Philipps Universität Marburg

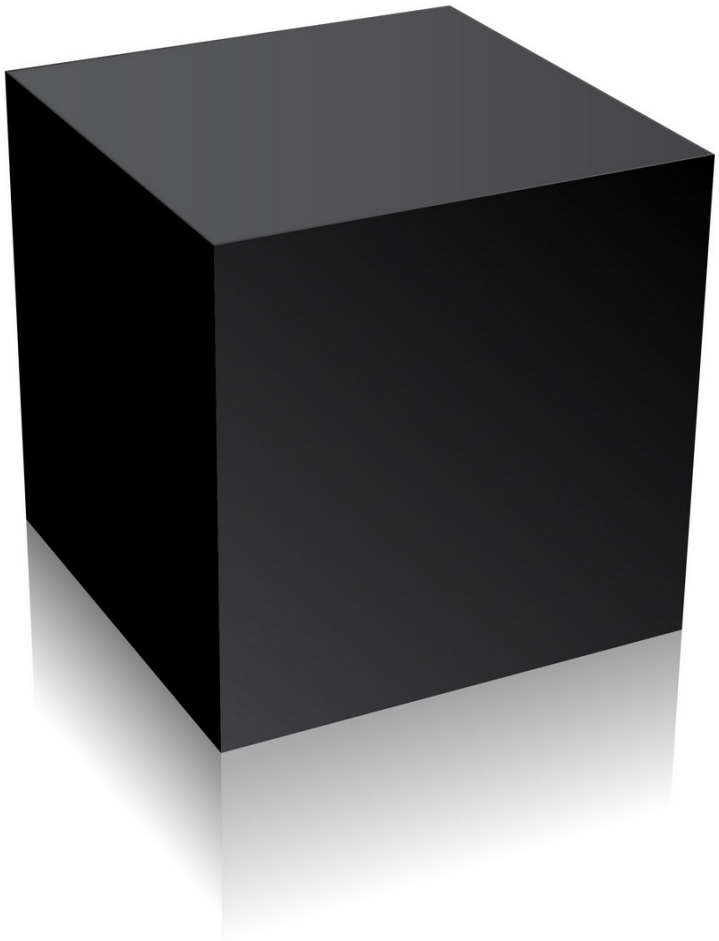
The Proposal for an EU Artificial Intelligence Act

Lecture Series on Ethical Implication of AI: Assessing Trustworthy AI in Practice – Seoul National University

Tuesday, Nov 8, 2022

Florian Möslein

How to regulate a black box?



How to regulate a black box?

:NU U.S. Chamber of Commerce

Above the Fold JUN 02, 2021 - 11:45AM

Europe's Gamble on AI Regulation

EVANGELOS RAZIS
Director, Center for Global Regulatory Cooperation

- **"In proposing the AI Act, European leaders seem to believe that their capacity and willingness to regulate is a competitive advantage over more innovative economies. This is a high-stakes gamble."**
- **Assumptions**
  - New Regulation Will Help, Not Hinder Europe's Competitiveness
  - Europe Should be the World's Leading AI Regulator
  - Handing over Proprietary Data, Source Code, and Algorithms to Regulators is a Good Idea
  - Europe Needs More Regulators and Regulation

Proposal for an EU Artificial Intelligence Act

- **Just a Proposal (for now!) – Timeline:**
  - April, 2021: Publication by EU Commission
  - March, 2022: JURI draft opinion by EU Parliament
  - September, 2022: Compromise text by EU Council
  - September, 2022: Draft opinion by EU Parliament
  - Mid-November, 2022: Latest version to be approved at the ambassador level; Plenary vote in the EU Parliament
  - December 2022: Trilogues begin
  - First half of 2023: Likely to be passed during the Swedish Council Presidency
- **Not an „Act“, but a Regulation:** EU Directives vs. EU Regulations
- **Cross-sectoral**
- **Public law, not (yet) private law**

Council of the European Union

Brussels, 3 November 2022
(OR. en)

13955/22

Interinstitutional File:
2021/0106(COD)

LIMITE

TELECOM 421
JAI 1369
COPEN 362
CYBER 337
DATAPROTECT 289
EJUSTICE 81
COSI 267
IXIM 246
ENFOPOL 520
FREMP 216
RELEX 1411
MI 772
COMPET 826
CODEC 1584

NOTE

From:	Presidency
To:	Delegations
No. prev. doc.:	13102/22
No. Cion doc.:	8115/21
Subject:	Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Preparation for Coreper

I. INTRODUCTION

1. The Commission adopted the proposal for a Regulation laying down harmonised rules on artificial intelligence **(Artificial Intelligence Act, AIA)** on 21 April 2021.

