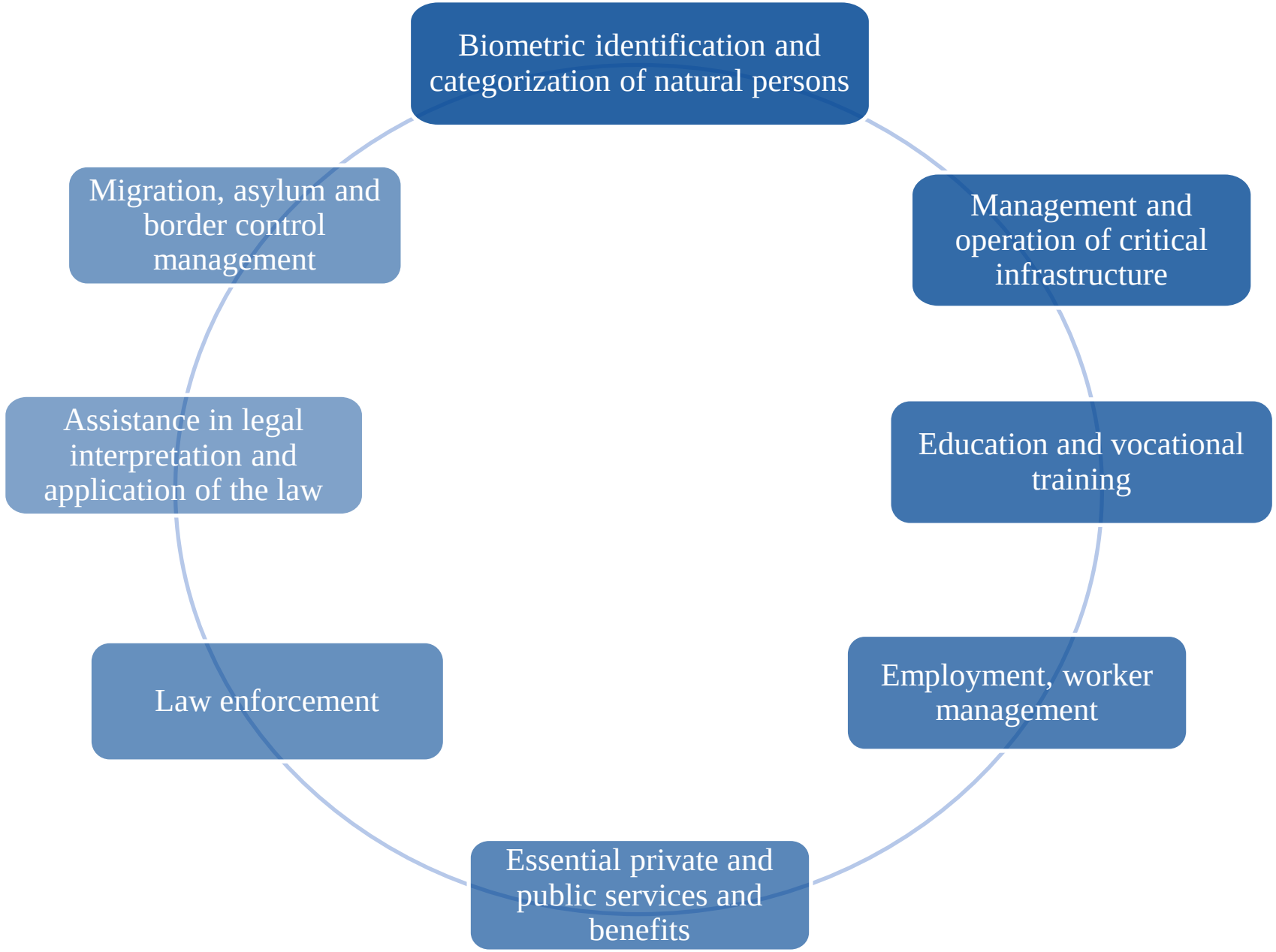
AI systems that can result in significant harm to people's health, safety, fundamental rights or the environment.

High-risk AI systems are the systems that are:

- intended to be used as a safety component of a product,
- themselves a safety component of a product;



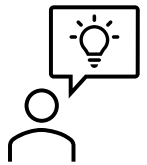
+ High-risk AI systems are the systems used in the following areas:



EU AI Act:
Stages for a Regulated High-risk AI system

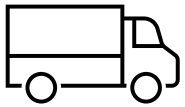
Design	Conforming to the design requirements for a high-risk AI system.
Implementation	The obligations towards putting the system into effect in a regulated manner.
Approval	Process of conformity assessment / assessing compliance with the Regulation.
Usage	Regulated usage of the system and the accompanied obligations.

