



# Designing Ethical and Trustworthy AI Research Policies at EU level

**ECAI 2024**

**Santiago de Compostela, 22 October 2024**

*Dr Mihalis Kritikos, Ethics & Research Integrity  
Sector*

# The Ethics and Integrity life-line – The policy and operational framework



## The ERA PACT

[https://ec.europa.eu/info/news/commission-adopts-proposal-pact-research-and-innovation-europe-2021-jul-16\\_en](https://ec.europa.eu/info/news/commission-adopts-proposal-pact-research-and-innovation-europe-2021-jul-16_en)



## HORIZON EUROPE REGULATION and SPECIFIC PROGRAMME

Ethics Appraisal and research  
financing in ethics and integrity

See for example Article 19

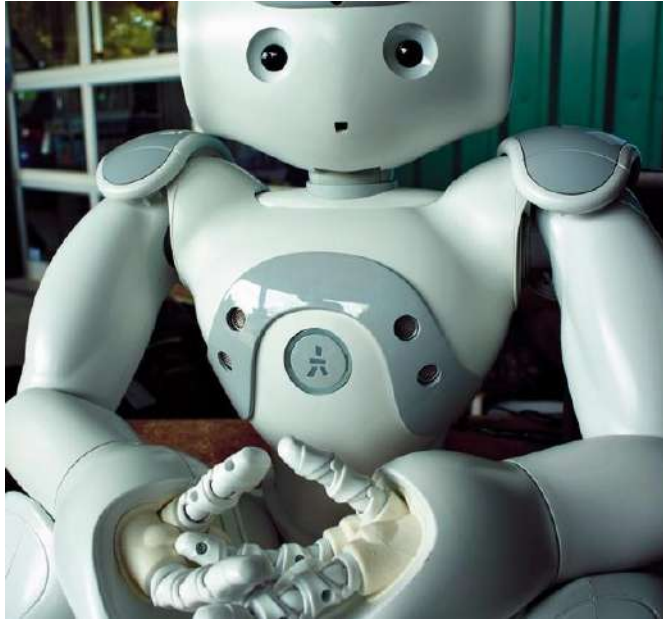


## The HE Grant Agreement

The responsibilities of the  
beneficiaries

See Ethics Article 14

# Ethics “big-ticket” Challenges



Support to  
researchers

(including emerging  
tech)



Ethics Committees support  
and workload at the  
national and local levels



Ethics Education and training

<http://www.eurecnet.org/information/index.html>

# REGULATION (EU) 2021/695 ESTABLISHING HORIZON EUROPE

---

## Article 19


‘Actions carried out under the Programme shall comply with **ethical principles** and **relevant Union, national and international legislation**, including the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights and its Supplementary Protocols.’



