Università Europea di Roma

Co-funded by the European Union

JEAN MONNET CHAIR IN DIGITAL TRANSFORMATION AND AI POLICY

DATA REGULATIONS AND FUNDAMENTAL RIGHTS

Course Market Law and Regulation a.y. 2023-2024
Course convenor: Professor Valeria Falce (Valeria.Falce@unier.it)

Università Europea di Roma

Co-funded by the European Union

Module 3.

Università Europea di Roma

Co-funded by the European Union

DIGITAL DECADE VISION

Digital Decade Vision

The European Commission has updated the EU's digital strategy in light of the importance of digital technology for the economy and society, as the coronavirus pandemic has recently highlighted.

It builds on the 2020 strategy on shaping Europe's digital future, which remains the overarching framework, while reconsidering the enormous changes brought about by Covid-19.

The pandemic has massively accelerated the use of digital tools, demonstrating their opportunities while exposing society's vulnerability to new digital divides. In the post-coronavirus environment, the EU aims to protect and reinforce its digital sovereignty in strategic areas to ensure strategic autonomy in the digital area, while also promoting common EU values and respecting fundamental freedoms, including data protection and privacy, safety and security.

On 9 March 2021, the European Commission presented its vision for Europe's digital transformation by 2030. Its communication on the "2030 Digital Compass: the European way for the Digital Decade" announced an update of the Commission's overall digital strategy from February 2020 and of its gigabyte society targets, set in 2020 and 2016 respectively. This new strategy has been put forward to address a number of digital vulnerabilities revealed by the coronavirus crisis, such as dependency on non-European technologies. Europe should fund and support the development of sectors that are crucial to its digital sovereignty, such as semiconductors and edge computing.

The 2030 Digital Compass

The Commission has identified four main areas for action:

1 Achieve a digitally-skilled population and highly-skilled digital professionals;

2 Implement secure and performant sustainable digital infrastructures;

3 Achieve the digital transformation of businesses; and

4 Achieve the digitalization of public services.

Each of the four cardinal points of the digital compass relates to one of the four digital decade goals.

1. A digitally skilled population and highly skilled digital professionals:

At least 80% of all adults should have basic digital skills by 2030: this indicator follows the European Pillar of Social Rights action plan.

Reach 20 million employed ICT specialists in the EU, with convergence between women and men, compared to 7.8 million in 2019. Currently, more than 70 % of businesses report a lack of staff with adequate digital skills as an obstacle to investment. There is also a severe gender imbalance, with only one in six information and communication (ICT) specialists and one in three science, technology, engineering, and mathematics (STEM) graduates being women.

The 2030 Digital Compass

2. Secure and performant sustainable digital infrastructure:

By 2030, all European households should be covered by 5G, as well as by a fixed gigabit network. All European households should have gigabit connectivity compared to 59% in 2020 and all populated areas covered by 5G, up from 14 % in 2021. High performance computing (HPC) will require terabit connections to allow real-time data processing.

The production of cutting-edge and sustainable semiconductors in Europe, including processors, should represent at least 20 % of world production in value, doubling from 10 % in 2020.

10 000 climate-neutral highly secure edge nodes should be deployed in the EU and distributed in a way that guarantees access to data with low latency (i.e. few milliseconds), wherever businesses are located.

The quantum revolution in the next decade will be a game-changer in the emergence and use of digital technologies. By 2025, Europe should have its first computer with quantum acceleration, paving the way for Europe to place at the cutting edge of quantum capabilities by 2030.

The 2030 Digital Compass

3. Digital transformation of businesses:

The transformation of businesses will depend on their ability to adopt new digital technologies rapidly and across the board, including in industrial and services ecosystems that are lagging behind. Three out of four companies should use cloud computing services, big data and artificial intelligence by 2030.

More than 90 % of European SMEs should reach at least a basic level of digital intensity, compared to 61% in 2019.

Creation of around 250 unicorns (start-ups valued at US$1 billion) should be supported in the EU, a 100 % increase compared to 2021.

4. Digitalisation of public services:

All key public services should be available online.

All citizens will have access to their e-medical records.

80 % citizens should use a digital identity (ID) solution.

