European Commission

AI Policy: the EU approach and implications for education and academia

Elinor Wahal,
Legal and Policy Officer
DG CNECT A2, AI Policy Development and Coordination

15.05.2024

AI: a powerful technology that needs to be regulated

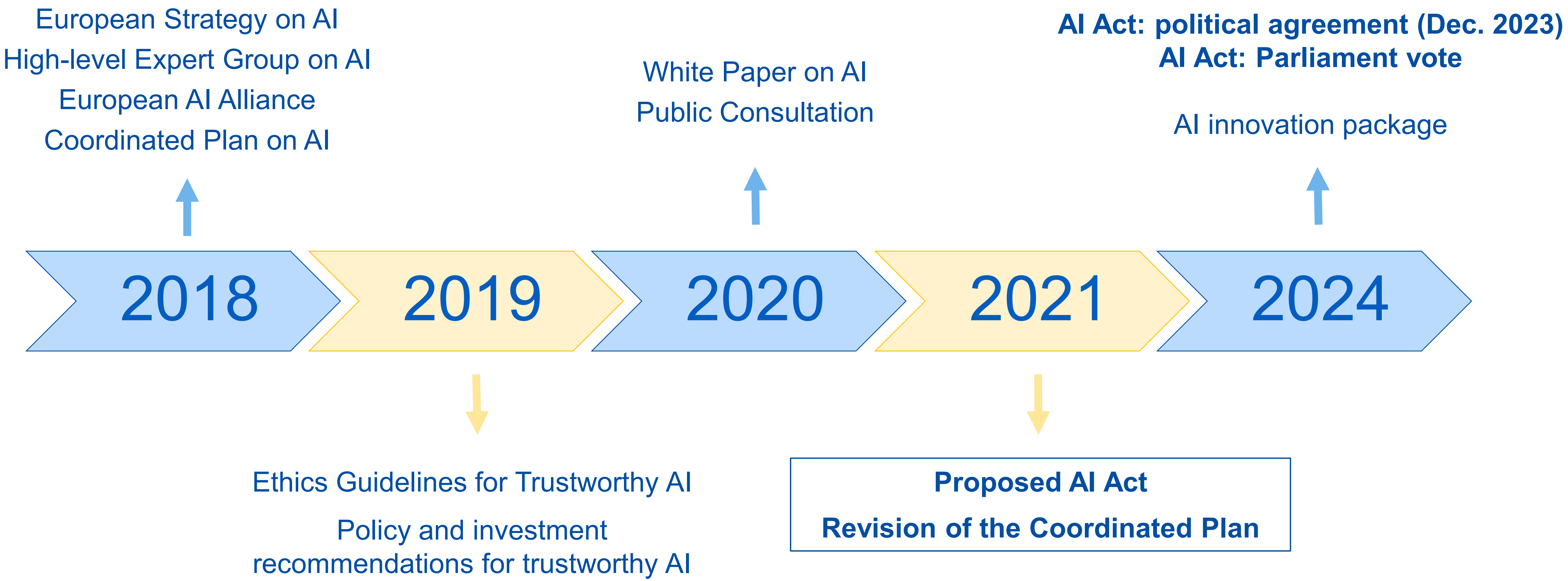
AI is good …

- For citizens
- For business
- For the public interest

… but it creates some risks

- For the safety of consumers and users
- For fundamental rights

EU strategy on AI – since 2018



2018

European Strategy on AI

High-level Expert Group on AI

European AI Alliance

Coordinated Plan on AI

2019

Ethics Guidelines for Trustworthy AI

Policy and investment recommendations for trustworthy AI

2020

White Paper on AI

Public Consultation

2021

Proposed AI Act

Revision of the Coordinated Plan

2024

AI Act: political agreement (Dec. 2023)

AI Act: Parliament vote

AI innovation package

An ecosystem of **excellence** and **trust**

- A **European legal framework for AI** that upholds fundamental rights and addresses safety risks specific to the AI systems

- A civil liability framework – adapting liability rules to the digital age and AI