# MODEL GRANT AGREEMENT (MGA) ARTICLE 14

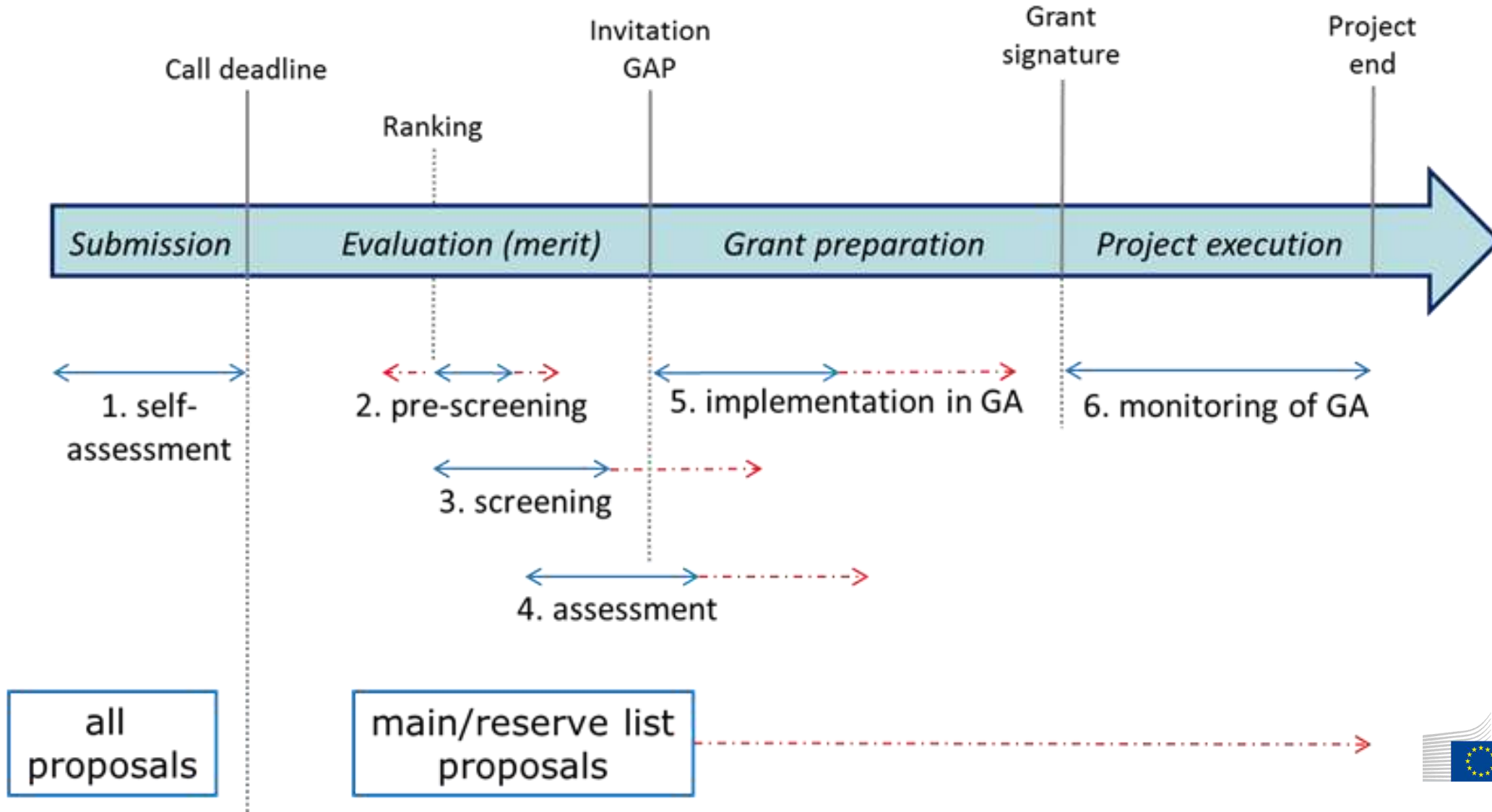
---

The action must be carried out in line with the **highest ethical standards** and the applicable EU, international and national law on ethical principles.'