The 2030 Digital Compass

Figure 1 – Digital decade compass

Skills

ICT Specialists: 20 million + Gender convergence
Basic Digital Skills: min 80% of population

Public Services

Key Public Services: 100% online
e-Health: 100% availability medical records
Digital Identity: 80% citizens using digital ID

Infrastructures

Connectivity: Gigabit for everyone, 5G everywhere
Cutting edge Semiconductors: double EU share in global production
Data – Edge & Cloud: 10,000 climate neutral highly secure edge nodes
Computing: first computer with quantum acceleration

Business

Tech up-take: 75% of EU companies using Cloud/AI/Big Data
Innovators: grow scale ups & finance to double EU Unicorns
Late adopters: more than 90% of European SMEs reach at least a basic level of digital intensity

Source: European Commisssion: Europe's digital decade.

Digital principles and rights

The Commission therefore tabled a proposed **declaration on digital rights and principles for a human-centred digital transformation on 26 January 2022**, aiming at raising awareness and creating an overarching reference framework to govern this process.

The proposal builds on previous work done in this respect: the eGovernment (Tallinn Declaration), digital society and value-based digital government (Berlin Declaration), and digital democracy with a purpose (Lisbon Declaration). However, this new declaration is the first dedicated entirely to the fundamental rights of EU citizens in the digital environment.

The declaration would not be legally binding; it is an instrument to raise understanding of the EU acquis in the digital field. It derives from primary and secondary EU law and the CJEU and the European Court of Human Rights case law. The principles of the declaration are based on the EU Charter of Fundamental Rights and the EU Treaties, adapted to the digital environment. Existing fundamental rights are applied online, so that the exact same safeguards and rights for citizens are applied in the same way as offline.

Digital principles and rights

The draft declaration does not replace other proposals – instead it complements them. It also does not confer new rights; it is a collection of existing rights serving as a reference for public and private entities when dealing with new technologies and digital transformation. It is complementary to existing rights already introduced in the EU Charter on Fundamental Rights, General Data Protection Regulation (GDPR), and ePrivacy legislation, to name just a few examples. However, it introduces new issues, such as transparency of artificial intelligence (AI) algorithms – dealt with in the proposed AI act – which it compliments in this regard.

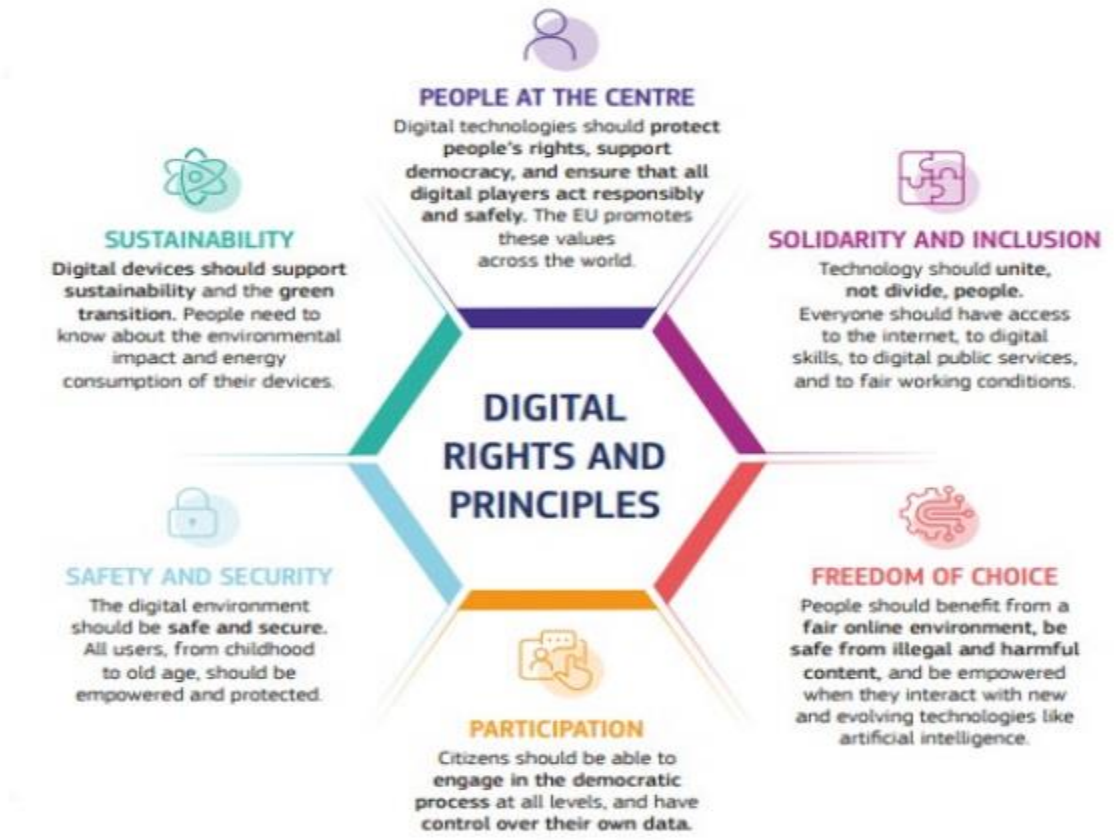
The draft declaration does not envisage direct enforcement. It however provides a framework for meeting the EU's digital decade targets and envisages an annual assessment of the digital transition.

