

SHAPING EUROPE'S DIGITAL FUTURE

The AI Act: moving to implementation

Antoine-Alexandre André – AI Office A.4

The EU AI Act: rules for trustworthy AI in Europe



**Product safety
regulation**



**Complementary with
other EU law**



Focus on AI systems



Innovation-friendly

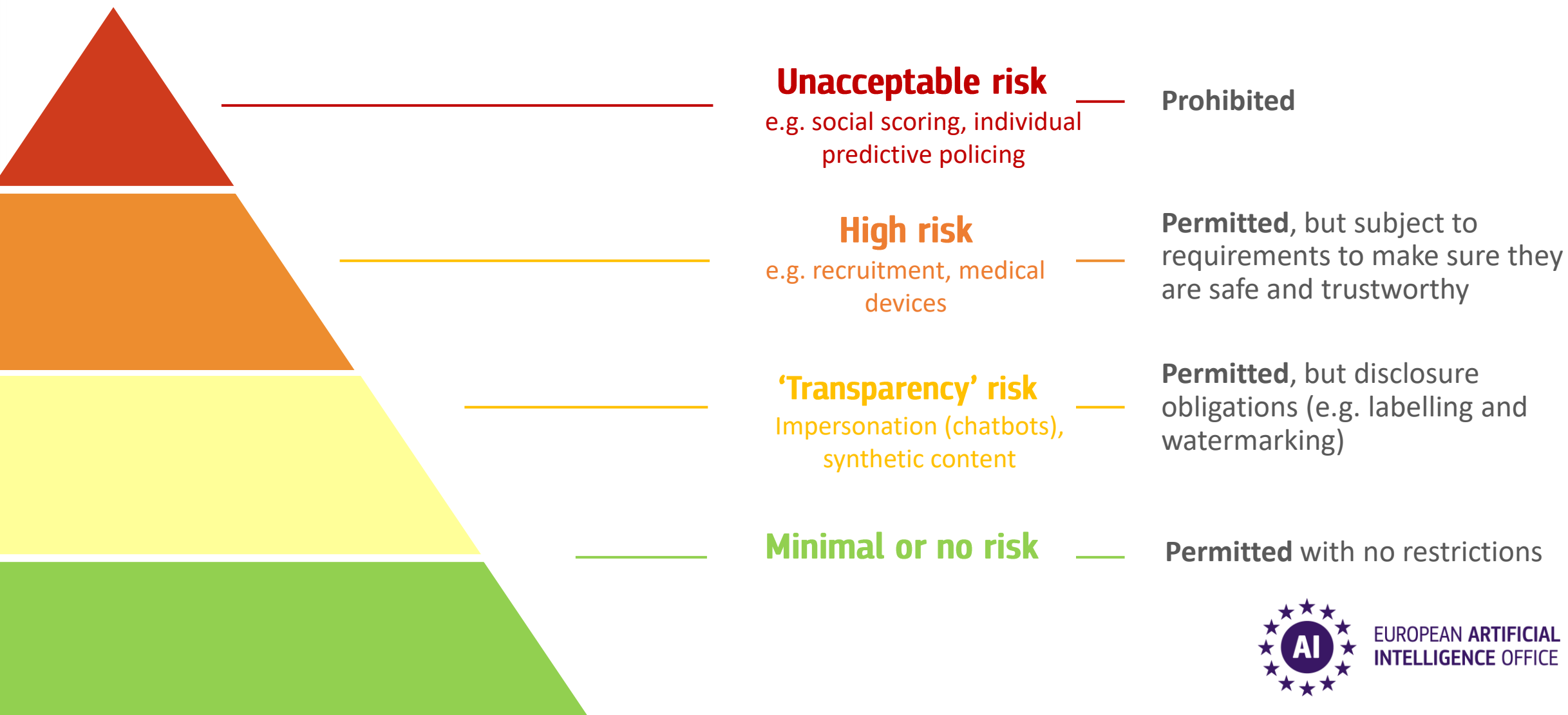


Risk-based approach



**EUROPEAN ARTIFICIAL
INTELLIGENCE OFFICE**

A risk-based approach



EUROPEAN ARTIFICIAL
INTELLIGENCE OFFICE

High-risk AI systems will have to comply with certain rules

1. High-risk systems embedded in products covered by Annex I

2. High-risk (stand-alone) use cases listed in Annex III:

- **Biometrics:** Remote biometric identification, categorization, emotion recognition;
- **Critical infrastructures:** e.g. safety components of digital infrastructure, road traffic
- **Education:** e.g. to evaluate learning outcomes, assign students in educational institutions
- **Employment:** e.g. to analyse job applications or evaluate candidates, promote or fire workers
- **Essential private and public services:** determining eligibility to essential public benefits and services; credit-scoring and creditworthiness assessment, risk assessment and pricing in health and life insurance
- **Law enforcement:**
- **Border management:**
- **Administration of justice and democratic processes**

