Decoding AI policy: AI governance in the European Union

GEORGINA NAVALLES

POLICY – BUSINESS ANALYST, INNOVATE UK KTN

November 2023

Table of Contents

Introduction

The Innovate UK BridgeAI programme aims to empower UK organisations to harness the power of artificial intelligence (AI) by funding and supporting innovators in agriculture and food processing, creative industries, construction, and transport, logistics and warehousing sectors. The programme is designed to help innovators assess, develop and implement trusted AI solutions; connect businesses with AI experts; and elevate their AI leadership skills through training and upskilling.

UK innovators and businesses interact and operate in multiple markets and geographies. Following the publication of our first article looking at the UK's AI regulation landscape, this second article in the series aims to help businesses gain a better understanding of the policy direction that the European Union is taking on AI.

*** This article is provided for general information purposes only. It does not constitute legal or formal advice in any form, and any information included in it should not be interpreted as such. For specific legal or formal advice, consult a qualified professional. ***

Context

In 2021, the European Union (EU) formally proposed the establishment of the world's first[1] comprehensive horizontal regulatory framework[2] for AI, known as the EU AI Act. The jurisdiction of the Act includes providers of AI systems in the EU independently from where the provider is located, deployers of AI systems based within the EU, and third-country providers and deployers "where the output produced by the system is used in the Union".[3]

UK businesses will be subject to the EU AI Act requirements when providing services or placing products in the EU market.

The Act is being developed alongside other major legislative pieces such as the Digital Services Act, the Digital Markets Act, and the Data Governance Act, and it has important overlaps with the 2016 General Data Protection Regulation (GDPR).

The EU AI Act overview

The purpose of the EU AI Act is to set a technology-neutral definition and classification of AI systems. Following a risk-based approach, the Act establishes a harmonised set of rules for

[1] https://www.adalovelaceinstitute.org/wp-content/uploads/2022/03/Policy-briefing-People-risk-and-the-unique-requirements-of-AI-18-recommendations-to-strengthen-the-EU-AI-Act.pdf
[2] Horizontal regulation refers to a regulatory framework that applies across sectors and applications
[3] https://www.brookings.edu/articles/machines-learn-that-brussels-writes-the-rules-the-eus-new-ai-regulation/

placing AI systems on the EU market with obligations for providers, deployers, distributors and importers.

This article is based on the EU AI Act version adopted by the European Parliament in June 2023. The Act is expected to be approved by the end of 2024 and to enter into force in 2027.

A new independent **European AI Office** will be responsible for monitoring the implementation of the EU AI Act.[4] The Act is set to be enforced with large penalties between 5 to 40 million EUR or 1% to 7% of global annual turnover for non-compliance.[5]

Nevertheless, the regulation will not apply to the research, non-real-world testing and development of AI systems prior to the products or services being in the market, provided that these follow fundamental rights[6] and relevant EU law. There are also exemptions for free and open-source licences, except for foundation models.[7]

In addition, the Act also includes measures to support innovation[8] such as the obligation for all EU states to have at least one AI regulatory sandbox.[9]

Definitions

The Act defines an '**AI system**' as: *'a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments'*[10]. This is an amended definition aligned with the one adopted by the OECD.

Actors defined in the Act[11]:

- **Provider** as '*a natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark, whether for payment or free of charge'.*
- **Importer** is '*any natural or legal person established in the Union that places on the market or puts into service an AI system that bears the name or trademark of a natural or legal person established outside the Union'*.

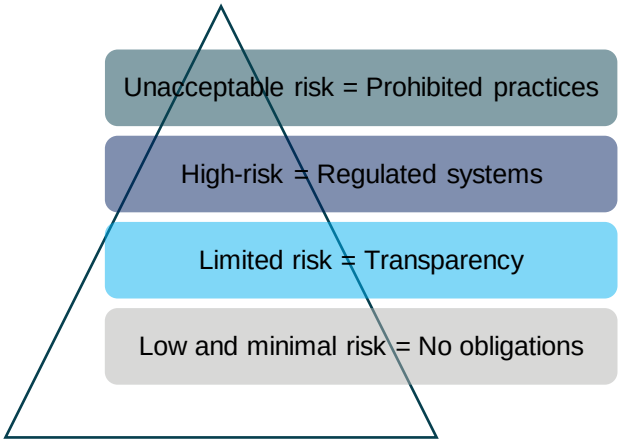
[4] EU AI Act P9_TA(2023)0236, Amendments 525-530
[5] Lilian Edwards. (2022). The EU AI Act proposal. Ada Lovelace Institute
EU AI Act P9_TA(2023)0236, Amendments 644 - 668
[6] Charter of Fundamental Rights of The European Union (2000)
[7] EU AI Act P9_TA(2023)0236, Amendments 163 and 164. Foundation models are defined in Amendment 168 as '*an AI system model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks'*.
[8] EU AI Act 2021/0106(COD), Title V
[9] EU AI Act P9_TA(2023)0236, Amendment 498
[10] EU AI Act P9_TA(2023)0236, Amendments 165
[11] EU AI Act 2021/0106(COD), Article 3; EU AI Act P9_TA(2023)0236, Amendment 172

