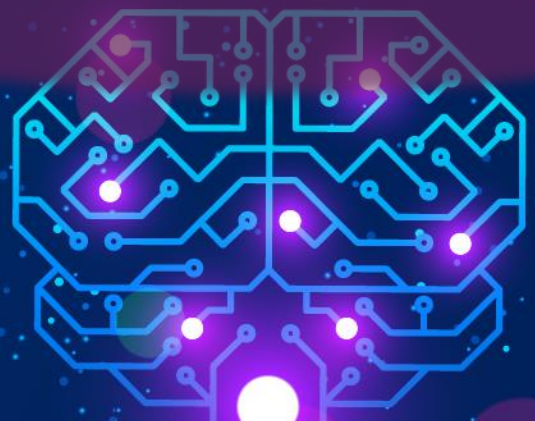
AI
ARTIFICIAL
INTELLIGENCE

EXCELLENCE
& TRUST



A European approach to Artificial Intelligence: The proposed AI Act

European Commission AI Policy 2018-2021

Key initiatives:

- **European Strategy on AI** (April 2018)
- **Guidelines for Trustworthy AI** developed in 2018/2019 by the High-Level Expert Group on Artificial Intelligence (HLEG), Followed by the ALTAI in 2020
- First **Coordinated Plan on AI** (December 2018)
- The Commission's **White Paper on AI** (February 2020)
  Ecosystem of trust & ecosystem of excellence
  Followed by a **public consultation**
- **AI package** (April 2021)

Proposal for the Artificial Intelligence Act
Why a EU regulation on AI?

Complexity — **Opacity**
Unpredictability —
Autonomy — **Data**

Solid framework of EU legislation already in place at EU and national level

HOWEVER

Certain **specific features of AI** can make application and enforcement of the existing rules more challenging and generate **risks to safety and fundamental rights**

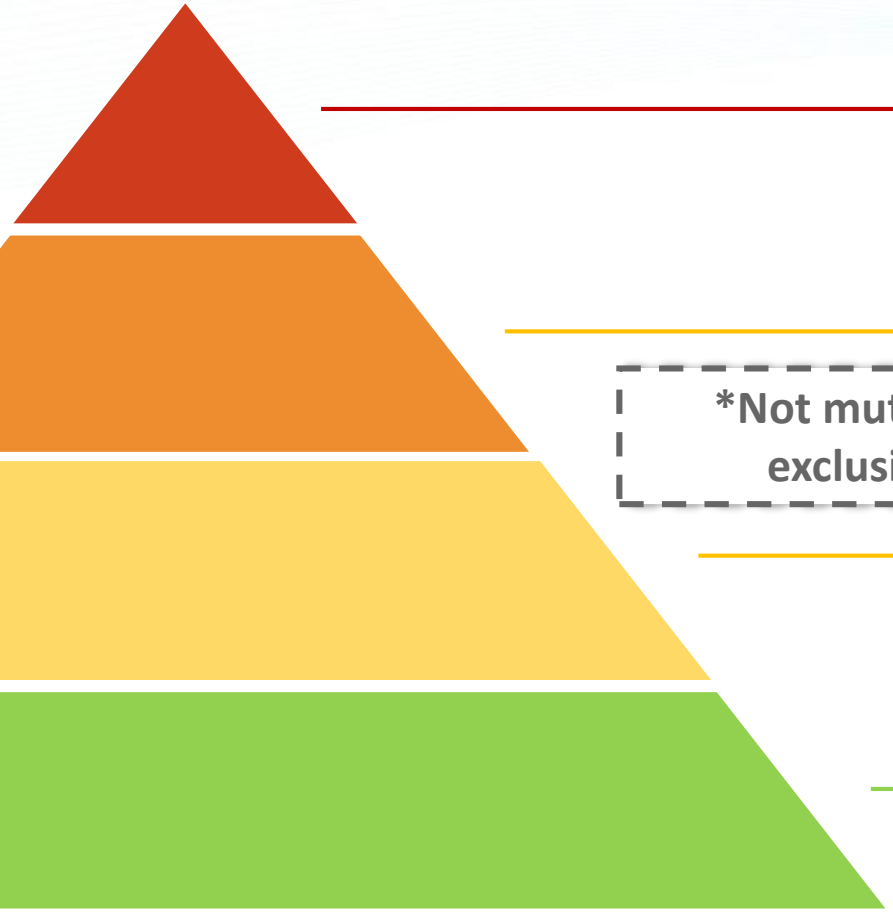
A **tailored regulatory response** needed

The Commission's **proposal for a regulatory framework on AI**

Proposal for the Artificial Intelligence Act- Foundations

- AI Act follows harmonisation legislation (Product Safety approach)
- Protection of safety and fundamental rights
- A risk based approach & Global level playing field
- A horizontal act
- Coherence and complementarity with existing legislation
- AI Act is innovation friendly