- A revision of sectoral safety legislation (e.g. Machinery Regulation, General Product Safety Directive)

© iStock by Getty Images - 1139760401 peshkov

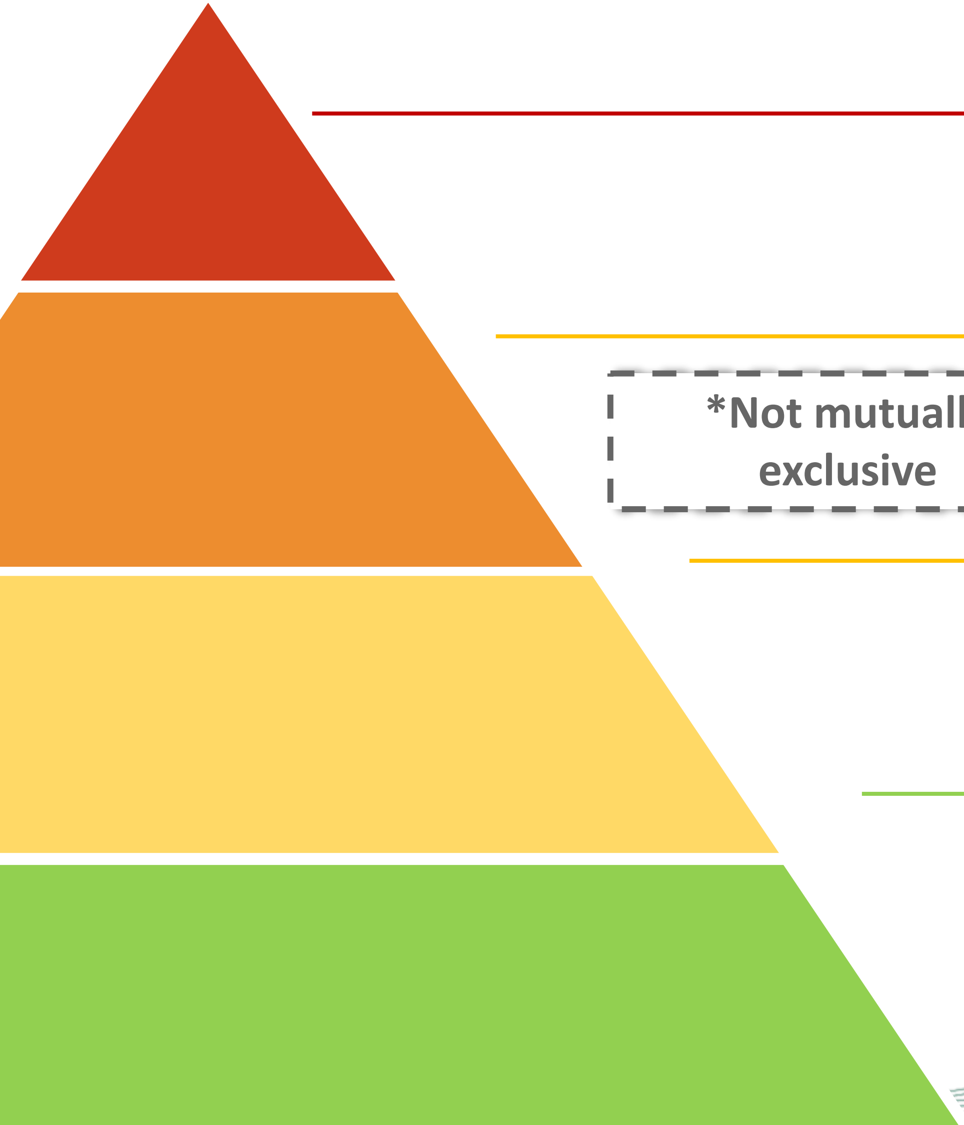
The EU Artificial Intelligence Act

European Commission

AI Act: foundations

- **Product safety** and **risk-based** approach
- Protection of **health, safety** and **fundamental rights**
- A **horizontal** act
- **Coherence and complementarity** with existing legislation
- **Innovation friendly**
- Will apply to **public and private** actors, **inside and outside the EU** (as long as the AI system is placed on the Union market or its use affects people located in the EU), **providers and deployers**