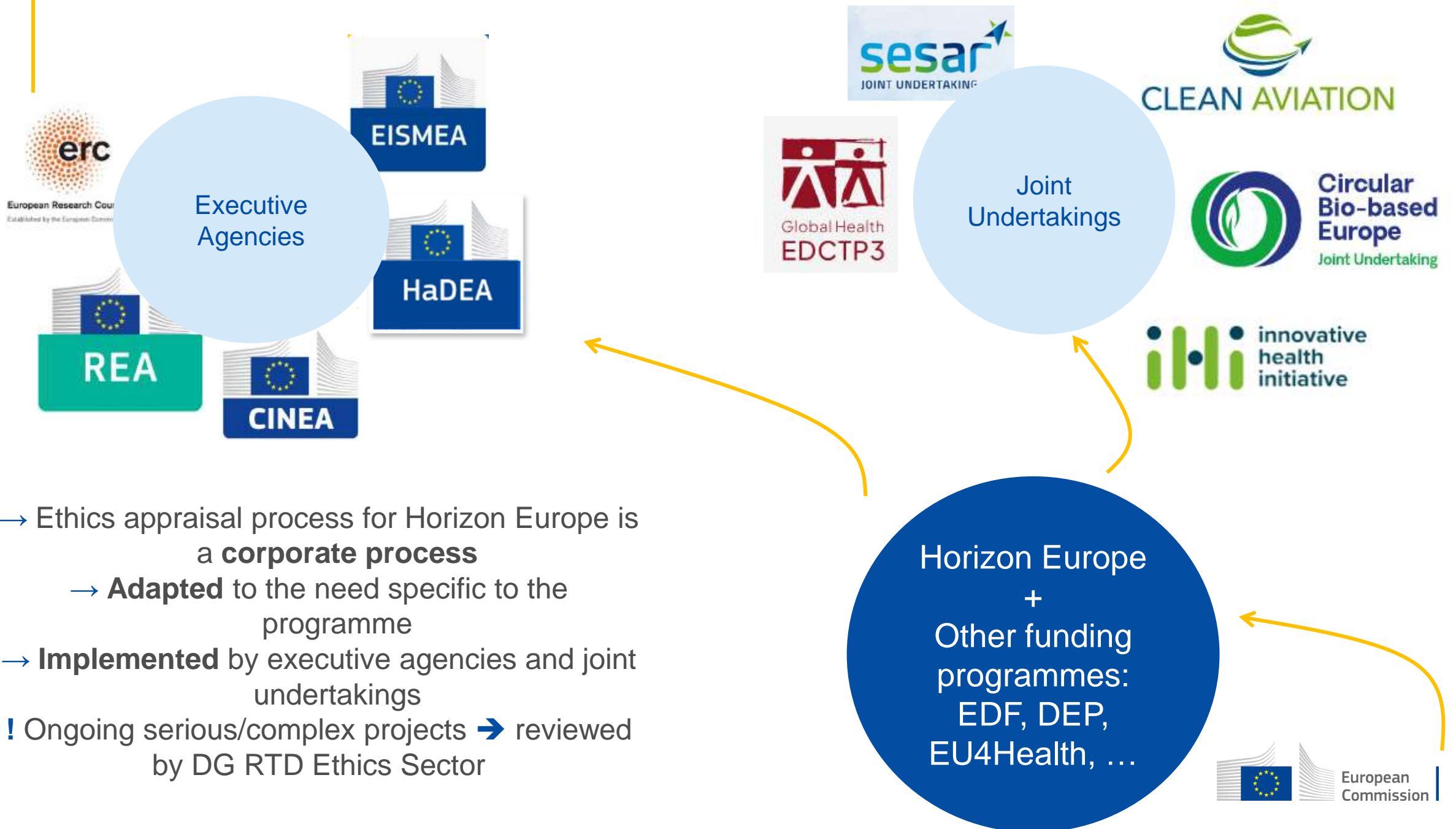
'The beneficiaries must commit to and ensure the respect of basic **EU values** (such as respect for human dignity, freedom, democracy, equality, the rule of law and human rights, including the rights of minorities).' 



# EU's Ethics Appraisal Mechanism







- Ethics appraisal process for Horizon Europe is a **corporate process**
  - **Adapted** to the need specific to the programme
- **Implemented** by executive agencies and joint undertakings
- ! Ongoing serious/complex projects → reviewed by DG RTD Ethics Sector

