EU REGULATION ON ARTIFICIAL INTELLIGENCE

OZNUR UGUZ

PHD RESEARCHER, SCUOLA SUPERIORE SANT'ANNA

INSTITUTE OF LAW, POLITICS AND DEVELOPMENT

Scuola Superiore Sant'Anna

Table of Contents

INSTITUTE OF LAW, POLITICS AND DEVELOPMENT
Scuola Superiore Sant'Anna

1. European Union AI Policies

INSTITUTE
OF LAW,
POLITICS AND
DEVELOPMENT

Scuola Superiore
Sant'Anna

European Commission

- Communication on Artificial Intelligence for Europe (European AI Strategy), 2018 – an ambitious plan of increasing investments, strengthening AI research and innovation, and facilitating access to data.
- Declaration of Cooperation on Artificial Intelligence, 2018 – EU Member States joined forces on AI to harness its potential and address the issues arising from the technology.
- Coordinated Plan on Artificial Intelligence, 2018 – sets out the European Union's strategic objectives and priorities for artificial intelligence.
- AI Watch, 2018 – monitors the development, uptake and impact of artificial intelligence for Europe.
- EU AI Alliance, 2018 – complements and supports the work of the AI High-Level Expert Group and provides input to European AI policy-making.

- High-Level Expert Group on Artificial Intelligence, 2018 – advises European Commission on its AI strategy.
  - Ethics Guidelines for Trustworthy AI, 2019
  - Policy and Investment Recommendations on Trustworthy AI, 2019
  - Assessment List for Trustworthy AI (ALTAI), 2020
  - Sectoral Considerations on the Policy and Investment Recommendations, 2020.
- White Paper on Artificial Intelligence, 2020 – lays out a future regulatory AI framework for the EU and contains specific actions for the support, development, and uptake of AI across the EU economy and public administration.
- AI, Data And Robotics Partnership in Horizon Europe, 2021 – provides strong leadership in the widespread deployment of AI, data and robotics in sectors and regions across Europe.

Council of Europe

- Ad Hoc Committee on Artificial Intelligence (CAHAI), 2019 – examines the achievability and possible elements of a potential AI legal framework.
  - Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law
- Committee on Artificial Intelligence (CAI), 2022 – established to produce a legal instrument on AI systems.
- European Committee on Legal Cooperation limited working group on administration and artificial intelligence (CDCJ-ADMIN-AI), 2022 – tasked with updating the Council of Europe handbook on "The Administration and You" in the light of the use AI and non-AI algorithmic systems in administrative law.
- Framework Convention on Artificial Intelligence, 2024 – aims to ensure that AI systems are fully consistent with human rights, democracy, and the rule of law.

European Parliament

- STOA Centre for AI, 2019 – established to contribute to the quality and coherence of discussion and policy-making for the coordination of the EU's efforts and influence on global AI standard-setting.
- Special Committee on Artificial Intelligence in The Digital Age, 2020 – studies the impact and challenges of AI.
- Resolution on a framework of ethical aspects of artificial intelligence, robotics and related technologies, 2020
- Resolution on a civil liability regime for artificial intelligence, 2020
- Resolution on intellectual property rights for the development of artificial technologies, 2020
- Resolution on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters, 2021
- Resolution on artificial intelligence in education, culture and the audiovisual sector, 2021

2. EU Artificial Intelligence Act

INSTITUTE
OF LAW,
POLITICS AND
DEVELOPMENT

Scuola Superiore
Sant'Anna

- Entered into force on August 1, 2024 and will be fully applicable by August 2, 2026.
- Introduced a uniform legal framework for the development, provision, deployment, and use of AI within the European Union.
- Takes a risk-based approach to AI and classifies AI systems according to the intensity and scope of the risks they could generate against people's health, safety or fundamental rights:
  - Unacceptable risk;
  - High-risk;
  - Minimal risk.