The AI Act follows a risk-based approach



*Not mutually exclusive

Unacceptable risk
e.g. social scoring, untargeted scraping — **Prohibited**

High risk
e.g. recruitment, medical devices — **Permitted** subject to compliance with AI requirements and ex-ante conformity assessment

‘Transparency’ risk
‘Impersonation’ (chatbots), deep fakes — **Permitted** but subject to information/transparency obligations

Minimal or no risk — **Permitted** with no restrictions, voluntary codes of conduct possible

A very limited set of particularly harmful AI uses are banned

Unacceptable risk

Subliminal techniques or exploitation of vulnerabilities	to manipulate people
Social Scoring	for public and private purposes
Biometric categorisation	to deduce or infer for example race, political opinions, religious or philosophical beliefs or sexual orientation, exceptions for labelling in the area of law enforcement
Real-time remote biometric identification	for the purpose of law enforcement, -with narrow exceptions and with prior authorisation by a judicial or independent administrative authority
Individual predictive policing	assessing or predicting the risks of a natural person to commit a criminal offence based solely on this profiling without objective facts
Emotion recognition	in the workplace and education institutions, unless for medical or safety reasons
Untargeted scraping of the internet	or CCTV for facial images to build-up or expand databases

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

High-risk use cases defined in Annexes II (embedded AI) and III:

Some examples from Annex III are related to

- **Certain critical infrastructures** such as road traffic, supply of water, gas, heating and electricity
- **Education and vocational training**, e.g. to evaluate learning outcomes
- **Employment, workers management**, e.g. to analyse job applications or evaluate candidates
- **Access to essential private and public services** and benefits, credit scoring
- **Remote biometric identification, categorization, emotion recognition; Law enforcement; border management; administration of justice and democratic processes**

Obligations for providers of high-risk AI systems:

- **Trustworthy AI requirements** such as data quality, documentation and traceability, transparency, human oversight, accuracy, cybersecurity and robustness
- **Conformity assessment** before placing the AI system on the market, to demonstrate compliance
- **Quality and risk management systems** to minimise risks for users and affected persons and to ensure compliance
- **Registration in an EU database**

This will be subject to **enforcement** to ensure that the high risk is effectively addressed.

The impact on fundamental rights must be assessed

The use of a high-risk AI system may produce an impact on fundamental rights.
This deserves a **fundamental rights impact assessment for most Annex III systems.**

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)

The most recent advancements in AI are addressed

So-called **‘general-purpose AI models‘ *** pose unique challenges

- General-purpose AI (GPAI) models can be used for a variety of tasks and are **becoming the basis for many AI systems in the EU**.
- Some of these models **could carry systemic risks** if they are very capable or widely used.
- For example, many individuals could be affected if a model propagates harmful biases across many applications.

General-purpose AI models are becoming too important for the economy and society not to be regulated.

***General-purpose AI model** = AI model, including when trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable to competently perform a wide range of distinct tasks regardless of the way the model is released on the market and that can be integrated into a variety of downstream systems or applications. *Research, development, and prototyping activities preceding the release on the market are not covered.*

Proportionate rules for GPAI models

Enabling downstream system providers to comply with the AI Act

All necessary information for providers wishing to build upon a GPAI model

Light-touch transparency obligations for all GPAI models

Documentation and information to downstream providers for instance through **model cards,** facilitated enforcement of **copyright rules,** info on **energy consumption**

Open-source models are exempted from transparency requirements, when they do not carry systemic risks.