Proposal for the Artificial Intelligence Act
Risk-based approach: overview



Unacceptable risk
e.g. social scoring

Prohibited

High risk
e.g. recruitment, medical devices

Permitted subject to compliance with AI requirements and ex-ante conformity assessment

*Not mutually exclusive

'Transparency' risk
'Impersonation' (bots)

Permitted but subject to information/transparency obligations

Minimal or no risk

Permitted with no restrictions

Proposal for the Artificial Intelligence Act
Requirements for high-risk AI systems

Establish and implement **risk management** processes

&

In light of the **intended purpose** of the AI system

- Use high-quality **training, validation and testing data** (relevant, representative etc.)
- Establish **documentation** and design logging features (traceability & auditability)
- Ensure appropriate certain degree of **transparency** and provide users with **information** (on how to use the system)
- Ensure **human oversight** (measures built into the system and/or to be implemented by users)
- Ensure **robustness**, **accuracy** and **cybersecurity**

Standardisation

The compliance and enforcement system

Pre-market – conformity assessment

AI that is safety component of products
(regulated by product legislation)

→ **Third-party ex ante conformity assessment**
(already existing under the relevant sectoral legislation)

Other high-risk AI systems ("stand-alone")

→ **Ex ante conformity assessment through internal checks**
(except biometrics)

→ **Registration**
(EU database)

Post-market

- **Market surveillance** (authorities)
- **Human oversight and monitoring** (users)
- **Post-market monitoring** (providers)
- **Reporting system** for serious incidents brought to the attention of the providers
- **Re-assessment** by the provider in case of substantial changes to AI systems

The governance structure

European level

Commission to act as Secretariat

- Day-to-day business
- Prepare/follow-up Board meetings

Artificial Intelligence Board	Expert Group
One member per Member State, European Data Protection Supervisor and representative of the Commission	Not foreseen in the regulation but may be introduced in the implementation process Independent experts recruited and paid by the Commission for technical and scientific advice.

Main tasks

- Facilitate **consistent application of the legal framework** by Member States
- Contributing to **market monitoring**
- Collect and **share best practices**
- Contribute to **standards / AI policy**
- Provide advice on AI issues

National level

National Competent Authority/ies

Responsible for the application and implementation of the regulation, including for notified bodies, market surveillance

Main tasks

- Responsible **for implementation and application of the regulation** within each relevant Member State territory

AI
ARTIFICIAL
INTELLIGENCE

EXCELLENCE
& TRUST

Thank you!

Most AI systems will not be high-risk

New transparency obligations for certain AI systems (Art. 52)

- **Notify humans** that they are **interacting with an AI system** unless this is evident
- Notify humans that emotional recognition or biometric categorisation systems are applied to them
- Apply **label to deep fakes** (unless necessary for the exercise of a fundamental right or freedom or for reasons of public interests)

Possible voluntary codes of conduct for AI with specific transparency requirements (Art. 69)

- No mandatory obligations
- Commission and Board to encourage drawing up of codes of conduct intended to foster the **voluntary application of requirements to low-risk AI systems**

Proposal for the Artificial Intelligence Act
The scope – what is covered?

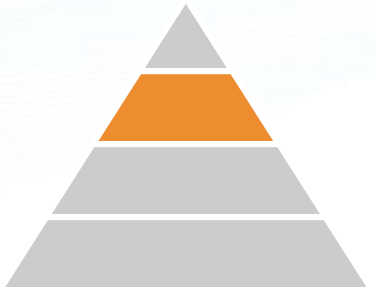
Definition of Artificial Intelligence

"a software that is developed with one or more of the techniques and approaches listed in Annex I (machine learning, logic and knowledge based, statistical approaches) and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with"

- Definition of AI should be **as neutral as possible** in order to cover techniques which are not yet known/developed
- **Overall aim is to cover all AI**, including traditional symbolic AI, Machine learning, as well as hybrid systems
- **Annex I**: list of AI techniques and approaches should provide for legal certainty. Adaptations over time may be necessary.

Proposal for the Artificial Intelligence Act
Risk-based approach: high risk AI systems

High-risk AI Systems



1 SAFETY COMPONENTS OF REGULATED PRODUCTS

(e.g. medical devices, machinery) which are subject to third-party assessment under the relevant sectorial legislation

2 CERTAIN (STAND-ALONE) AI SYSTEMS IN THE FOLLOWING AREAS

- ✓ Biometric identification and categorisation of natural persons
- ✓ Management and operation of critical infrastructure
- ✓ Education and vocational training
- ✓ Employment and workers management, access to self-employment
- ✓ Access to and enjoyment of essential private services and public services and benefits
- ✓ Law enforcement
- ✓ Migration, asylum and border control management
- ✓ Administration of justice and democratic processes

