The EU AI Act: Caution against a potential "Ultron"

Srishti Dutta

7th January, 2024

BACKGROUND

The European Union Artificial Intelligence Act (The EU AI Act) was proposed essentially with the idea of laying down harmonised rules on artificial intelligence. Artificial Intelligence (AI) is a quick evolving branch of technology that can bring a wide array of economic and societal benefits across the socio-economic spectrum. Improving prediction, optimising operations and resource allocation and personalising service delivery, the use of artificial intelligence can support beneficial social and environmental outcomes, providing key competitive advantages to companies and the European economy. However, the same elements and techniques that power the socio-economic benefits of AI can also bring about new risks or negative consequences for an individual and in turn society at large. Considering the speed of technological change and possible challenges, the EU committed to come up with a balanced approach.

The European Council called for a 'sense of urgency to address emerging trends', in 2017, including 'issues such as artificial intelligence …, while at the same time ensuring a high level of data protection, digital rights and ethical standards'. In its 2019 "Conclusions on the Coordinated Plan on the development and use of artificial intelligence", the European Council highlighted the necessity of ensuring that its citizens rights are fully respected, in turn calling for a review of the existing relevant legislation in this field to make it more relevant with the new opportunities and challenges raised by AI. It proposed for further clarifications and determination of the AI applications that should and would be considered high-risk. The most recent conclusions further called for addressing the opacity, complexity, bias, degree of unpredictability and partially autonomous behaviour of certain AI systems, to ensure that they align and are compatible with the fundamental rights of citizens, and can also be brought under proper jurisdiction.

OBJECTIVES OF THE PROPOSAL

- ensuring that AI systems placed in the European Union market are used safely, respecting the existing laws on fundamental rights and Union values
- ensuring legal certainty in facilitating investment and innovation in AI and research related to it
- enhancing governance and effective enforcement of existing laws on fundamental rights and the safety requirements that may be applicable to AI systems

- facilitating the development of a singular market for lawful, safe and trustworthy AI applications to prevent the fragmentation of major and minor markets.

To achieve these objectives the EU AI Act proposal also presented a balanced regulatory approach to AI, that entails the minimum necessary requirements to address the risks and problems linked to AI, without unnecessarily constraining or hindering technological advancements and research in this field. It puts in place a proportionate regulatory system that revolves around a well-defined risk-based regulatory approach that does not create unnecessary restrictions to trade, whereby legal intervention is tailored to concrete situations, where there may be a justified cause for concern, or where such concern can reasonably be anticipated soon. The legal framework at the same time includes flexible mechanisms that enable it to be dynamically adapted as AI evolves and new concerning situations emerge. The promotion of AI-driven innovation is also closely linked to the Data Governance Act, the Open Data Directive and other initiatives under the EU strategy for data, which will establish trusted mechanisms and services for reusing, sharing and the pooling of data, which are essential for the development of data-driven AI models of high quality.

NOTION OF BIOMETRIC DATA

The notion of remote biometric identification system used in this Regulation should be defined functionally. It is defined as "**an AI system intended for the identification of natural persons at a distance through the comparison of a person's biometric data with the biometric data contained in a reference database, and without prior knowledge whether the targeted person will be present and can be identified, irrespectively of the particular technology, processes or types of biometric data used."**
A classification should be made between 'real-time' and 'post' remote biometric identification systems.

- In 'real-time' systems, the biometric data is captured, and the comparison and identification all occur instantaneously, near-instantaneously or at most without a significant delay. In this case, there should be no scope of finding a way around the rules of this Regulation on the 'real-time' use of the AI systems in question by providing for minor delays.
- In 'post' systems, the biometric data have already been captured and the comparison and identification occur only after a significant delay. This involves data, which has been generated before the use of the system in respect of the natural pupils concerned.

The use of Real Time Biometric Systems for the purpose of law enforcement should be prohibited, except in three exhaustively listed and narrowly defined situations:

1. Search for potential victims of crime, including missing children
2. Investigation regarding certain threats to the life or physical safety of natural persons or of a terrorist attack
3. Detection, localisation, identification or prosecution of perpetrators or suspects of the criminal offences referred to in Council Framework Decision 2002/584/JHA.

Only after specific authorisation by a judicial authority or an independent administrative authority of a Member State.