Addressing systemic risks of a few GPAI models

Strict rules and oversight for very capable (at least 10^25 FLOPs*) or individually designated GPAI

Additional obligations for "GPAI models with systemic risk"

Managing risks and **monitoring serious incidents**, performing **model evaluation** and adversarial testing, **cybersecurity**

Operationalised through Codes of Practice developed by industry, the scientific community civil society and other experts, together with the AI Office

* The AI Office may update this threshold in light of technological advances, and may in specific cases designate other models as such based on further criteria (e.g. number of users, or the degree of autonomy of the model)

Rules for AI systems which are not high risk

Transparency obligations for certain AI systems

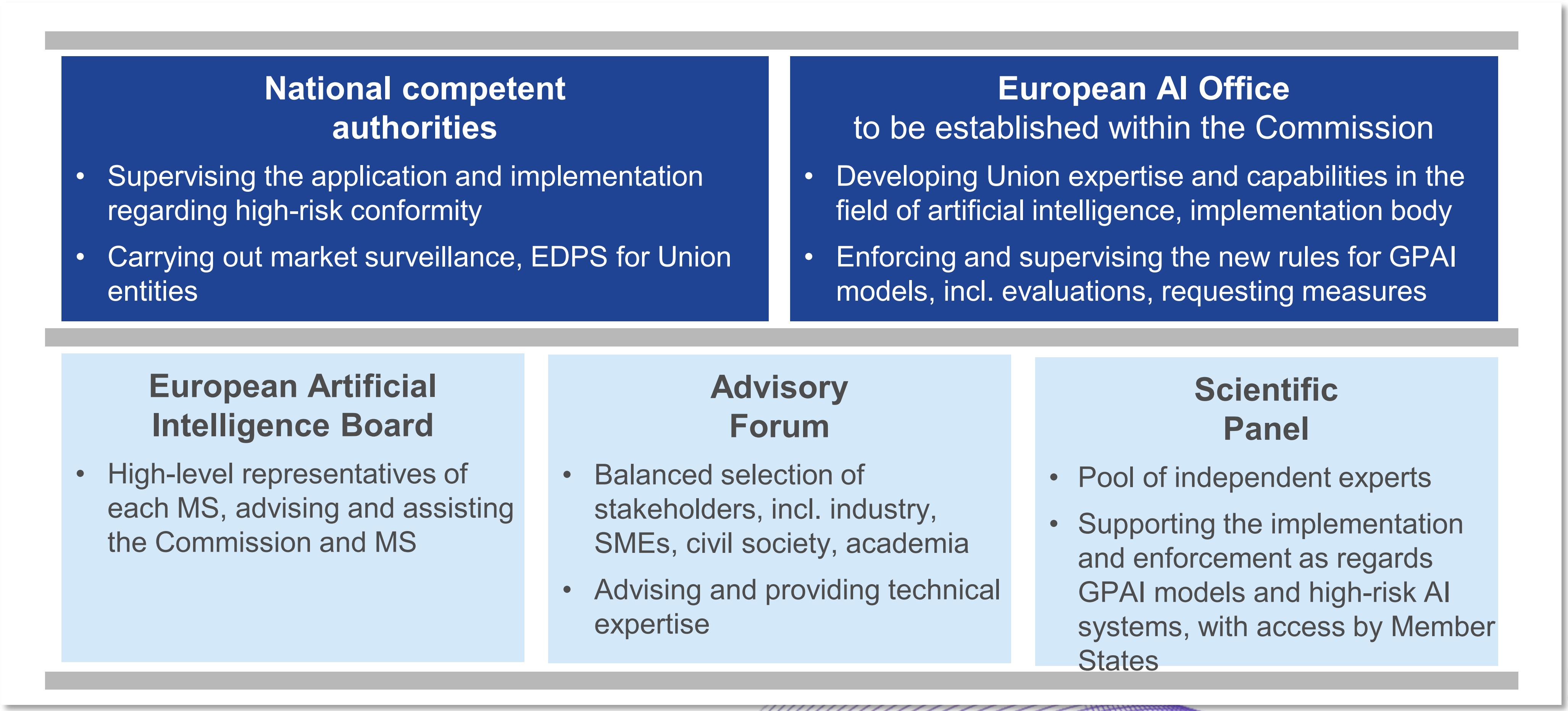
- **Notify humans** that they are **interacting with an AI system** unless this is evident.
- Ensure that synthetic audio, image, video or text content generated by an AI system **is marked in a machine-readable format and detectable as artificially generated.**
- **Label text as artificially generated** if it is published with the purpose of informing the public on matters of public interest.
- Apply **label to deep fakes generated by AI** (unless necessary for the exercise of a fundamental right or freedom or for reasons of public interests).
- Notify humans that **emotion recognition or biometric categorisation systems** are applied to them.