Unacceptable risk AI systems

- Unacceptable-risk AI applications are prohibited entirely or exempting certain circumstances.
- Article 5 of the AI Act lays out the prohibited AI systems:
  - deploying subliminal, manipulative, or deceptive techniques to distort behaviour and impair informed decision-making, causing significant harm (Art 5(1)(a));
  - exploiting vulnerabilities related to age, disability, or socio-economic circumstances to distort behaviour, causing significant harm (Art 5(1)(b));
  - social scoring, i.e., evaluating or classifying individuals or groups based on social behaviour or personal traits, causing detrimental or unfavourable treatment of those people (Art 5(1)(c));

- assessing the risk of an individual committing a criminal offence solely based on profiling or personality traits, except when used to augment human assessments (Art 5(1)(d));
- compiling facial recognition databases by untargeted scraping of facial images from the internet or CCTV footage (Art 5(1)(e));
- inferring emotions in workplaces or educational institutions, except for medical or safety reasons (Art 5(1)(f));
- biometric categorisation systems inferring sensitive attributes such as race, political opinions, trade union membership, religious or philosophical beliefs, sex life, or sexual orientation, except labelling or filtering of lawfully acquired biometric datasets or when law enforcement categorises biometric data (Art 5(1)(g));

- real-time’ remote biometric identification (RBI) in publicly accessible spaces for law enforcement (Art 5(1)(h)), except when;
  - searching for missing persons, abduction victims, and people who have been human trafficked or sexually exploited; or
  - preventing substantial and imminent threat to life, or foreseeable terrorist attack; or
  - identifying and localising suspects of serious crimes punishable in the concerned Member State for a maximum period of at least four years.

High-risk AI systems

- An AI system will be considered high-risk when (Art. 6):
  - used as a safety component or a product covered by EU laws in Annex I and required to undergo a third-party conformity assessment under those Annex I laws; or
  - referred to in Annex III, except where it does not pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including when an AI system;
    - performs a narrow procedural task; or
    - improves the result of a previously completed human activity; or
    - detects decision-making patterns or deviations from prior decision-making patterns and not meant to replace or influence the previously completed human assessment, without proper human review; or
    - performs a preparatory task to an assessment relevant to the purposes of the use cases listed in Annex III.

High-risk AI systems under Annex III

- Biometrics that are not prohibited under EU law,
- Those used as a safety component of critical infrastructure,
- Those used to determine access to education and training, monitor prohibited student behaviour, and evaluate learning outcomes,
- Those used for recruitment and making decisions affecting work-related relationships, promotion, termination, task allocation based on individual traits, monitoring and evaluating work performance and behaviour,
- Those used to determine access to and enjoyment of essential private services and essential public services and benefits,
- Those used in law enforcement unless prohibited under relevant EU or national law,
- Those used in migration, asylum, and border control management unless prohibited under relevant EU or national law,
- Those used to influence the outcome of an election or referendum or the voting behaviour of natural persons and used by judicial authorities or in dispute resolution to assist them with researching and interpreting the facts and law and application of law to the facts.

- Nevertheless, an AI system referred to in Annex III will always be considered high-risk where the AI system performs profiling of natural persons.
- High-risk AI systems can only enter the EU market or be put into use if they comply with certain mandatory requirements regarding compliance, risk management, data governance, human oversight, transparency, and accuracy listed under Section 2 of the AI Act (Art. 8-15).
- In addition to those, Section 3 contains obligations to be complied with by the providers, deployers, importers, distributors of high-risk AI systems and other parties in the value chain, including those pertaining to quality assessment, documentation, and fundamental rights impact assessment (Art. 16-27).

General-purpose AI models

- Systems based on general-purpose AI models, which display significant generality and are capable of competently performing a wide range of distinct tasks along with the ability to be integrated into various downstream systems or applications (Article 3(63)).
- Article 53 lays down obligations for general-purpose AI models:
  - Providing up-to-date technical documentation of the model, including its training and testing process and the results of its evaluation, which includes, at least, the information set out in Annex XI.
  - Providing up-to-date information and documentation to providers of AI systems who intend to integrate the general-purpose AI model into their AI systems which;
    - enables them to have a good understanding of the capabilities and limitations of the general-purpose AI model and to comply with their obligations under the Act, and
    - contains, at least, the elements set out in Annex XII.

General-purpose AI models with systemic risk

- Systemic risk refers to risks specific to such models' high-impact capabilities that match or exceed the capabilities recorded in the most advanced general purpose AI models in terms of their reach or actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole.
- A general-purpose AI model is classified as «with systemic risk» if it meets any of the following conditions (Art. 51):
  - it has high-impact capabilities evaluated on the basis of appropriate technical tools and methodologies, including indicators and benchmarks;
  - based on a decision of the Commission, ex officio or following a qualified alert from the scientific panel, that it has high capability or impact.

- Under Article 55 of the AI Act, providers of general-purpose AI models with systemic risk must;
  - perform model evaluation in accordance with standardised protocols and tools reflecting the state of the art,
  - assess and mitigate possible systemic risks that may stem from the development, deployment, and use of such systems,
  - keep track of, document, and report relevant information about serious incidents and possible corrective measures,
  - ensure an adequate level of cybersecurity protection for the model and its physical infrastructure.