# What is the EU doing in the field of AI?

LEGAL GOVERNANCE OF AI: AI ACT/EU PACT

INNOVATION: FUNDING OF AI RESEARCH – AI FACTORIES

COORDINATED PLAN ON AI + INTERNATIONAL POLICY WORK

ETHICAL GOVERNANCE OF AI: GUIDELINES AND OTHERS

STANDARDISATION



# AI Act: general remarks

Risk based approach

Emphasis on the protection of fundamental rights

Multi-level governance structure

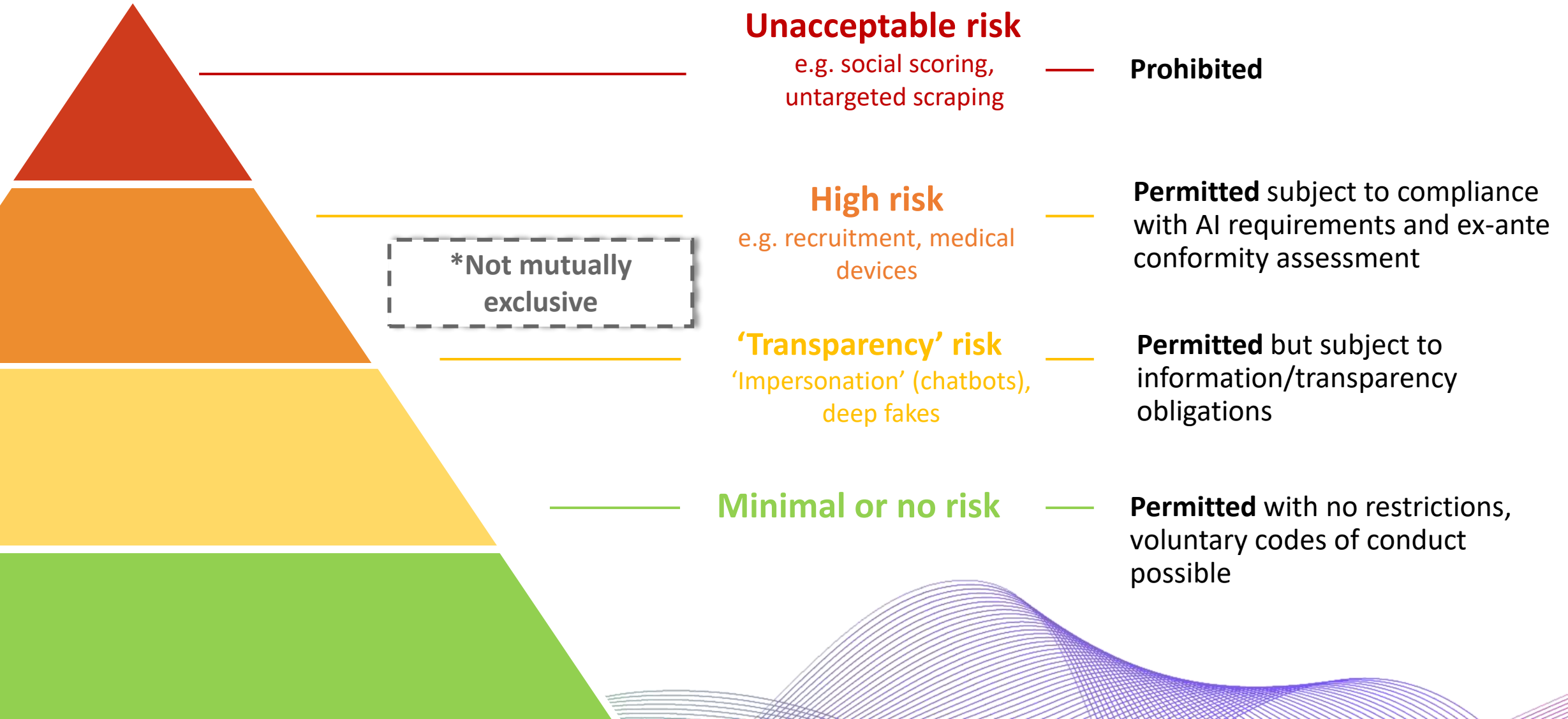
Strict fines and penalties

150-page long document

Lots of delegated/implementing acts clauses

Extra-territorial application

# The AI Act follows a risk-based approach





# The impact on fundamental rights has to be assessed

▶ The use of a high-risk AI system may produce an impact on fundamental rights after deployment  
Prior to first use, some deployers must do a **fundamental rights impact assessment for Annex III systems** (except critical infrastructure)

## Consisting of an assessment of:

- ▶ **Deployers' processes**, in which the high-risk AI system is intended to be used
  - ▶ **Categories of natural persons and groups** likely to be affected by its use in the specific context
    - ▶ **Specific risks of harm** likely to impact the affected categories of persons or group of persons
    - ▶ Description of **human oversight measures**
- ▶ Measures to be taken **in case of materialization of the risks**