Its adoption could however enable initiating legislation to transform rights into enforceable legal instruments. As European Commissioner Margrethe Vestager notes, the principles of the declaration provide "a blueprint for the digital transition".

As such, the proposed declaration is above all a political document, combining the policy and constitutional approaches and has primarily an advocacy role aimed at raising public awareness as well as promoting digital rights worldwide.

Digital principles and rights

Figure 2– Digital rights and principles



Source: European Commission: Digital Rights and Principles Factsheet.

Università Europea di Roma

Co-funded by the European Union

THE CASE OF THE ARTIFICIAL INTELLIGENCE ACT («AI ACT»)

The case of the AI Act

Co-funded by the European Union

On 1 August 2024, the European Artificial Intelligence Act (AI Act) enters into force. The Act aims to foster responsible artificial intelligence development and deployment in the EU.

Proposed by the Commission in April 2021 and agreed by the European Parliament and the Council in December 2023, the AI Act addresses potential risks to citizens' health, safety, and fundamental rights. It provides developers and deployers with clear requirements and obligations regarding specific uses of AI while reducing administrative and financial burdens for businesses.

Recently, the Commission has launched a consultation on a Code of Practice for providers of general-purpose Artificial Intelligence (GPAI) models. This Code, foreseen by the AI Act, will address critical areas such as transparency, copyright-related rules, and risk management. GPAI providers with operations in the EU, businesses, civil society representatives, rights holders and academic experts are invited to submit their views and findings, which will feed into the Commission's upcoming draft of the Code of Practice on GPAI models.

The provisions on GPAI will enter into application in 12 months. The Commission expects to finalize the Code of Practice by April 2025. In addition, the feedback from the consultation will also inform the work of the AI Office, which will supervise the implementation and enforcement of the AI Act rules on GPAI.

The case of the AI Act

Co-funded by the European Union

The EU AI Act introduces a sophisticated ‘product safety regime’ constructed around a set of 4 risk categories. It imposes requirements for market entrance and certification of High-Risk AI Systems through a mandatory CE-marking procedure. This pre-market conformity regime also applies to machine learning training, testing and validation datasets.

The AI Act combines a risk-based approach based on the pyramid of criticality, with a modern, layered enforcement mechanism. This means that as risk increases, stricter rules apply. Applications with an unacceptable risk are banned. Fines for violation of the rules can be up to 6% of global turnover for companies.

The EC aims to prevent the rules from stifling innovation and hindering the creation of a flourishing AI ecosystem in Europe, by introducing legal sandboxes that afford breathing room to AI developers.

The case of the AI Act

Co-funded by the European Union

The EU AI Act sets out horizontal rules for the development, commodification and use of AI-driven products, services and systems within the territory of the EU. The draft regulation provides core artificial intelligence rules that apply to all industries.

The EU AI Act introduces a sophisticated ‘product safety framework’ constructed around a set of 4 risk categories. It imposes requirements for market entrance and certification of High-Risk AI Systems through a mandatory CE-marking procedure. To ensure equitable outcomes, this pre-market conformity regime also applies to machine learning training, testing and validation datasets.