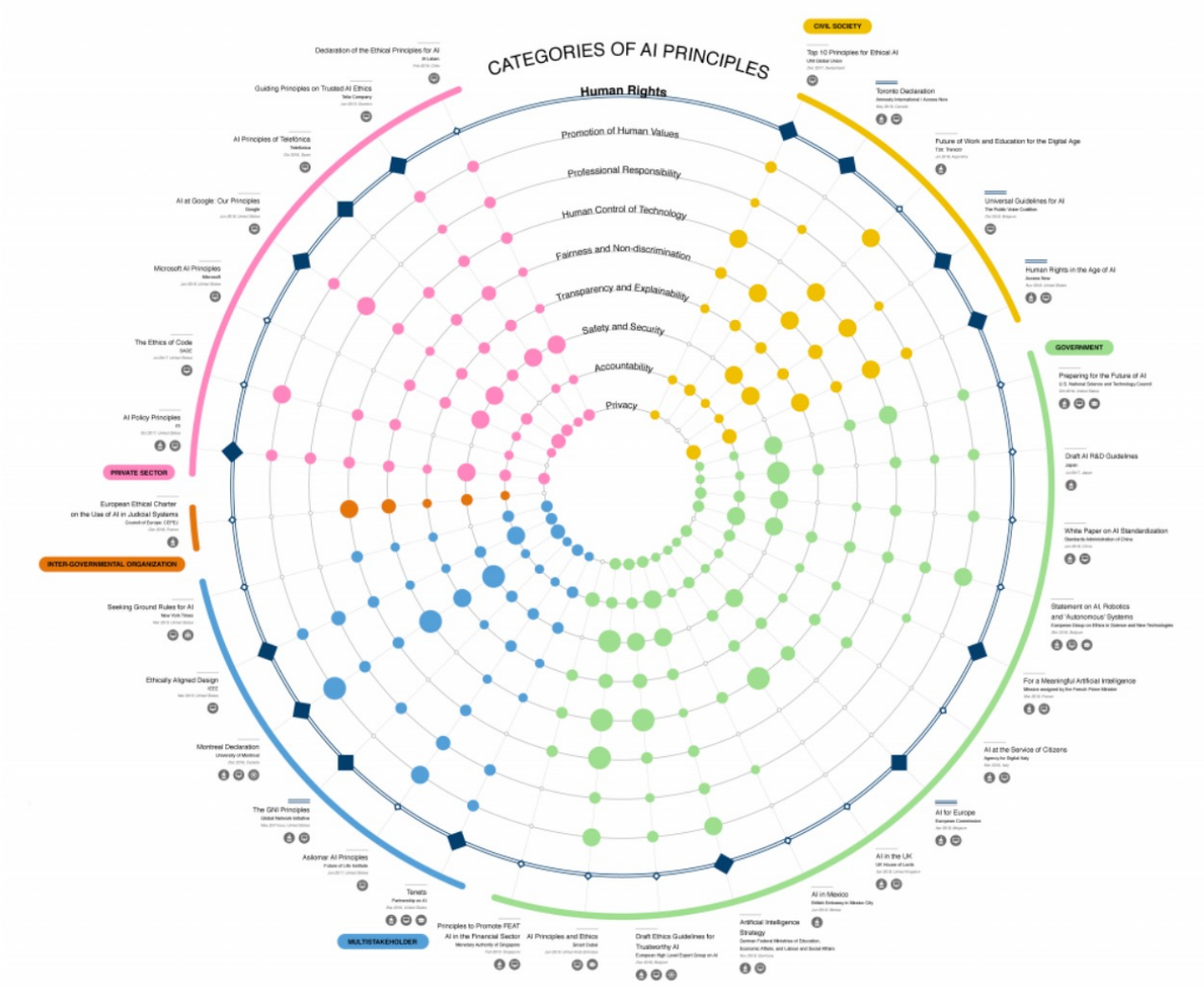
Data law framework

- **Digital Agenda for Europe, 2010**
  - Lisbon Strategy
- **Digital Single Market Strategy for Europe, 2015**
- **Regulatory framework for the Data economy**
  - Directive on certain aspects of the law of obligations relating to the provision of digital content and digital services (May 2019).
  - Regulation to promote fairness and transparency for commercial users of online intermediation services (P2B-VO, April 2019).
  - Data Governance Act (June 2022), Digital Markets Act (October 2022), Digital Services Act (still requires formal approval by the Council), Data Act (Proposal)
- **Sector-specific data economy legal acts, e.g. Digital Finance Package (September 2020), among others:**
  - Proposal for Regulation on DLT Pilot Regime for Market Infrastructures
  - Proposal for Digital Operational Stability Regulation for the Financial Sector (DORA)
  - Strategy for retail payments