NOTION OF PUBLICLY ACCESSIBLE SPACES

The notion of publicly accessible spaces, referring to "**any physical place that is accessible to the public, irrespective of whether the place in question is privately or publicly owned.**" Hence the notion "**does not cover** places that are private in nature and normally not freely accessible for third parties, including law enforcement authorities, unless those parties have been specifically invited or authorised, such as homes, private clubs, offices, warehouses and factories. Online spaces are not covered either, as they are not physical spaces."
Determination of whether a given space is accessible to the public should however be determined on a case-by-case basis, having regard to the specificities of the individual situation at hand.

NOTION REGARDING AI SYSTEMS FALLING UNDER THE REGULATION

Considering the digital nature, the EU AI Act suggests AI systems should fall within the scope of this Regulation even when they are :

- not placed on the market
- not put into service
- not used in the Union

This is the case where an operator established in the EU contracts certain services to an operator established outside the EU, in relation to 'high risk' AI systems that may hold impact on the European citizen. To prevent the circumvention of this Regulation and ensuring an effective protection of natural persons located in the Union, "this Regulation should also apply to providers and users of AI systems that are established in a third country, to the extent the output produced by those systems is used in the Union."

NOTION REGARDING HIGH-RISK AI SYSTEMS

"High-risk" AI systems will include systems that pose a threat or unacceptable risks to Union Public interests provided and protected by Union Law.
These systems can only be placed in the Union market or into service if they comply with certain mandatory requirements enlisted in the EU AI Act. However, AI systems identified as high-risk should be limited only to those that have a significant harmful impact on the health, safety and fundamental rights of person in the Union, minimising any potential restriction to international trade, if any.
Following AI Systems will be considered high risk:

- ‘Real-time’ and ‘post’ remote biometric identification systems should be classified as high-risk considering the risks that they pose and should be subject to specific requirements on “logging capabilities and human oversight.”
- Systems managing and operating critical infrastructure
- AI systems used in education or vocational training, for determining access or assigning people to educational and vocational training institutions or to evaluate people on tests as part of or as a precondition for their education.
- AI systems used in employment, workers management and access to self-employment, notably for the recruitment and selection of persons or for making decisions on promotion and termination and for task allocation, monitoring or evaluation of persons in work-related contractual relationships
- AI systems used to evaluate the credit score or creditworthiness of an individual
- AI systems providing access to and enjoyment of certain essential private and public services and benefits necessary for people to fully participate in society or to improve one’s standard of living in the Member States.
- AI Systems designed to take actions by law enforcement authorities involving certain uses of AI systems are characterised by “a significant degree of power imbalance” and may lead to surveillance, arrest or deprivation of a natural person’s liberty as well as other adverse impacts on fundamental rights guaranteed in the Charter.
- AI systems used in migration, asylum and border control management.
- AI systems intended for the administration of justice and democratic processes

High quality training, validation and testing data sets require the implementation of appropriate data governance and management practices. The training, validation and testing of data sets should be relevant, representative and free of errors and complete in view of the intended purpose of the system along with having the appropriate statistical properties.

The stakeholders also agreed upon a risk-based framework as a better option than blanket evaluation and regulation of all AI Systems. The types of risks and threats should be based on sector by sector and in turn case by case approach. The risks shall also be calculated after considering the impact on human rights and safety of the EU citizen.

NOTION REGARDING TRANSPARENCY OF HIGH-RISK AI SYSTEMS

Having information on how high-risk AI systems have been developed and how they perform throughout their lifecycle is essential. Thus, keeping records and proper technical documentation becomes mandatory. Such information should include the “general characteristics, capabilities and limitations of the system, algorithms, data, training, testing and validation processes used as well as documentation on the relevant risk management system.” along with up-to-date technical documentation. Users should be able to interpret the system output and use it appropriately. Therefore, the AI Systems concerned should be accompanied by relevant documentation and