The Act seeks to codify the high standards of the EU trustworthy AI paradigm, which requires AI to be legally, ethically and technically robust, while respecting democratic values, human rights and the rule of law.

CONTENTS

THE NUMBERS OF THE WORLD AI MARKET

1.9 TRILLION BY 2030

According to the latest estimates provided by Statista, the global AI market has been valued at over EUR 130 billion in 2023 and is expected to grow substantially to almost EUR 1.9 trillion by 2030.

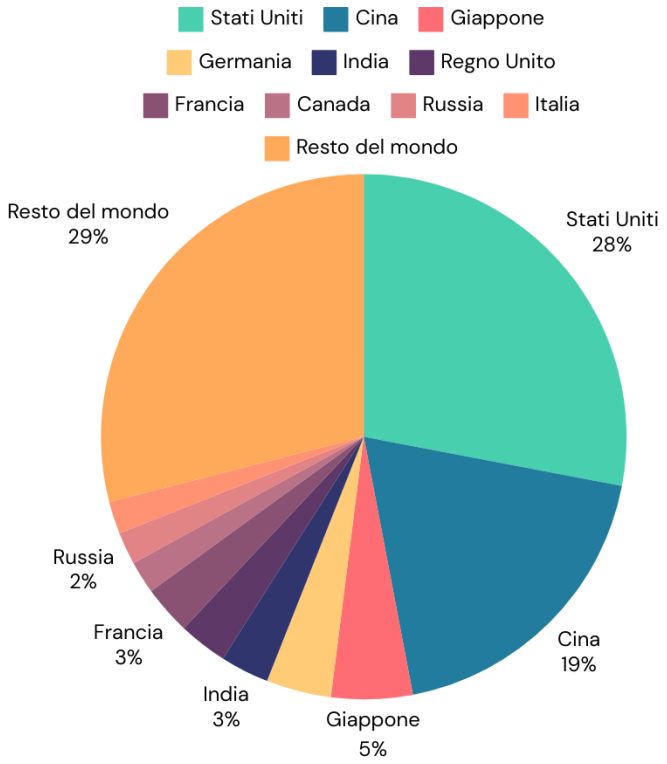
PREDOMINANCE OF PRIVATE INVESTMENT

Private investment accounts for the majority of investments in AI.

120 BILLION IN US COMPANIES

Between 2018 and the third quarter of 2023, almost EUR 32.5 billion was invested in EU AI companies, compared to more than EUR 120 billion in US AI companies.

THE WORLD AI MARKET BY COUNTRY (IN % OF TOTAL VALUE, 2024)

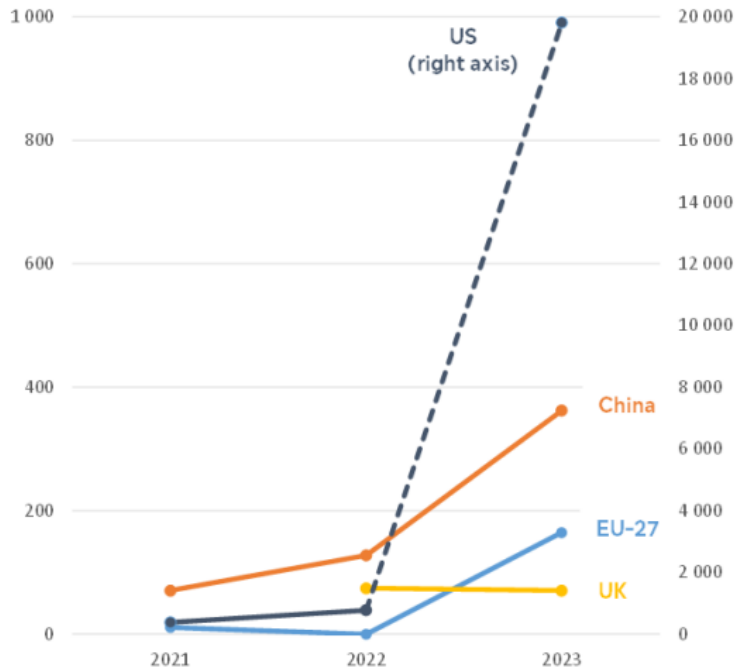


SOURCE: I-COM

PRIVATE INVESTMENT IN AI BY COUNTRY, 2023 (BILLION EURO)

Source: Stanford University, 2024 AI Index Report

VENTURE CAPITAL INVESTMENTS IN GENERATIVE A BY COUNTRY (MILLIONS OF EURO)



Source: OECD/Preqin, 2024

THE NUMBERS OF THE EUROPEAN AI MARKET

2.1 BILLION EURO INVESTMENT

Public investment in AI is growing. The EU's Digital Europe programme will fund AI with a total of EUR 2.1 billion over the period 2021-2027.