- **Distributor** is '*any natural or legal person in the supply chain, other than the provider or the importer, that makes an AI system available on the Union market without affecting its properties'*.
- **Deployer** is '*any natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity*'.

The EU AI Act places most responsibilities and obligations on the **provider**, with some responsibilities on the **deployers** (i.e., a local authority using an AI fraud system scheme or an employer using an AI hiring system). The obligations on importers and distributors aim to prevent dangerous products from entering the EU market.[12]

Categorisation

The key section of the proposed EU AI Act is the classification of AI systems uses according to their risks, and the emanating legal requirements associated with these risks. The EU AI Act distinguishes the four following levels of risk:



1. Unacceptable risks – Prohibited AI practices[13]

The proposed Act will ban the practices that are considered a threat to people's safety, livelihood and rights due to their unacceptable level of risk associated, dubbed as AI practices that could cause *significant harm*. The banned practices will include:

- **Subliminal techniques**: AI systems that deploy subliminal[14] or purposefully manipulative or deceptive techniques in a way that could distort a person's behaviour causing significant harm to that person or others.[15]

[12] Lilian Edwards. (2022). The EU AI Act proposal. Ada Lovelace Institute.
[13] EU AI Act 2021/0106(COD), Title II
[14] Meaning: 'beyond a person's consciousness in order to materially distort a person's behaviour.'
[15] EU AI Act P9_TA(2023)0236, Amendment 215. This prohibition excludes subliminal techniques used for approved and consented therapeutical purposes.

- **Manipulation of vulnerable groups**: AI systems that exploit specific vulnerable groups (age, physical or mental ability, social economic situation or predicted personality traits), in a way that could cause significant harm to that person or others.[16]
- **Biometrics classification based on protected attributes**: Biometric classification systems that categorise natural persons according to sensitive or protected characteristics or based on the inference of those characteristics.[17]
- **Social scoring**: AI systems used for social scoring purposes leading to detrimental or unfavourable treatment of certain persons or groups for unrelated or disproportionate causes.[18]
- **Real-time biometric identification**: 'Real-time' remote biometric identification systems in publicly accessible spaces.[19]
- **Predictive policing systems**: AI system for making risk assessments of natural persons or groups to assess the risk of offending or reoffending or for predicting the occurrence or reoccurrence of an actual or potential criminal or administrative offence.[20]
- **Indiscriminate biometric online scraping:** AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage.[21]
- **Emotion recognition:** AI systems to infer emotions of a natural person in law enforcement, border management, in workplace and education institutions.[22]
- **Post remote biometric identification systems**: AI systems for the analysis of recorded footage of publicly accessible spaces through 'post' remote biometric identification systems (unless they are subject to a pre-judicial authorisation and strictly connected to a specific serious criminal offense).[23]

2. High-risk – Regulated high-risk AI systems[24]

The main focus of the Act is on high-risk AI systems, which will be extensively regulated. The proposed Act divides high-risk AI systems in two categories:

- AI systems used as a safety component of a product, or product itself, or falling under EU health and safety harmonisation legislation (i.e., toys, aviation, cars, medical devices, lifts).[25]

[16] EU AI Act P9_TA(2023)0236, Amendment 216
[17] EU AI Act P9_TA(2023)0236, Amendment 217. This prohibition excludes AI systems used for approved and consented therapeutical purposes.
[18] EU AI Act P9_TA(2023)0236, Amendments 218 and 219
[19] EU AI Act P9_TA(2023)0236, Amendments 220-223
[20] EU AI Act P9_TA(2023)0236, Amendment 224
[21] EU AI Act P9_TA(2023)0236, Amendment 225
[22] EU AI Act P9_TA(2023)0236, Amendment 226
[23] EU AI Act P9_TA(2023)0236, Amendment 227
[24] EU AI Act 2021/0106(COD), Title III
[25] EU AI Act 2021/0106(COD), Annex II-A and Annex II-B