Filter mechanism:

Excludes systems from the high-risk list that:

- perform narrow procedural tasks,
- improve the result of previous human activities,
- do not influence human decisions or
- do purely preparatory tasks,

NB. Profiling of natural persons always high-risk



EUROPEAN ARTIFICIAL
INTELLIGENCE OFFICE

High-risk AI systems will have to comply with certain rules

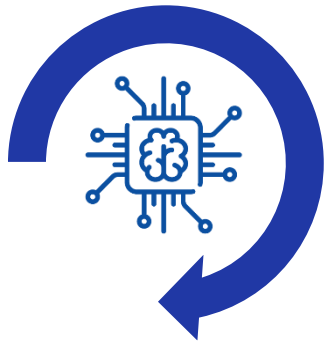
Provider obligations

- ▶ **Risk management system** to minimise risks for deployers and affected persons
- ▶ **Trustworthy AI requirements:** data quality and management, documentation and traceability, transparency and information to deployers, human oversight, accuracy, **cybersecurity** and robustness
- ▶ **Conformity assessment** to demonstrate compliance prior to placing on the market
- ▶ **Quality management system**
- ▶ **Register** standalone AI system in EU database (listed in Annex II)
- ▶ Conduct **post-market monitoring** and report **serious incidents**
- ▶ Non-EU providers to appoint **authorized representative in the EU**

Deployer obligations

- ▶ Operate high-risk AI system in accordance with **instructions of use**
- ▶ Ensure **human oversight:** persons assigned must have the necessary competence, training and authority **Monitor** for possible risks and **report problems and any serious incident** to the provider or distributor
- ▶ Public authorities to **register the use in the EU database**
- ▶ **Inform affected workers** and their representatives
- ▶ **Inform people** subjected to decisions taken or informed by a high risk AI system and, upon request, provide them with **an explanation**

Transp. and risk manag. for powerful AI models



General-purpose AI models

= highly capable AI models used at the basis of AI systems such as ChatGPT

**Transparency for all
general-purpose AI models**



**Risk management for those with
systemic risk**



Codes of practice developed together with stakeholders will detail out rules



EUROPEAN ARTIFICIAL
INTELLIGENCE OFFICE

A robust governance framework

Rules for AI systems

National level:

EU Member States to designate supervisors

Rules for general-purpose AI models

EU level:

AI Office within Commission



AI Board

with EU Member States to coordinate at EU level



Scientific Panel

supports with independent technical advice



Advisory Forum

supports with stakeholder input

What's next

Commission actions and priorities:

Setting up the governance structure

Growing the AI Office, set-up of advisory bodies with priority on AI Board and Scientific Panel and supporting build-up of national governance

Coordination of Code of practice on GPAI

COM coordinates and guides drawing up by GPAI developers and other stakeholders (due 9 months after entry into force), process will be launched soon

Contributing to preparation of standards

COM has mandated CEN/CENELEC to develop standards and actively follows the process; next step is amendment of mandate after AI Act entry into force

Preparing guidelines, implementing and delegated acts

Priority deliverables are the guidelines on AI system definition and on prohibitions

Entry into
force:
Aug. 2024

Prohibitions

6 months

Governance & GPAI

12 months

All other rules

24 months

Rules for Annex I
high-risk AI systems

36 months

Interplay AI Act and CRA: the basics

▶ **Presumption of conformity:** high-risk AI systems compliant with the essential requirements of the CRA are deemed to be compliant with the cybersecurity requirement of the AI Act.

▶ **Conformity assessment procedure:** the general rule -> the AIA is a reference act, i.e. the procedure under Article 43 of the AI Act is to be followed.

Exception, high risk AI systems qualified as critical products under the CRA + subject to Annex VI (conformity assessment based on internal control), in this case the conformity assessment provisions of the CRA apply to the essential requirements + for all the other aspects covered by the AI Regulation, the respective provisions on conformity assessment based on internal control set out in Annex VI to the AI Regulation apply.

For further information

Analysis of the preliminary AI standardisation work plan in support of the AI Act

Publications Office of the European Union

<https://publications.jrc.ec.europa.eu/repository/handle/JRC132833>

Cybersecurity of artificial intelligence in the AI Act

Publications Office of the European Union

<https://publications.jrc.ec.europa.eu/repository/handle/JRC134461>

AI Watch: Artificial Intelligence Standardisation Landscape Update

Publications Office of the European Union

<https://publications.jrc.ec.europa.eu/repository/handle/JRC131155>

The role of explainable AI in the context of the AI Act

2023 ACM Conference on Fairness, Accountability and Transparency

<https://dl.acm.org/doi/pdf/10.1145/3593013.3594069>