## Carried out by

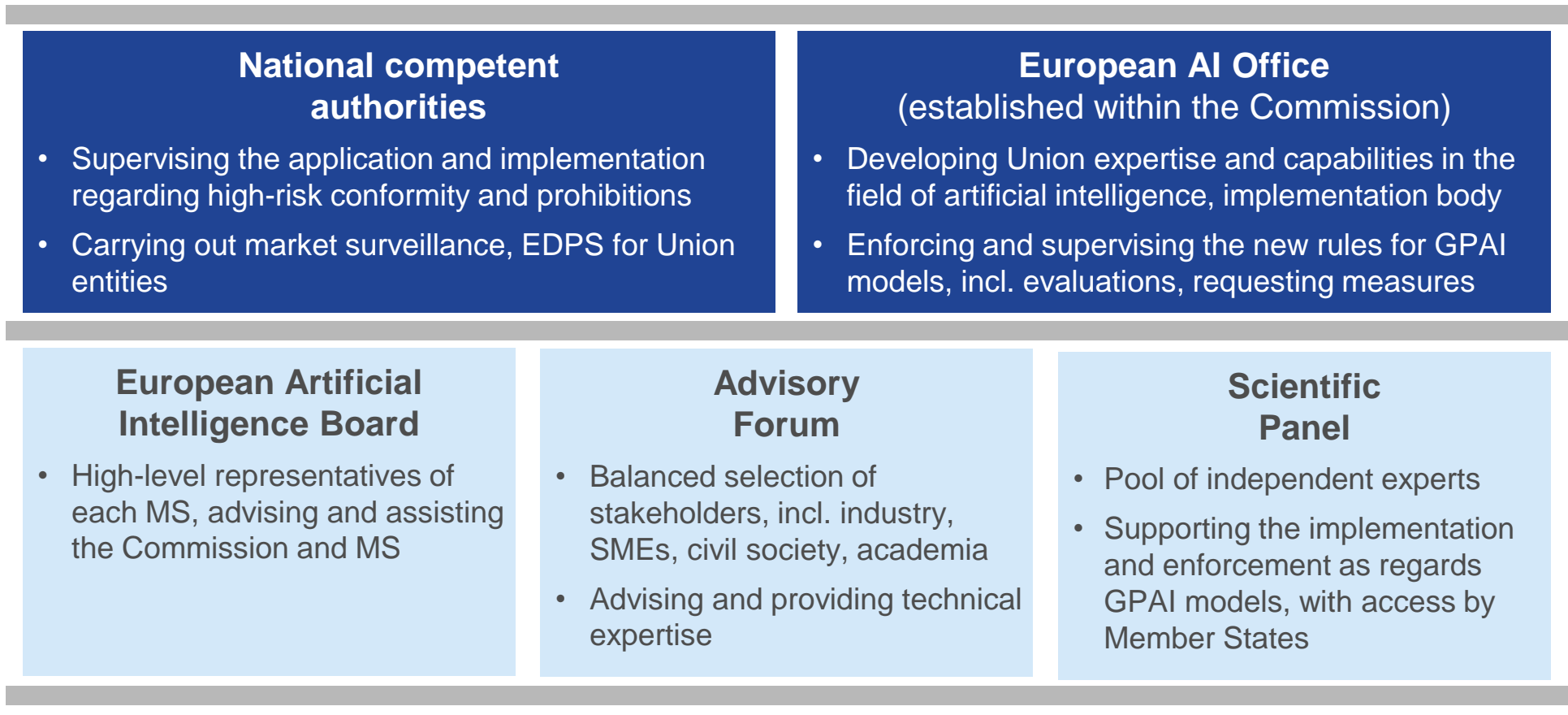
Deployers that are

1. Bodies governed by **public law**
2. Private operators providing **public services**
3. Certain other **private providers** (credit scoring/ credit worthiness assessment of health and life insurances)