Overview: obligations of operators

Provider obligations

- Undergo **conformity assessment** and potentially re-assessment
- Establish and Implement **quality management** system
- Draw-up and keep up to date **technical documentation**
- Conduct **post-market monitoring**
- **Collaborate** with market surveillance authorities
- **Logging** obligations to enable users to monitor the operation of the high-risk AI system

User obligations

- Operate AI system in accordance with **instructions of use**
- Ensure **human oversight** when using of AI system
- **Monitor** operation for possible risks
- **Existing legal obligations** continue to apply (e.g. under GDPR)

Other operators with certain obligations to providers include:

- Authorised representatives
- Importers
- Distributors
- Other third parties substantially changing the system

Proposal for the Artificial Intelligence Act
AI that contradicts EU values are prohibited

✗ **Subliminal manipulation** resulting in physical/ psychological harm

Example: An **inaudible sound** is played in truck drivers' cabins to push them to **drive longer than healthy and safe**. AI is used to find the frequency maximising this effect on drivers.

✗ **Exploitation of children or mentally disabled persons** resulting in physical/psychological harm

Example: A doll with an integrated **voice assistant** encourages a minor to **engage in progressively dangerous behavior** or challenges in the guise of a fun or cool game.

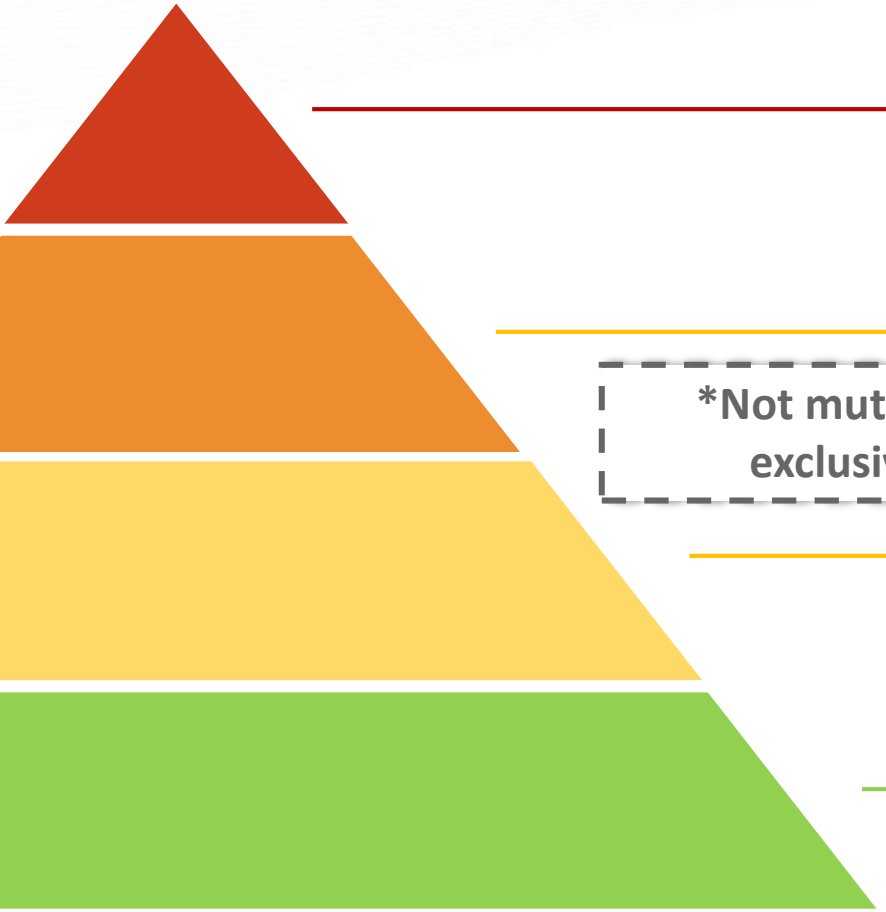
✗ **General purpose social scoring**

Example: An AI system **identifies at-risk children** in need of social care **based on insignificant or irrelevant social 'misbehavior'** of parents, e.g. missing a doctor's appointment or divorce.

✗ **Remote biometric identification for** law enforcement purposes in publicly accessible spaces (with exceptions)

Example: All faces captured live by video cameras checked, in real time, against a database to identify a terrorist.

Regulatory approach to biometrics



*Not mutually exclusive

Unacceptable risk
Real-time RBI systems for law enforcement purposes in publicly accessible spaces

Prohibited (with limited exceptions)

High risk
All RBI systems

Permitted subject to compliance with AI requirements and ex-ante conformity assessment

AI with specific transparency obligations
Emotional recognition and categorisation systems

Permitted but subject to information/transparency Obligations

Minimal or no risk
Biometric authentication/ verification
Closed set identification/ controlled environment

Permitted with no restrictions