42 BILLION EURO MARKET

Statista indicates that the AI market in Europe is expected to stand at just over EUR 42 billion by the end of 2024, almost doubling the value of the market compared to 2020. The market is then expected to grow further, adding over EUR 190 billion by 2030.

ACCESS TO EUROPEAN FUNDING

In January 2024, the EU introduced measures to support European start-ups and SMEs in the development of reliable AI by granting access to funding, including the VentureEU, Horizon Europe, Digital Europe, EIC accelerator and InvestEU programmes.

THE 4 MAIN INDUSTRIAL SECTORS THAT HAVE ADOPTED AI TO DATE

As of 2023, the banking industry has seen a significant increase in AI adoption of 43%, transforming customer service, enhancing security and increasing operational efficiency. Financial institutions now leverage AI to provide tailored banking experiences and implement sophisticated fraud detection systems.

The IT sector, with an AI adoption rate of 13.8 per cent, is a crucial driver of AI integration, especially in areas such as cybersecurity, data analytics and software development. In addition to infrastructure, AI drives advances in cloud computing, data privacy and user experience.

The integration of AI in healthcare has significantly transformed medical diagnostics. Algorithms analyse medical images with greater speed and accuracy, aiding doctors in the early diagnosis of diseases. By processing large data sets, AI identifies patterns not noticed by humans.

An AI adoption rate of 12 per cent indicates the advent of smart manufacturing, characterised by AI-driven robotics, predictive maintenance and optimised supply chains. AI leads to greater efficiency and sustainable practices, highlighting its transformative role in manufacturing.

THE EUROPEAN APPROACH TO AI

The European approach to AI is inspired by two principles: technological sovereignty for strategic autonomy and the centrality of people in digital transformation. The objective is twofold: enhancing research and industrial capacity while guaranteeing fundamental rights.

However, the EU remains a secondary player in the development of AI and suffers from chronic delays in innovation. Lack of investment, incomplete single market, unattractiveness for talent, data scarcity, and regulatory complexity hinder the EU's emergence as a technological powerhouse.

A second major brake is the absence of an innovation ecosystem for European AI excellence. Among the 20 largest tech companies, only three are European (Accenture, SAP and ASML).

THE EUROPEAN STRATEGY ON AI

The Commission's AI strategy was launched with the adoption of the communication 'Artificial Intelligence for Europe' in April 2018.

The main assumption behind the strategy is that Europe can lead the way in the development and use of AI for the benefit of all, building on its values and strengths.

The European AI strategy is based on three distinct but complementary commitments:

- Increase investment to a level that matches the economic weight of the European Union;
- Leave no one behind - with particular reference to education - and ensure a smooth transition to the AI era in the workplace;
- Ensure that new technologies reflect European values.

THE TURNING POINT OF THE EU AI STRATEGY

The EU's AI le strategy reached a turning point in December 2019 with the arrival of the new European Commission led by Ursula von der Leyen. Following the appointment of Thierry Breton as Commissioner for the Single Market, the Commission also intensified its efforts on the European Data Strategy.

IOn 19 February 2020, the Commission launched a comprehensive package containing its ideas and actions on digital transformation, including a White Paper on Artificial Intelligence and a European Data Strategy.

The package marks another step forward in Europe's quest for ‘human-centric’ AI.

A EUROPEAN APPROACH TO AI

In 2021, the Commission is publishing a Communication on the promotion of a European approach to artificial intelligence.

The Communication includes 4 main objectives:

- Establish favourable conditions for the development and adoption of AI in the EU;
- Make the EU the place where excellence thrives ‘from the lab to the market’;
- Ensure that AI serves people as well as being a positive factor for society;
- Establishing strategic leadership in high-impact areas.