Possible voluntary codes of conduct for AI with specific transparency requirements

- No mandatory obligations

A holistic structure ensures 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



National competent authorities

- Supervising the application and implementation regarding high-risk conformity
- Carrying out market surveillance, EDPS for Union entities

European AI Office
to be established within the Commission

- Developing Union expertise and capabilities in the field of artificial intelligence, implementation body
- Enforcing and supervising the new rules for GPAI models, incl. evaluations, requesting measures

European Artificial Intelligence Board

- High-level representatives of each MS, advising and assisting the Commission and MS

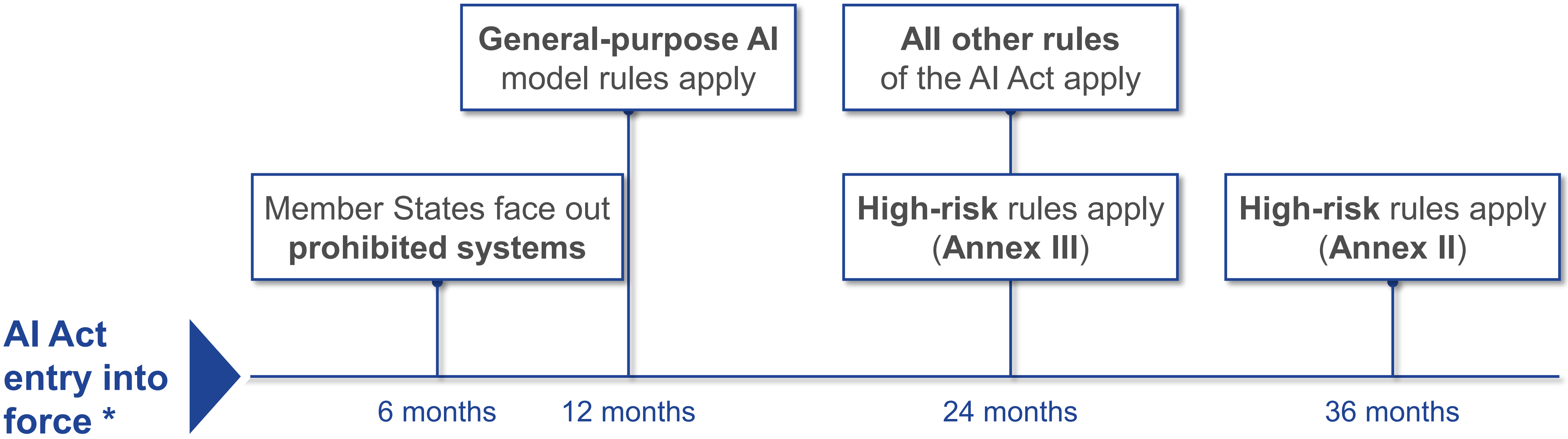
Advisory Forum

- Balanced selection of stakeholders, incl. industry, SMEs, civil society, academia
- Advising and providing technical expertise

Scientific Panel

- Pool of independent experts
- Supporting the implementation and enforcement as regards GPAI models and high-risk AI systems, with access by Member States

The AI Act enters into application in a gradual approach



AI Act entry into force *

6 months: Member States face out **prohibited systems**

12 months: **General-purpose AI** model rules apply

24 months: **All other rules** of the AI Act apply

24 months: **High-risk** rules apply (**Annex III**)

36 months: **High-risk** rules apply (**Annex II**)

*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.