Relevant
Roles in the supply chain



Provider

develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark;



Distributor

makes an AI system available on the Union market without affecting its properties;

Importer

established **in the Union** that places on the market or **puts into service an AI system** that is **established outside the Union**;

User

using an AI system under its authority

Design of High-Risk AI Systems

High-Risk AI Systems designed in a way that complies with having the following seven **requirements**:

Risk management system

Iterative risk management system

1. Identify and analyze known and foreseeable risks
2. Estimate and evaluate risks during intended use and foreseeable misuse
3. Evaluate the emerging risks based on data gathered after market release
4. Adopt suitable risk management measures

Image Sources: Generated by AI with *Canva Text to Image*

Data and data governance

Collect and process data appropriately.

Choose relevant design of and methods for training, validation, and testing of data sets.

Assess suitability, availability, sufficiency and biases of data.

Assess what the data is representing and measuring.

Identify possible gaps and ways to address them.

Image Sources: Generated by AI with *Canva Text to Image*

Technical documentation

- Demonstrative of compliance with the (design) requirements and provide the system design, development and usage specifications
- EU declaration of conformity
- Notification of authority

Record-keeping

- Designed with capability of recording events or logs during the use of the system for monitoring purposes

Image Sources: Generated by AI with *Canva Text to Image*

Transparency and providing information to the users

- Sufficiently transparent so that it enables to user to interpret outputs of the system
- Provide appropriate information (when applicable)

  Examples:
  - *Identity of providers*
  - *Capabilities and limitations*
  - *Intended purpose*
  - *Foreseeable risks*
  - *Human oversight measures*

Human oversight

System should be designed in a way that is effectively overseen by natural persons during its use to prevent and minimize risks to health, safety or fundamental rights.

Human overseeing should:

- Monitor operation and anomalies
- Be aware of automation-bias (not over-trust)
- Be able to make decisions and override the system

Example: driver in an autonomous car

Image Sources: Generated by AI with *Canva Text to Image*

Accuracy, robustness, and cybersecurity

- Designed for **intended purpose**
- **Robust** through fail-safe plans
- **Resilient** to errors and malicious use
  (able to recover)
- **Feedback loops**
  (to avoid biases from new data collected after market placement)

Image Sources: Generated by AI with *Canva Text to Image*

Implementation

What are the obligations (of providers) towards putting the system into effect in a regulated manner?

1. **Compliant with design requirements**
2. **Quality management system**
3. Technical documentation
4. Conformity assessment
5. Comply with registration obligations
6. Corrective measures if not compliant
7. Notify authorities
8. Conformity marking (Conformite Europeene CE)
9. Demonstrate conformity with the design requirements

Quality management is a system that ensure compliance with the Regulations.

- Documented in form of policies, procedures, instructions, containing (among other specifications):
  - **Strategy for a regulatory compliance** with conformity assessment;
  - **Standards** and **technical specifications**;
  - **Procedures for data management**;
  - **Risk management** system from design requirements;
  - Implementation of **post-market monitoring system**;
  - Procedure for **reporting of malfunction**;
  - **Communication with authorities**;
  - **Accountability framework** or responsibilities of management.

1. Compliant with design requirements
2. Quality management system
3. **Technical documentation**

+ Obligation of importers

4. **Conformity assessment**

- Ensuring to undergo relevant conformity assessment procedures.
  - Compliance with design requirements

+ Obligation of importers and distributors

5. Comply with registration obligations
6. Corrective measures if not compliant
7. Notify authorities
8. Conformity marking (Conformite Europeene CE)
9. Demonstrate conformity with the design requirements

1. Compliant with design requirements
2. Quality management system
3. Technical documentation
4. Conformity assessment
5. **Comply with registration obligations**
6. **Corrective measures if not compliant**
7. Notify authorities
8. Conformity marking (Conformite Europeene CE)
9. Demonstrate conformity with the design requirements

- Register the high-risk AI system in *EU database for stand-alone high-risk AI systems*, accessible to the public.

- If the system is not compliant with the Regulation, take necessary corrective actions to achieve conformity with the design requirements.

1. Compliant with design requirements
2. Quality management system
3. Technical documentation
4. Conformity assessment
5. Comply with registration obligations
6. Corrective measures if not compliant
7. **Notify and inform authorities**
8. **Conformity marking** **(Conformite Europeene CE)**
9. **Demonstrate conformity with the design requirements**

- Inform the competent authorities (in the Member State where the system is made available) about the system and the conformity assessment .
- Affix the CE to the system.
  CE is EU mandatory conformity marking for regulating the goods sold within the European Economic Area (EEA).
- Capable of demonstrating conformity with the design requirements upon request from authorities.