# A holistic governance structure for effective enforcement

**Enforcement by national competent authorities and the AI Office**

**with a supportive structure for close collaboration with Member States and for additional technical expertise**



# Introducing the European AI Office

## 360 degrees vision on AI:

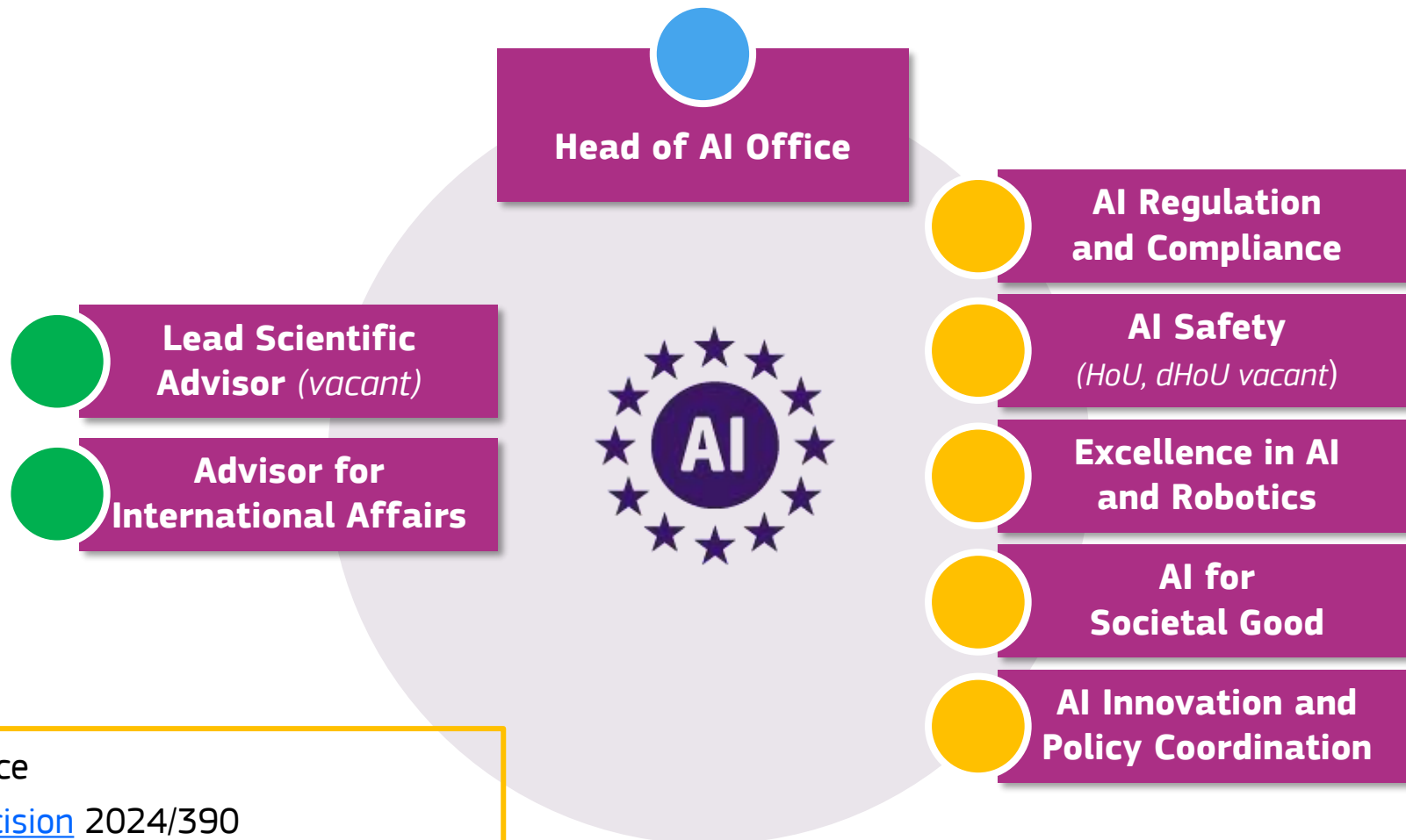
Key role in the implementation of the AI Act, in particular in the enforcement of rules on general-purpose AI models

Cooperation with all relevant EU bodies, Member States and the stakeholder community

Cross-sectoral cooperation within the Commission

Fostering research and innovation in trustworthy AI

Positioning the EU as a leader in international discussions and contributor to AI for good



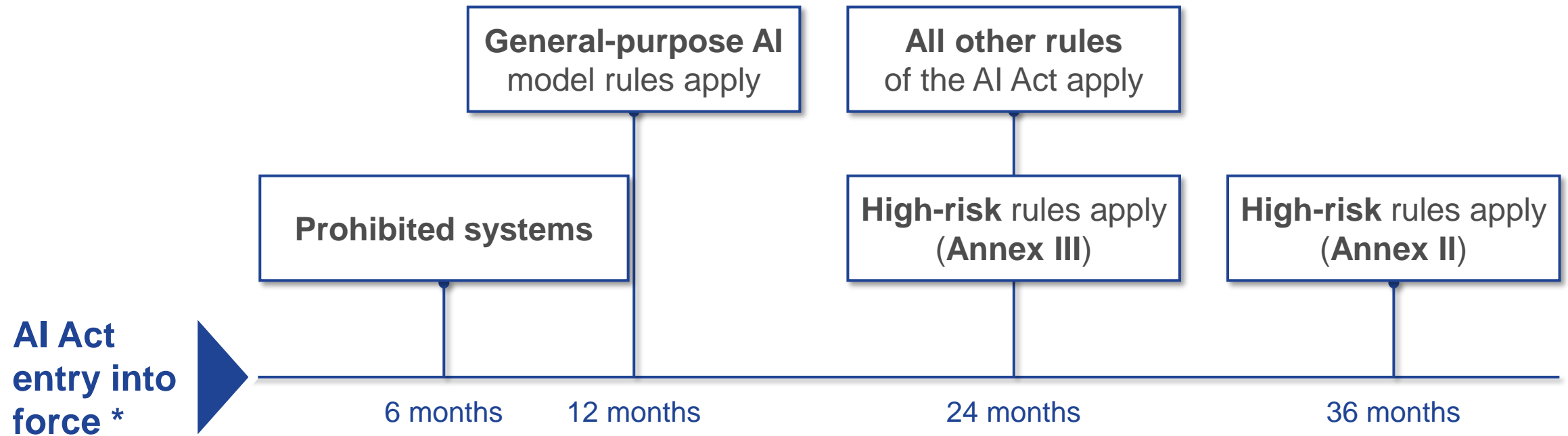
### Context:

- ❖ AI Act introduces role of AI Office
- ❖ Established by [Commission Decision](#) 2024/390
- ❖ Part of DG CNECT, launched in June 2024
- ❖ Tech specialist recruitments ongoing, more to come



EUROPEAN ARTIFICIAL  
INTELLIGENCE OFFICE

# The AI Act enters into application in a gradual approach



\*Following its adoption by the European Parliament and the Council, the AI Act shall enter into force on the twentieth day following that of its publication in the official Journal.





# AI ethics governance in the EU

EU GUIDELINES ON  
AI/ALTAI LIST

HUMAN RIGHTS  
IMPACT  
ASSESSMENT

AI ETHICS  
APPRAISAL  
MECHANISM

AI ETHICS  
STANDARDS AND  
TRAINING  
ACTIVITIES

EU-FUNDED  
RESEARCH ON AI  
VALUES/ETHICS

OPERATIONAL  
GUIDELINES /SAM  
OPINION

# Ethics Guidelines for Trustworthy AI

On 8 April 2019, **the High-Level Expert Group on AI** presented Ethics Guidelines for Trustworthy Artificial Intelligence.

According to the Guidelines, TRUSTWORTHY AI should be:

- **lawful - respecting all applicable laws and regulations**
  - **ethical - respecting ethical principles and values**
- **robust - both from a technical perspective while taking into account its social environment**

# Ethics Guidelines for Trustworthy AI



## 7 core requirements:

1. Human agency and oversight
2. Technical robustness and safety
3. Privacy and Data Governance
4. Transparency
5. Diversity, non-discrimination and fairness
6. Societal and environmental wellbeing
7. Accountability

# Guidance on Ethics By Design for AI



Ethics By Design and Ethics of Use Approaches  
for Artificial Intelligence

Version 1.0  
25 November 2021

Builds on the work of the **Independent High-Level Expert Group on AI** and their **'Ethics Guidelines for Trustworthy AI'**

**Addressing ethical issues during research and development**

# Ethics by design

---

Ethics by Design is an approach for **systematically and comprehensively including ethical considerations in the design and development process of new technological systems and devices**

---

The central idea behind it is that technology can be made ethical **through its design process**

---

Technology is not neutral / **values can be embedded into the design process**

---

**Design choices are not morally neutral** but can have significant ethical consequences

---

**Design choices can sometimes generate particular consequences** across a broad range of uses or use contexts

---

# EU Operational Guidelines in ethics

Guidance Note on *human-centered AI: algorithmic bias and fairness*



Guidance Note on *AI and informed consent*



Guidance Note on *AI Ethics Audits and Checks*



Guidance Note on *Human Rights Impact Assessment*



# International engagement

EU leadership in global AI governance and promoting a risk-based and human-centric approach to trustworthy AI

**Council of Europe:** AI Convention

**G7 Hiroshima AI Process:** Guiding principles & Code of Conduct

**OECD & Global Partnership on AI**

**Series of International AI Safety and Action Summits**

**UN:** AI Advisory Body, AI Resolution, Global Digital Compact (GDC)

**TTC & Digital Partnerships:** US, IN, SG, JP, KR, CA

**Collaboration with AI Safety Institutes:** US, UK, CA, KR, JP, SG

# Thank you



© European Union 2020

Unless otherwise noted the reuse of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.

Slide xx: **element concerned**, source: e.g. [Fotolia.com](https://www.fotolia.com/); Slide xx: **element concerned**, source: e.g. [iStock.com](https://www.istock.com/)

