The AI Act and the use of AI systems in the context of justice

Eike GRAEF

18th November 2021

Intro

- All data is about the past
- We cannot predict the future with certainty
- Opacity is a challenge when the stakes are high

Risks to fundamental rights

- Use of AI can pose risks in different ways, e.g. use for decision-making (HR) or for surveillance (RBI), or by adding opacity to existing processes.
- Key rights: data protection, privacy, **non-discrimination,** consumer protection, **good administration,** social security and assistance, freedom of expression, freedom of assembly, education, asylum, collective bargaining and action, fair and just working conditions, and to access preventive care.
- In the context of justice: **procedural rights**

Objectives of the AI proposal

Protect safety & fundamental rights, foster uptake.

Not the objective: creation of new rights.
Rather: help compliance and enforcement of existing acquis in view of opacity. Need for info to ascertain legality.

Challenge: horizontal proposal for a broad array of different technologies. (Annex IV preceded by „as applicable“).

Definition and technological scope of the regulation (Art. 3)

Definition of Artificial Intelligence

- 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)

> *"a software that is developed with one or more of the techniques and approaches listed in Annex I 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"*

Annex I

AI techniques and approaches covered

- Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;
- Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;
- Statistical approaches, Bayesian estimation, search and optimization methods.

High-risk Artificial Intelligence Systems (Title III, Annexes II and III)

Certain applications in the following fields:

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 FIELDS

- ✓ 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

Annex III

HIGH-RISK AI SYSTEMS REFERRED TO IN ARTICLE 6(2)

6. Law enforcement:

(d) AI systems intended to be used by law enforcement authorities for evaluation of the reliability of evidence in the course of investigation or prosecution of criminal offences;

8. Administration of justice and democratic processes:

(a) AI systems intended to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts.

Requirements for high-risk AI (Title III, chapter 2)

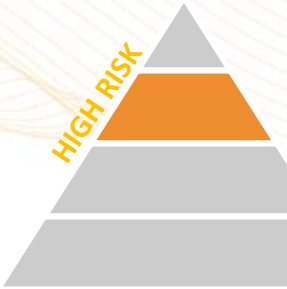
Establish and implement **risk management** processes

&

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**

Overview: obligations of operators (Title III, Chapter 3)



Provider obligations

- Establish and Implement **quality management** system in its organisation
- Draw-up and keep up to date **technical documentation**
- **Logging** obligations to enable users to monitor the operation of the high-risk AI system
- Undergo **conformity assessment** and potentially re-assessment of the system (in case of significant modifications)
- Register AI system in EU database
- Affix CE marking and sign declaration of conformity
- Conduct **post-market monitoring**
- **Collaborate** with market surveillance authorities

User obligations

- Operate AI system in accordance with **instructions of use**
- Ensure **human oversight** when using of AI system
- **Monitor** operation for possible risks
- **Inform the provider or distributor about any serious incident** or any malfunctioning
- **Existing legal obligations** continue to apply (e.g. under GDPR)

Art. 10 - Data and data governance

Relevant, representative, appropriate statistical properties, including as regards the persons on which the system is to be used

Take into account **specific geographical, behavioural or functional setting** within which the AI system is intended to be used

- **presumption of compliance (Art. 42)** if AI system has been trained and tested on data concerning that specific geographical, behavioural and functional setting

Legal basis to process special categories of personal data for the purposes of ensuring **bias monitoring, detection and correction**

Art. 11 & Annex IV - Technical documentation

Demonstrate compliance with requirements & enable authorities and notified bodies to assess such compliance

Drawn up **before AI system is placed on the market** and **kept up-to date**

Contain at least (Annex IV):

- general description of the AI system
- detailed description of **the elements of the AI system** and of the **process for its development**
- detailed **information about the monitoring, functioning and control of the AI system**
- detailed **description of the risk management system**
- (as applicable) description of any **change made to the system through its lifecycle**
- list of **harmonised standards applied** or description of **other technical solutions adopted**
- copy of the **EU declaration of conformity**
- detailed description of the **system to evaluate the AI system performance in the post-market phase**

Art. 13 - Transparency & information to users

(Users have to comply with fundamental rights obligations)

- ►**Instructions**
- ►**Characteristics, capabilities and limitations of performance** of the AI system
  - ►when appropriate, specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets used, taking into account the intended purpose of the AI system.
- ►**(…)**

Art. 64 - Access to data and documentation

- Enhances ex-post supervision in view of fundamental rights
  - Access to documentation (64.3)
  - Testing (64.5)
- Dynamic: „grows“ with Annex III
- Coordination to keep burden for users + providers low

Next steps and other activities

- Co-legislators discuss
- Standard setting bodies prepare their activities
- Important to ensure capacities for authorities
- DG JRC is doing research e.g. into risk assessment systems
- EU Justice Scoreboard keeps track of the digitalisation of national justice systems
- Various funding activities

European Commission

Thank you