Origins of the AI Act Proposal

- **Common European AI Strategy, COM (2018) 795 final.**
- **High-Level Expert Group, Ethics Guidelines for Trustworthy AI, April 2019 and Policy and Investment Guidelines for Trustworthy AI, June 2019.**
- **White Paper on Artificial Intelligence - A European Approach to Excellence and Trust, COM(2020) 65 final.**
- **European Parliament**
  - Special Committee on Artificial Intelligence in the Digital Age (in advance of the Commission's proposal).
  - Own proposals for EU rules on artificial intelligence (AI), in particular on ethics frameworks for AI, civil liability for AI-related damages, and intellectual property rights (October 2020 reports).
  - Report on Shaping Europe's Digital Future (May 2021)

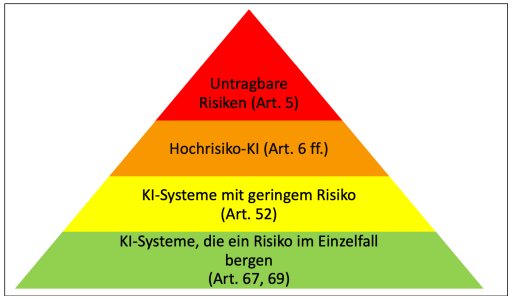
Standards diversity in the field of AI



(c) Berkman Klein Center - https://cyber.harvard.edu/story/2019-06/introducing-principled-artificial-intelligence-project

Basics and scope of application

- **Legal basis Regulatory character, regulatory objectives**
  - Art. 114 TFEU (internal market competence)
  - (P) concrete internal market barriers due to lack of national AI legislation currently not (yet) demonstrable
  - Regulation: directly applicable
  - Preventive prohibition law => Goal: "ecosystem of trust"; legal certainty
  - Risk-based approach
    - Prohibited practices (Art. 5)
    - High-risk systems (Art. 6 ff.)
    - Medium risk systems (Art.52)
    - Systems with low risk (Art. 69)



  - Additional measures to promote innovation (Title V)
    - AI reallabs (Art. 53), but with relatively strict minimum requirements; (P) obligation to comply with GDPR? Cf. Art. 54
    - Measures for small providers and small users (Art. 54), but no reduction of regulatory requirements, only priority access to AI reallabs, fee reduction, etc.

Basics and scope of application

- **Legal basis Regulatory character, regulatory objectives**
- **Material scope of application, cf. Art. 3 No. 1**
  - "Artificial Intelligence System" (AI System):
    - ***software that*** *has* ***been developed using one or more of the techniques and concepts listed in Annex I*** *and is* ***capable of producing results such as content, predictions, recommendations, or decisions that influence the environment with which it interacts, with respect to a set of*** *human-defined* ***objectives;***
  - Techniques and concepts according to the annex (power of amendment of the commission!):
    - Machine learning concepts, with supervised, unsupervised, and reinforcement learning using a wide range of methods, including deep learning;
    - Logic and knowledge-based concepts, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deduction engines, (symbolic) reasoning and expert systems;
    - Statistical approaches, Bayesian estimation, search and optimization methods.
  - (P) boundless expanse, almost any computer program => "software VO".

Basics and scope of application

- **Legal basis Regulatory character, regulatory objectives**
- **Material scope of application, cf. Art. 3 No. 1**
- **Personal scope**
  - "Provider, Art. 3 No. 2
    - *Natural or legal person, public authority, institution or other body that develops an AI system or has it developed with a view to placing it on the market or putting it into operation* under *its own name or trademark - whether in return for payment or free of charge;*
    - Cf. product safety regulations (manufacturing and placing on the market).
  - "User", Art. 3 No. 4
    - *natural or legal person, authority, institution or other entity that uses an AI system under its own responsibility, unless the AI system is used in the course of a personal and not professional activity;*
    - (P) Use of third-party infrastructures, e.g. IaaS
    - (P) Use of cloud applications (SaaS)

Basics and scope of application

- **Legal basis Regulatory character, regulatory objectives**
- **Material scope of application, cf. Art. 3 No. 1**
- **Personal scope**
- **Spatial scope**
  - (P) Migration possibility of "incorporeal" AI.
  - Scope of application according to art. 2 par. 1
    - *Providers who place AI systems on the market or put them into operation in the Union, regardless of whether those providers are established in the Union or in a third country;*
    - *Users of AI systems located in the Union;*
    - *Providers and users of AI systems established or located in a third country, if the result produced by the system is used in the Union.*
  - Market location principle; (P) location of AI systems at lit. b)? Hardware reference
  - Relevance of lit. c) => What is the result produced by a "system"? (In-)Indirectness reference?
  - (P) Networking of technical systems