+ Obligation of importers

Approval Processes for High-Risk AI Systems

Conformity Assessment
Process of verifying the compliance with the design requirements.

When?

- Before placing the high-risk AI system on the EU market or in service
- With substantial modifications

Who?

- Performed by the system provider
- Manufacturer, distributor, importer, or by a third party

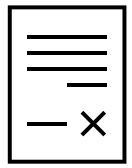
How?

- Verify Quality Management System
- Examine technical documentation
- Verify design requirements
- Verify post-market monitoring

Conformity and certificate

- EU Technical assessment certificate (with limited time validity)
- EU declaration of conformity
- CE marking of conformity (by provider)
- Registration in EU database (by provider)

Consequences in case of non-conformity

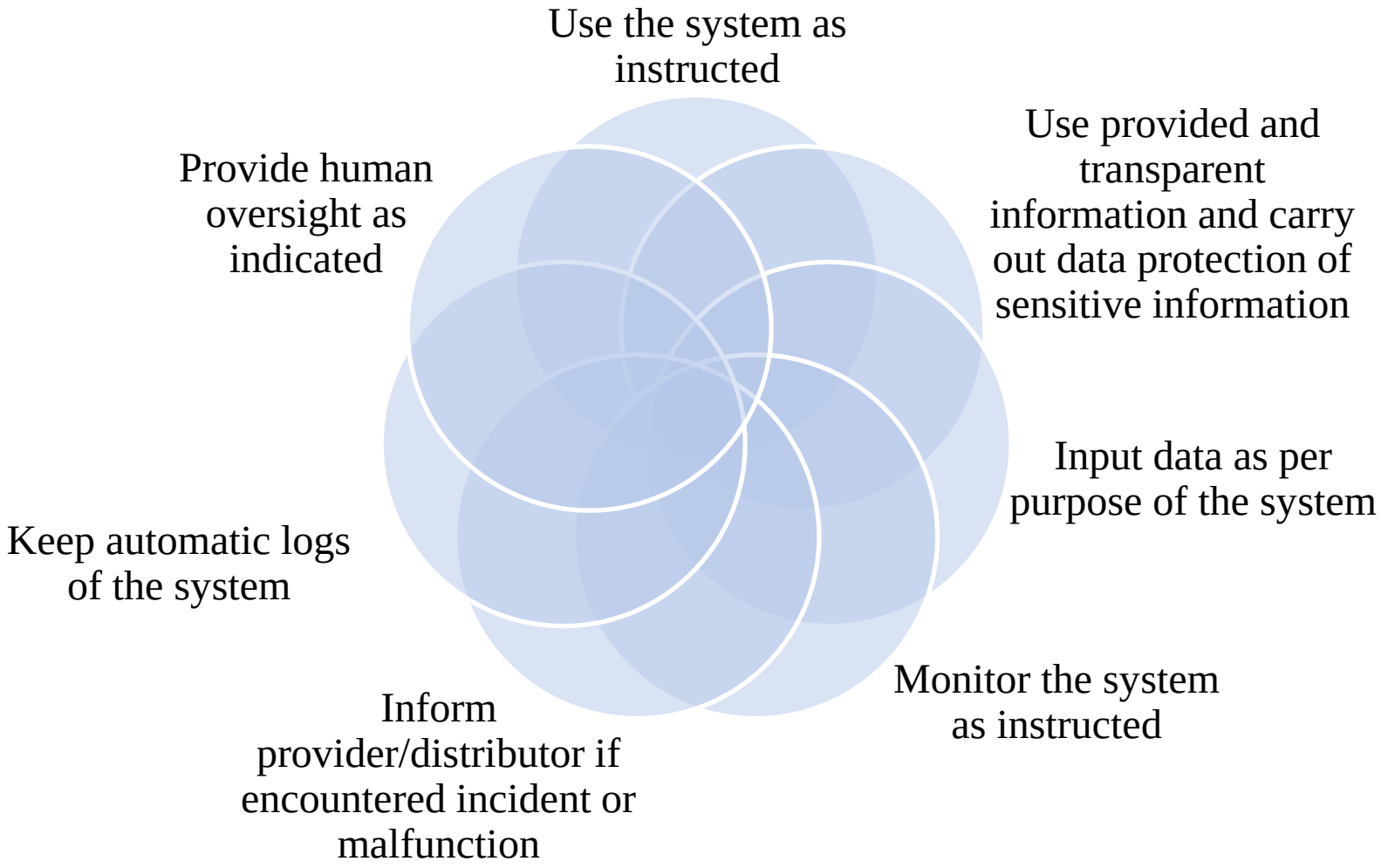


- Notified body communicates in detail the reasoning of non-conformity.