instructions of use and include concise and clear information (including in relation to possible risks to fundamental rights and discrimination). Cybersecurity will play a crucial role in ensuring that AI systems are resilient against attempts to alter their use, behaviour, performance or compromise their security properties by malicious third parties exploiting the system's vulnerabilities.
The legal person, defined as the provider, who had taken the responsibility for the placing of the AI System in the market or into service, regardless of whether that natural or legal person is the person who designed or developed the system, will be held responsible for it. To ensure a high level of trustworthiness of high-risk AI systems, those systems should be subject to a conformity assessment prior to their placing on the market or putting into servic. It is also considered appropriate that an AI system undergoes a new conformity assessment whenever a change occurs which may affect the compliance of the system with this Regulation or when the intended purpose of the system changes. Providers will be required to register their high-risk AI system in an EU database, established and managed by the European Commission.
High-risk AI systems should bear the CE marking to indicate their conformity with this Regulation so that they can move freely within the internal market. However, under certain conditions, rapid availability of innovative technologies may be crucial for health and safety of persons and for society, thus making it appropriate that under "exceptional reasons of public security or protection of life and health of natural persons and the protection of industrial and commercial property", Member States could authorise the placing on the market or putting into service of AI systems which have not undergone a conformity assessment.

NOTIONS REGARDING AI RESEARCH AND DEVELOPMENT

To ensure a legal framework that is innovation and research friendly, future-proof and resilient to disruption, competent national authorities from one or more Member States should establish artificial intelligence regulatory sandboxes to facilitate the development and testing of innovative AI systems under strict regulatory oversight before these systems are placed on the market or into service. The objectives of the regulatory sandboxes should be "to foster AI innovation by establishing a controlled experimentation and testing environment in the development and pre-marketing phase with a view to ensuring compliance of the innovative AI systems with this Regulation and other relevant Union and Member States legislation". To promote and protect innovation, it is important that the interests of small-scale providers and users of AI systems are considered as well.

NOTIONS REGARDING SETTING UP OF AN "EUROPEAN ARTIFICIAL INTELLIGENCE BOARD"

In order to facilitate a smooth, effective and harmonised implementation of this Regulation a European Artificial Intelligence Board would be established and would be responsible for a number of advisory tasks, including "issuing opinions, recommendations, advice or guidance on

matters related to the implementation of this Regulation, including on technical specifications or existing standards regarding the requirements established in this Regulation and providing advice to and assisting the Commission on specific questions related to artificial intelligence."

The Board shall be composed of the national supervisory authorities, represented by the head or equivalent high-level official of that authority, and the European Data Protection Supervisor. (Other national authorities may be invited to the meetings if the discussed issues are of relevance to them)

OTHER NOTIONS

- Effective protection of rights and freedoms of individuals across the Union, the rules established by this Regulation would apply to providers of AI systems in a non-discriminatory manner, irrespective of whether they are established within the Union or in a third country, and to users of AI systems established within the Union.
- The Regulations apply to "Union institutions, offices, bodies and agencies", providng AI Systems to the citizens. (Excluding military purposes defined under the Act)
- Research for legitimate reasons in relation to high risk AI systems should not be stifled by the Regulation, However, "such research does not amount to use of the AI system in human-machine relations that exposes natural persons to harm and such research is carried out in accordance with recognised ethical standards for scientific research."
- AI systems related to products that are not high-risk in accordance with this Regulation will not require to comply with the requirements set out` under the act and are nevertheless safe when placed on the market or into service

GENERAL PROVISIONS

The Regulation mentioned in the EU AI Act apply to:

- Providers of the AI systems in the Union (irrespective of whether those providers are established within the Union or in a third country)
- Users of these AI Systems based within the Union
- providers and users of AI systems that are in a third country, where the output produced by the system is used in the Union

The Regulations however do not apply to:

- AI systems developed or used exclusively for military purposes by the Union
- public authorities in a third country nor to international organisations falling within the scope of the EU AI Act

PROHIBITED ARTIFICIAL INTELLIGENCE PRACTICES

The prohibited practices pertaining to AI will remain prohibited under the Act:

- Putting into service AI Systems that deploy subliminal techniques beyond an individual's consciousness, that may in turn cause physical or psychological harm to the concerned person
- AI Systems that segregate and exploit the vulnerabilities of an individual or community of people based on their differences.
- Unfavourable treatment of certain people or groups in social contexts which are unrelated to the contexts in which the data was originally generated or collected by the AI Systems
- the use of ‘real-time’ remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement (unless otherwise excluded in the EU AI Act)