Limitations and Loopholes

- Exclusion of limited-risk AI systems from the regulation
- Lack of a total ban on biometrics and emotion recognition systems
- Far-reaching exceptions that diminish the safeguarding effect of the regulations
- Many self-assessed and self-enforced obligations and exemptions
- Broad discretion granted to member states in implementation for national security purposes
- Weak and self-enforced standards for fundamental rights protection
- Narrow definition of general-purpose AI models with systemic risk and lenient approach towards their regulation

3. Framework Convention on Artificial Intelligence

INSTITUTE
OF LAW,
POLITICS AND
DEVELOPMENT

Scuola Superiore
Sant'Anna

- The first-ever international legally binding treaty aimed at ensuring the consistency of the use of AI systems with human rights, democracy, and the rule of law.
- Drafted by the 46 member states of the Council of Europe and 11 extra-EU countries.
- Adopted by the Council of Europe Committee of Ministers on May 17, 2024.
- Opened to signature on September 5, 2024.
- Open to ratification and compliance by both EU Member States and non-member countries.
- Enters into force on the first day of the month following the expiration of a period of three months after the date on which five signatories, including at least three Council of Europe member states, have ratified it.
- So far 37 countries signed the Convention.

- Applicable to the use of AI systems by public and private actors.
- Covers the activities within the lifecycle of artificial intelligence systems that have the potential to interfere with human rights, democracy, and the rule of law.
- Provides remedies (Art. 14), procedural safeguards (Art. 15), and risk and impact management requirements (Art. 16).
- Not required to apply to activities related to the protection of the national security interests of signatory states.
- Does not apply to;
  - national defence matters; and
  - research and development activities regarding artificial intelligence systems not yet made available for use unless they have the potential to interfere with human rights, democracy, or the rule of law.

Core Principles

- Human dignity and individual autonomy
- Equality and non-discrimination
- Respect for privacy and personal data protection
- Transparency and oversight
- Accountability and responsibility
- Reliability
- Safe innovation

Shortcomings

- A framework of «minimum and broadly defined standards» with a few rights and obligations.
- No strict compliance mechanism but rather a focus on monitoring, consultation, and cooperation.
- Grants signatory states discretion regarding the application of the Convention to private actors. They can either directly apply the principles and obligations of the Convention to the private sector entities or take «other appropriate measures» to fulfil the obligations.
- Individuals cannot bring their claim for the violation of the Convention directly before the European Court of Human Rights and will need to seek legal remedies at the domestic level.
- Leaves matters of national defence out of the scope while partially excluding the research and development activities and keeping the application of the Convention in the area of national security optional.

4. Recent European Case Law on AI

INSTITUTE
OF LAW,
POLITICS AND
DEVELOPMENT

Scuola Superiore
Sant'Anna

The Mevaluate Case

- Mevaluate Onlus Association v. Garante per la Protezione dei Dati Personali
- Considers the legality of the «reputation rating platform» operated by Mevaluate Onlus Association and owned by Mevaluate Holding Ltd., which assigns reputational scores to individuals based on algorithmic processing of their profiles.
- Through the platform, users could generate their own reputational profile or verify the credibility of third parties.

- In 2016, data protection authority Garante per la Protezione dei Dati Personali sanctioned the association on the ground that the platform violated the privacy legislation:
  - no legal basis for profiling, and
  - the consents of the non-users whose profile were processed to verify their credibility cannot be considered free.
- The association challenged the sanction before the Court of Rome.
- The court found only the profiling of the third parties illegitimate and partially annulled the sanction.
- The data protection authority filed an appeal against the decision before the Court of Cassation.

- **La Corte Suprema di Cassazione, Prima Sezione Civile (Civil Court of Cassation) ruling of 25 May 2021, n. 14381**
- In the appeal, the Court of Cassation found the profiling of the users also unlawful, stating that for the consent of users to be considered free and valid, the procedure and elements of the algorithm must have been known to the users.
- The ruling of the Court of Rome was quashed with a referral.
- In its new judgement, the Court of Rome rejected the association's appeal, saying that users should have been informed how the output is produced by the algorithm, indicating the «executive scheme» with which the rating is generated and the «specific weight» given to factors in the evaluation, including the interaction between them.
- The association appealed against the new decision.