- Provider has the **right to appeal** against the decision of the notified body and must take the necessary **corrective actions**.

Usage of a High-Risk AI System

What are the obligations for the user of a high-risk AI system?



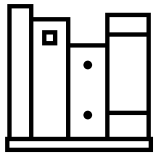
IV
Influence:

Influence
Criticism
Proposals outside of EU

Influence

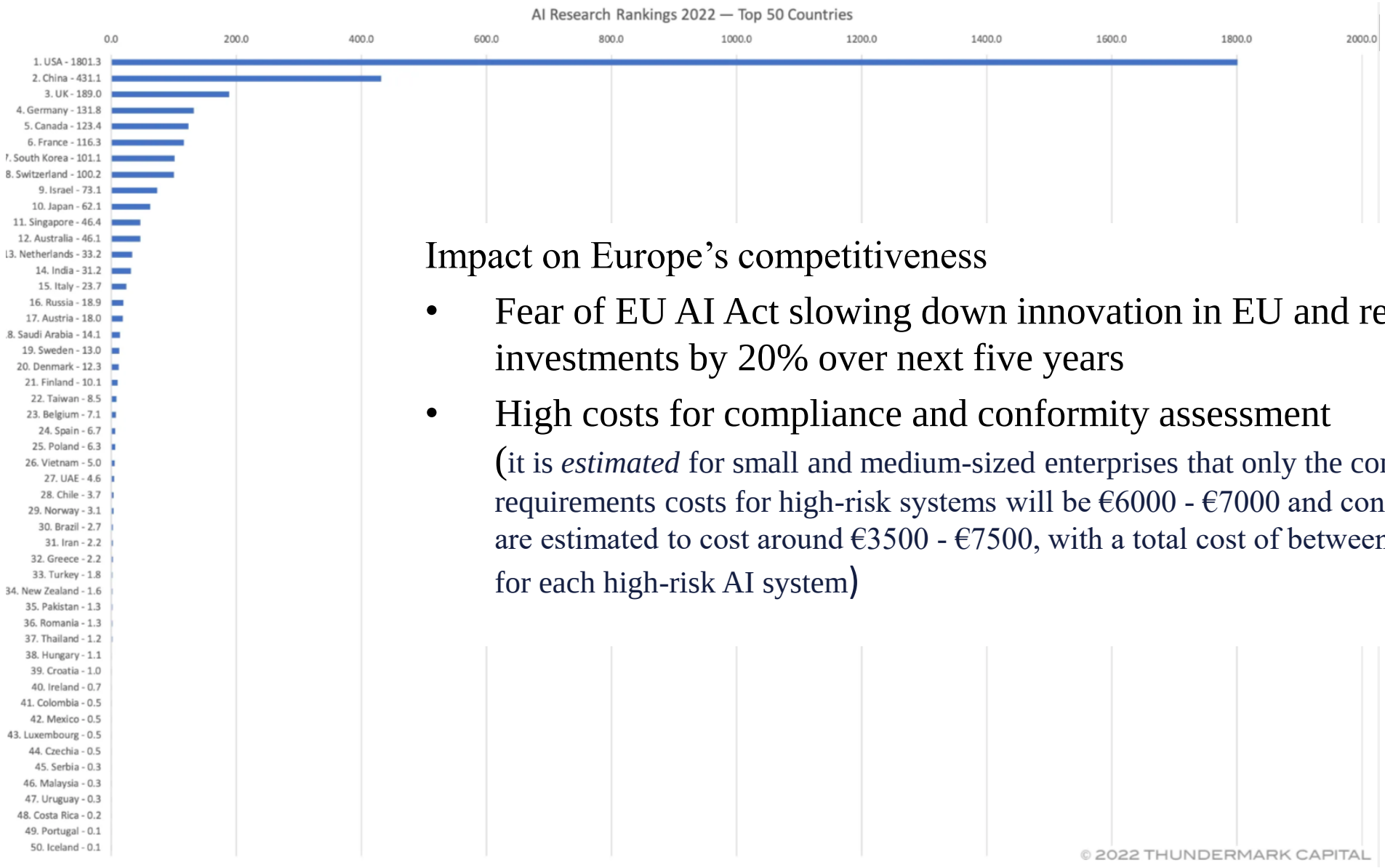
- EU AI Act is expected to have a global effect :
  Brussel effect (process of unilateral regulatory globalization caused by EU externalizing laws outside of EU through market regulations)

- Similar to General Data Protection Regulation (GDPR)
  - GDPR contains Data Protection Impact Assessment
    (with many similarities to EU AI Act Conformity Assessment for High-Risk AI Systems)
  - requires explicit consent from individuals for data sharing and exposure to automated decision-making processes



- Applicable to already regulated areas (e.g., medical devices) and are in conformity with harmonized standards.