Prohibited AI applications (Art. 5)

- **Legal concept: violation of fundamental values of the Union**
- **Case groups**
  - Behavioral influence (lit. a)/b): *AI system that uses techniques of subliminal influence outside of a person's awareness - or: exploits a weakness or vulnerability of a particular group of people due to their age or physical or mental disability - to substantially influence a person's behavior in a way that causes or is likely to cause physical or psychological harm to that person or another person.*
    - (P) threshold of significant influence, e.g. personalized advertising
  - Social scoring (lit. c): *AI system used by or on behalf of public authorities to assess or classify the trustworthiness of natural persons over a period of time based on their social behavior or known or predicted personal characteristics or personality traits.*
  - Remote biometric identification for law enforcement purposes (lit. d)

High-risk systems (Art. 6 et seqs.)

- **Categories (Art. 6)**
  - Cumulative conditions according to par. 1:
    - *AI system is intended to be used as a safety component of a product covered by Union harmonization legislation listed in Annex II or is itself such a product;*
    - *the product of which the AI system is the safety component, or the AI system itself as a product, shall be subject to a third party conformity assessment with regard to the placing on the market or putting into service of that product in accordance with the Union harmonisation legislation listed in Annex II.*
  - Notified systems according to par. 2:
    - *the AI systems listed in Annex III*
    - *E.g., biometric identification; access to essential private or public services; law enforcement; administration of justice.*
    - *partly duplications with exemptions according to Art. 2(2), e.g. flight and transport sector (?) drafting oversights*

High-risk systems (Art. 6 et seqs.)

- **Categories (Art. 6)**
- **Requirements**
  - Data and data governance (Art. 10):
    - Strict data quality requirements; avoidance of biases, etc.
  - Documentation and record-keeping requirements (Art. 11, 12)
    - Ex-ante technical documentation
    - Logging requirements ("logs") => ensure "*that the functioning of the AI system is traceable throughout its lifecycle to an extent appropriate to the system's intended purpose.*"
  - Transparency and provision of information to users (Art. 13)
    - *sufficiently transparent so that users can appropriately interpret and use the results of the system.*
    - (P) Black box AI
    - Meaning of Explainable AI (AI); cf. evidence-based studies.
  - Human supervision (Art. 14)
    - Develop systems in such a way *that they can be effectively supervised by natural persons during the period of use of the AI system (...).*
    - Understanding, monitoring, possibility of intervention (cf. par. 4).

High-risk systems (Art. 6 et seqs.)

- **Categories (Art. 6)**
- **Requirements**
- **Conformity assessment (Art. 19, 43)**
  - Differentiation depending on Art. 6-AI system: external or internal evaluation
  - (P) technical limitations of ex-ante testing of self-learning systems.
  - Re-examination according to Art. 43 Par. 4 only in case of significant changes
    - (P) Threshold value
    - (P) autonomous changes to the system => fully automated compliance testing?

Medium risk systems (Art.52)

- **"intended for interaction with natural persons"**
- **Special labeling requirements vis-à-vis natural persons**
  - Information about interaction with AI (No. 1)
  - Information about special properties (No. 2)

Systems with low risk (Art. 69)

- No special requirements of the AI Act
- Applicability of general legal provisions, e.g. GDPR
- Incentive to introduce codes of conduct, but on a voluntary basis

Enforcement mechanisms

- **European Artificial Intelligence Board (Art. 56 et seq.)**
  - National authorities and European Data Protection Supervisor
  - Consulting function
- **Market monitoring**
  - Robust monitoring
  - Among other things, introduction of an EU-wide database
  - MS authorities
- **Rules on fines, Art. 71**
  - Up to 30 million euros or 6% of global annual sales
  - Graduation according to violated standard
- **(!) No rights of complaint or enforcement for "AI affected persons".**
  - Difference to the GDPR
  - Need for supplementation by civil liability standards

Thank you for your attention.

INSTITUT FÜR DAS RECHT
DER DIGITALISIERUNG

PROFESSOR DR.
FLORIAN MÖSLEIN
LL.M. (LONDON)

MARBURG UNIVERSITY SCHOOL OF LAW
UNIVERSITÄTSSTR. 6
D-35032 MARBURG

T: +49 6421 28 – 21704
F: +49 6421 28 – 27046
MOESLEIN@IRDI.INSTITUTE

WWW.IRDI.INSTITUTE