- **La Corte Suprema di Cassazione, Prima Sezione Civile (Civil Court of Cassation) ruling of 10 October 2023, n. 28358**
- The question was whether the information provided by the association regarding the algorithm was adequate to make the consent of the data subjects valid.
- The association argued that the disclosure of the «executive scheme» required by the court of first instance for the consent to be valid was excessively broad that it corresponds to the revealing the mathematical functioning of the algorithm.
- The Court of Cassation held that what users must know ex-ante and with certainty is the procedure that leads to the final evaluations, not the elements such as the "specific weight" of the factors evaluated by the algorithm and found the data processing lawful based on valid consent.

The SyRI Case

- **Rechtbank den Haag (District Court of the Hague, Netherlands) ruling of 5 February 2020, C-09-550982**
- Concerns the use of the Risk Indication System (SyRI) by the Dutch government to detect fraud in citizens' access to social benefits, exemptions, and tax benefits.
- The system connects and analyzes data from various government agencies and public bodies, generating a risk report in the event of the identification of a citizen suspected of fraud.
- The applicants claimed that the system violates the right to privacy codified in Article 8 of the European Convention on Human Rights (ECHR), Article 7 of the EU Charter of Fundamental Rights, and Article 17 of the International Covenant on Civil and Political Rights (ICCPR) as well as the protection of personal data protected under EU and national legislation and Article 8 of the EU Charter.
- The state argued that the decision-making process of SyRI is based on objective criteria and that any infringement to privacy caused by the system is limited to what is strictly necessary.

- The Court assessed whether the measure;
  - had a legal basis,
  - had a legitimate aim,
  - was proportionate and necessary to the pursued aim.
- The court recognised the existence of a legal basis in Article 65 of the SUWI Act and the legitimacy of the purpose pursued by the legislator, which was preventing fraud in the interest of the economic well-being of the State.
- However, judges held that the procedures and guarantees of protection provided for by the legislation do not integrate the requirements of necessity and proportionality, assessed on the basis of the principle of transparency, the principle of purpose limitation, and the principle of data minimization.

- They found that:
  - neither the court nor individual citizens can ascertain the software's decision-making process, in violation of the principle of transparency.
  - The lack of accessibility and knowledge of the data and indicators used can lead to discrimination that cannot be corrected or challenged by citizens.
  - The large amount of data that can be processed under Art. 65 of the SUWI Act, without a specification of limits and necessity, entails a violation of the principle of data minimization.
- Accordingly, the Court declared the use of the system unlawful on the grounds that it violates Article 8 of the European Convention on Human Rights (ECHR).

The SCHUFA Holding Case

- **The Court of Justice (ECJ) ruling of 7 December 2023, C-634/21 (request for a preliminary ruling from Verwaltungsgericht Wiesbaden, Germany) - OQ v Land Hessen**
- Concerns whether the credit-scoring that decisively influences contractual relationships constitutes an automated decision-making process within the meaning of Art. 22 GDPR.
- The applicant OQ applied to SCHUFA Holding for a loan and was refused on the basis of an automated calculation of his creditworthiness.
- The applicant requested SCHUFA to provide them with a detailed account of the logic involved in the determination of their credit score, and the significance and consequences of the processing of their data, which was refused based on business secrecy.
- The applicant filed a complaint to the Commissioner for Data Protection and Freedom of Information for the State of Hesse (HBDI) and requested an order for the disclosure behind the reasoning of the decision as well as the significance and consequences of the processing of their data.

- HBDI dismissed the complaint, stating that the processing was compliant with domestic law. The applicant then appealed to the Administrative Court of Wiesbaden, under Article 78(1) GDPR.
- The Administrative Court of Wiesbaden made a preliminary reference to CJEU and asked:
  - Is a credit score issued by a third party considered a decision for the purposes of Article 22(1) GDPR, where the decision-making party draws strongly from it to reach a decision?
- The Court analysed three elements when assessing the applicability of Article 22 GDPR:
  - There must be a decision,
  - The decision must be based solely on automated processing, including profiling,
  - The decision must produce legal effects concerning the data subject, or similarly significantly affect them.

- ECJ stated that the conditions could be met at different times and by different parties and when the use of credit scoring by a third-party company decisively determines the stipulation, execution or termination of a contractual relationship, it constitutes an automated decision-making process under Article 22 of GDPR.
- The Court clarified that if the credit-scoring was considered a preparatory act and only the final decision based on the scoring was classified as a decision within the meaning of Article 22(1), there would be a risk of circumventing the safeguards provided under Article 22 of GDPR for the protection of the individuals subjected to automated processing of their data.

Thank you for your attention!

INSTITUTE OF LAW, POLITICS AND DEVELOPMENT

Scuola Superiore Sant'Anna