Criticism



Impact on Europe's competitiveness

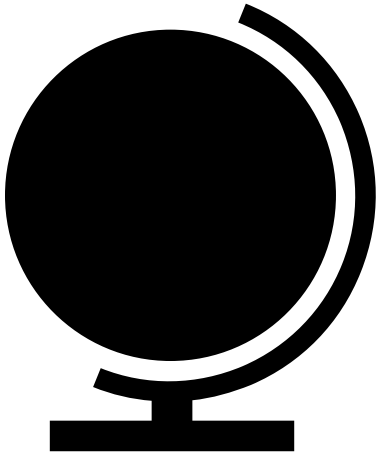
- Fear of EU AI Act slowing down innovation in EU and reducing investments by 20% over next five years
- High costs for compliance and conformity assessment

  (it is *estimated* for small and medium-sized enterprises that only the compliance with requirements costs for high-risk systems will be €6000 - €7000 and conformity assessments are estimated to cost around €3500 - €7500, with a total cost of between €9500 and €14500 for each high-risk AI system)

Proposals outside of EU

expected to closely follow the EU AI Act

- **United Kingdom** released a ‘Pro-Innovation’ approach to regulating AI (March 2023).
- **United States** Federal Trade Commission clarified authority to pursue enforcement actions against organizations failing to mitigate AI bias or that engage in harmful uses of AI.
- **Canadian** Federal Government requires algorithmic impact assessments for autonomous decision-making systems.



V
Conclusion

How to begin preparations for Regulatory changes?
Takeaways

How to begin preparations for the Regulatory changes?

- ❑ Get familiarized with the EU AI Act and harmonized standards
- ❑ Consider the scope of regulations
- ❑ Consider the risk in case of non-compliance

Risk Assessment

- Conduct the conformity assessment of current or planned AI systems

Governance Framework

- Facilitate best practices and development of responsible AI systems

AI Database

- Register all AI systems along with their technical documentation

Takeaways

AI systems should be overseen by people, rather than automation, to prevent harmful outcomes.

EU AI Act is ensuring a safe, transparent, traceable, non-discriminatory, sustainable approach to AI.

- **Suppliers** of artificial intelligence should be subject to a minimal but clear set of requirements, thereby assuring legal certainty and access to the entire single market.
- **Users** of AI should have legal assurance that the high-risk AI systems they purchase adhere to European laws and values.
- **Consumers** should benefit by having reduced risk of their safety and fundamental rights being violated.

References

- Proposal for a Regulation of the European Parliament and of the Council https://artificialintelligenceact.eu/the-act/
- ANNEXES to the Proposal for a Regulation of the European Parliament and of the Council https://artificialintelligenceact.eu/annexes/
- WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdfhttps://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf
- The EU AI Act: Adoption Through a Risk Management Framework, Adeline Chan, CISM (10 July 2023) https://www.isaca.org/resources/news-and-trends/industry-news/2023/the-eu-ai-act-adoption-through-a-risk-management-framework#:~:text=AI%20Act%20Risk%20Categories,that%20create%20an%20unacceptable%20risk. (Accessed: 20 August 2023)
- EU AI Act: first regulation on artificial intelligence, European Parliament News (14 June 2023) https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence (Accessed: 20 August 2023)
- EU AI Act: How risk is classified, Trail, https://www.trail-ml.com/blog/eu-ai-act-how-risk-is-classified (Accessed: 20 August 2023)
- Beyond the Individual: Governing AI's Societal Harm, Nathalie A. Smuha, Internet Policy Review 10 (3) DOI: 10.14763/2021.3.1574
- Center for Data Innovation, How Much Will the AI Act Cost Europe? Benjamin Mueller, 2021
- AI Research Rankings 2022, Thundermark Capital (May 20, 2022) https://thundermark.medium.com/ai-research-rankings-2022-sputnik-moment-for-china-64b693386a4 (Accessed: 22 August 2023)
- What considerations have been made for SMEs under the EU AI Act? HollisticAI, August 22, 2023 https://www.holisticai.com/blog/how-will-smes-be-supported-under-the-eu-ai-act#:~:text=Indeed%2C%20it%20is%20estimated%20that,each%20high-risk%20AI%20system.