- Systems deployed in specific areas should also be considered high-risk if they pose a significant risk of harm to natural persons established fundamental rights or the environment:[26]
    - **Biometric and biometric-based systems**[27]
    - **Management and operation of critical infrastructure** (i.e., road traffic but also utilities such as supply of water and gas)[28]
    - **Education and vocational training** admission and assessment[29]
    - **Employment, worker management and access to self-employment**[30]
    - **Access to and enjoyment of essential private services and public services and benefits** (i.e., healthcare, housing, electricity, credit checker, life insurance, emergency response systems, etc.)[31]
    - **Law enforcement** uses permitted under relevant EU and national law[32]
    - **Migration, asylum and border control management** (i.e., examination of evidence and documents veracity, risk assessment, etc.)[33]
    - **Administration of justice and democratic processes**, including AI systems used to influence voters or used in recommender systems of very large online platforms (VLOPs).[34]

If providers fall into one of the specific areas listed but consider their AI systems do not pose a significant risk, they will have to submit a reasoned notification to the relevant supervisory authority.[35]

Essential requirements for high-risk AI systems[36]

The main obligation for high-risk AI system **providers** will be to conduct a conformity assessment (prior to placing products or services on the market). If approved, providers will be able to add the 'CE' mark to their systems. The conformity assessment should be based on 'internal control' (self-assessment) except for the biometric identification use case, in which providers will have to hire a third-party ('notified body') to audit their conformity.

Providers will have to register their system in a publicly accessible EU-wide database.[37] The conformity assessment involves conformity with harmonised standards and with essential requirements covering:

[26] EU AI Act 2021/0106(COD), Identified in Annex III
[27] EU AI Act P9_TA(2023)0236, Amendments 710-712
[28] EU AI Act P9_TA(2023)0236, Amendments 713-714
[29] EU AI Act P9_TA(2023)0236, Amendments 715-718
[30] EU AI Act P9_TA(2023)0236, Amendments 719-720
[31] EU AI Act P9_TA(2023)0236, Amendments 721-724
[32] EU AI Act P9_TA(2023)0236, Amendments 725-731
[33] EU AI Act P9_TA(2023)0236, Amendments 732-727
[34] EU AI Act P9_TA(2023)0236, Amendments 72-73 and 738-740
[35] EU AI Act P9_TA(2023)0236, Amendment 235
[36] EU AI Act 2021/0106(COD), Title III, Chapter II
[37] EU AI Act 2021/0106(COD), Article 60; EU AI Act P9_TA(2023)0236, Amendments 567-573

- **Data and data governance**[38]: to address concerns about error and discrimination caused by biased, incomplete or erroneous data or assumptions.
- **Technical documentation**[39]
- **Record-keeping**[40]
- **Transparency and provision of information**[41]
- **Human oversight**[42]: AI system design and development must allow a human overseer to effectively identify anomalies, automated biases and be able to interpret the AI system outputs to override or disregard the system if necessary.
- **Accuracy, robustness, and cybersecurity**[43]

A risk management system must be established, implemented, documented, and maintained in relation to high-risk AI systems, throughout the entire lifecycle of the AI system. The aim is to identify, estimate and evaluate the known or reasonably foreseeable risks when the AI system is used as intended or under reasonably foreseeable misuse, to minimise them to a residual level. When risks cannot be eliminated adequate mitigation and control measures must be established and it must be communicated to the deployers.[44]

In addition, **providers established outside the EU** will have to, by written mandate, appoint an authorised representative which is established in an EU Member State.[45]

Lastly, **providers** will also have post-market obligations: They must establish and document a post-market monitoring system proportionate to the risks of the AI system through its lifetime. Furthermore, they must report any serious incident or malfunctioning no later than 15 days after becoming aware of it. [46]

Although the bulk of the requirements is the responsibility of providers, the Act also outlines the responsibilities of importers, distributors and deployers. Furthermore, **deployers**, **distributors** or **importers** become **providers** in terms of legal obligations if they modify substantially the AI system.[47] This also applies to deployers, distributors or importers of general purpose AI systems[48], when placed on the market or used in a manner that they become high-risk.[49]