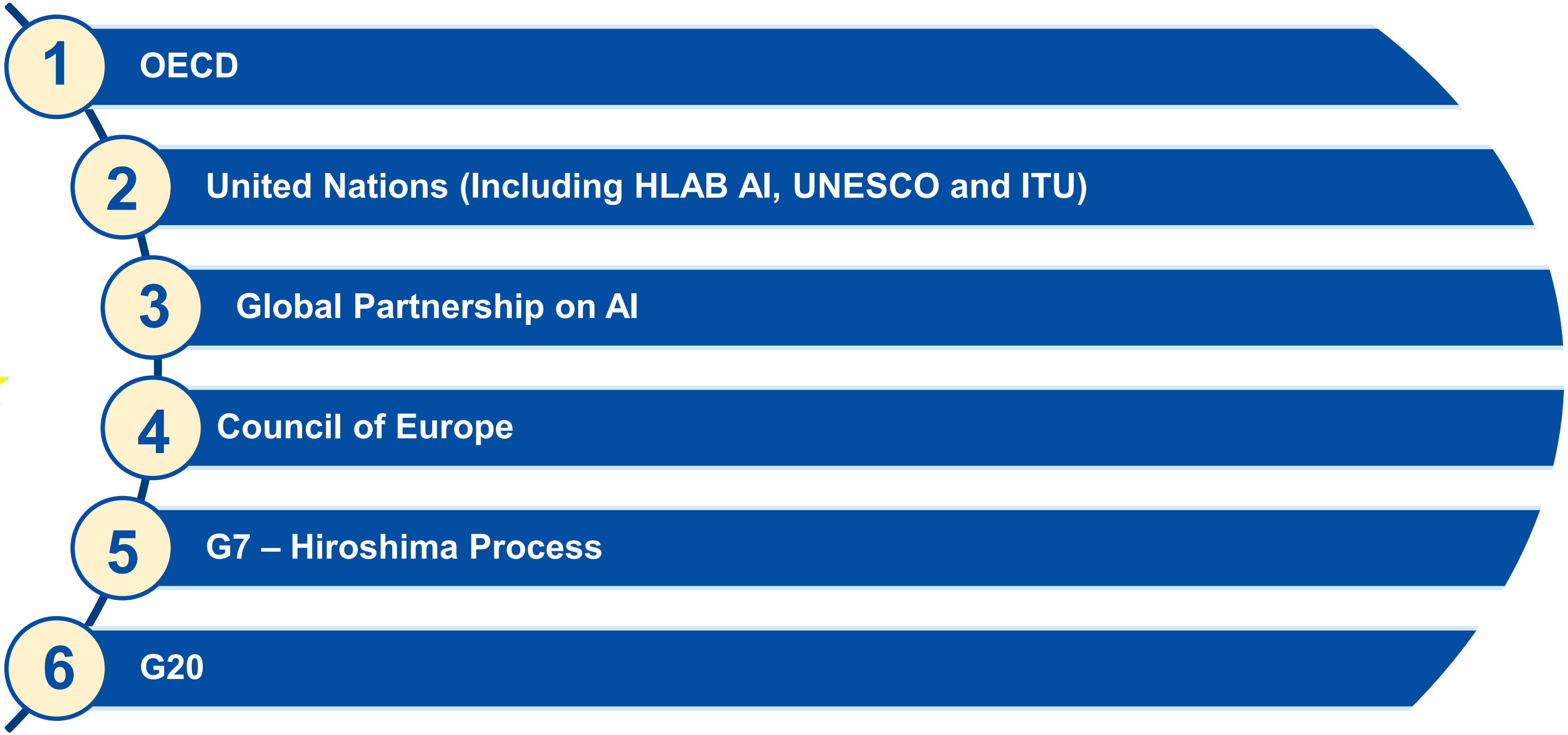
© iStock by Getty Images - 1139760401 peshkov

“

Multilateral cooperation plays an important role in the European approach to Artificial Intelligence

European Commission

International multilateral activities



1. OECD
2. United Nations (Including HLAB AI, UNESCO and ITU)
3. Global Partnership on AI
4. Council of Europe
5. G7 – Hiroshima Process
6. G20

Thank you

Elinor Wahal,
Legal and Policy Officer
DG CNECT A2, AI Policy Development and Coordination