[38] EU AI Act 2021/0106(COD), Article 10
[39] EU AI Act 2021/0106(COD), Article 11
[40] EU AI Act 2021/0106(COD), Article 12
[41] EU AI Act 2021/0106(COD), Article 13
[42] EU AI Act 2021/0106(COD), Article 14
[43] EU AI Act 2021/0106(COD), Article 15
[44] EU AI Act 2021/0106(COD), Article 9; EU AI Act P9_TA(2023)0236, Amendments 261-277
[45] EU AI Act P9_TA(2023)0236, Amendment 366
[46] EU AI Act 2021/0106(COD),Title VIII, Chapters I-II
[47] EU AI Act 2021/0106(COD), Article 28
[48] General purpose AI system is '*an AI system that can be used in and adapted to a wide range of applications for which it was not intentionally and specifically designed*'. EU AI Act P9_TA(2023)0236, Amendment 169
[49] EU AI Act P9_TA(2023)0236, Amendment 394

Requirements for providers of Foundation Models[50]

The European Parliament added explicit obligations for **providers of foundation models**[51] including reasonably foreseeable risks assessment and registry into an EU database.[52] Generative AI systems based on such models (i.e., ChatGPT) will have to comply with transparency requirements (i.e., disclosing that the content is AI generated, and helping differentiate deep fake[53] images from real ones), establish quality management systems, technical documentation, and ensure safeguards against generating illegal content. Summaries of the copyrighted data used for their training will also have to be made available.[54]

3. Limited risk – Transparency obligations[55]

The use cases subject to these obligations are:

- **Chatbots**: **providers** must ensure that AI systems intended to interact with natural persons are designed and developed in such a way that the natural person exposed to an AI system is informed they are interacting with an AI system in a timely, clear and intelligible manner, unless this is obvious from the circumstances and the context.[56]
- **Emotion recognition and biometric categorisation systems** (the types not included in the banned uses): **deployers** of such systems must inform in a timely, clear and tangible manner and obtain prior consent in accordance with the GDPR.[57]
- **Systems generating deep fake** or synthetic content: **deployers** of an AI system that generates or manipulates text, audio or visual content that would falsely appear to be authentic and which features depictions of people without their consent, must disclose in an appropriate, timely, clear and visible manner that the content has been artificially generated or manipulated, as well as, the name of the natural or legal person that generated or manipulated it where possible.[58]

[50] EU AI Act P9_TA(2023)0236, Amendment 399
[51] Foundation model is '*an AI system model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks'.* EU AI Act P9_TA(2023)0236, Amendment 168
[52] EU AI Act P9_TA(2023)0236, Amendment 399
[53] Deep fake is '*manipulated or synthetic audio, image or video content that would falsely appear to be authentic or truthful, and which features depictions of persons appearing to say or do things they did not say or do, produced using AI techniques, including machine learning and deep learning*'. EU AI Act P9_TA(2023)0236, Amendment 203
[54] https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai
[55] EU AI Act 2021/0106(COD), Title IV
[56] EU AI Act P9_TA(2023)0236, Amendment 484
[57] EU AI Act P9_TA(2023)0236, Amendment 485
[58] EU AI Act P9_TA(2023)0236, Amendment 486: Exceptions apply when authorised by law related to criminal offences.

4. Minimal risk – Codes of conduct[59]

Lastly, for minimal risk AI systems, such as **spam filters** or **videogames**, the EU AI Act encourages providers to draw up **voluntary codes of conduct** including: providing AI literacy among staff and users, assessing the impact of the AI system on vulnerable groups, diversity and equality and democratic processes, and evaluating if the AI system reinforces existing biases or inequalities, among others.[60]

What are the next steps?

EU legislators are still negotiating and amending some of the provisions of the EU AI Act, which is scheduled to be approved the end of next year. You can follow the latest developments on the EU AI Act in the European Commission dedicated site.

The **BridgeAI** programme offers a wide range of funding and support available for businesses, particularly in the agriculture and food processing, creative industries, construction, and transport, logistics and warehousing sectors, to facilitate the adoption of AI. Find out more about the upcoming events and resources available here.

Subscribe to the **BridgeAI** programme updates here.

[59] EU AI Act 2021/0106(COD), Title X
[60] EU AI Act P9_TA(2023)0236, Amendment 634