Co-funded by the European Union

Università Europea di Roma

THE EUROPEAN REGULATION ON ARTIFICIAL INTELLIGENCE: THE AI ACT

OBJECTIVE 1

Establishing favourable conditions for the development and adoption of AI in the EU

- Acquiring, pooling and sharing strategic information
- Exploiting the potential of data
- Promoting critical computing skills

OBJECTIVE 2

Making the EU the place where excellence thrives 'from the lab to the market'

- Collaborate with stakeholders, e.g. the European Partnership on AI, Data and Robotics and expert groups
- Build and mobilise research capacity
- Provide an environment in which developers can test and experiment and SMEs and P.A. can adopt AI
- Fund and scale up innovative AI ideas and solutions

OBJECTIVE 3

Ensuring that AI serves people

- Cultivate talent and improve the supply of skills needed to enable a thriving AI ecosystem
- Develop a strategic framework to ensure trust in AI systems
- Promote the EU's vision for sustainable and trusted AI to the world

OBJECTIVE 4

Establishing strategic leadership in high impact sectors

- Using AI in climate and environment
- Using the next generation of AI to improve health
- Preserving Europe's leadership: A strategy for robotics in the AI world

THE EU AI ACT

APRIL 2021

In April 2021, with a risk-based approach, the Commission presented its proposal for a ‘future-proof’ Artificial Intelligence Act, which establishes horizontal rules on AI, focusing on damage prevention.

MARCH 2024

On 13 March 2024, the European Parliament passed the AI Act, which became the world's first AI regulation.

MAY 2024

On 20 May 2024, the EU Council gave final approval to the AI Act, which will enter into force twenty days after its publication in the EU Official Journal.

JULY 2024

On 12 July 2024, the Artificial Intelligence Act, (Regulation (EU) 2024/1689) was published in the EU Official Journal.

KEY OBJECTIVES (ART. 1; RECITALS 1-8)

- "Improve the functioning of the internal market by laying down a uniform legal framework" for the development, placing on the market, commissioning and use of AI systems in the EU.
- "Promote the deployment of human-centric and trustworthy artificial intelligence", centred on respect for EU values, ensuring a high level of protection of health, safety, the environment, democracy, the rule of law and the fundamental rights enshrined in the Charter (set out in recital 48).
- Preventing and mitigating the risks of AI by prohibiting or restricting the use of AI systems that present unacceptable risks to the safety, health, dignity or autonomy of individuals, or that violate democratic values.
- Supporting innovation, with a focus on SMEs, including start-ups, by providing priority access to regulatory sandboxes, reduced fees for conformity assessment and simplified forms for technical documentation for high-risk AI systems.

SCOPE OF APPLICATION (ART. 2; RECITALS 22, 24, 25)

- The AI Act does not apply to areas outside the scope of EU law.
- The regulation will not apply to AI systems that have “military, defence or national security purposes, regardless of the type of entity carrying out these activities”, nor to AI systems used exclusively for research and innovation purposes, nor to persons using AI for non-professional purposes.
- The regulation will apply to deployers of AI systems who place such systems on the EU market, as well as to operators, even if located outside the EU, if the output produced by the AI system is used in the EU.
- Importers, distributors, manufacturers and authorised representatives of AI systems are also included in the scope. Systems used in commercial activities, systems addressed to natural persons, both embedded and stand-alone systems.

AI SYSTEM (ART. 3(1); RECITAL 12)

- Automated ("machine-based system").
- Designed to operate with varying levels of autonomy.
- Can exhibit "adaptability to learn new, distinct tasks" after deployment, i.e. ability to change during use (due to self-learning).
- Characterised by inferential capacity, i.e. the "capability to derive models or algorithms, or both, from inputs or data", to generate from the input it receives, for implicit or explicit purposes*, outputs, content, predictions, recommendations or decisions capable of influencing physical or virtual environments. Inference is possible through the use of machine learning techniques and logic and knowledge-based approaches in the construction of the system.

*Explicit goals: encoded by the developer directly in the system;

*Implicit goals: underlying human-specified rules or embedded in training data and derived through learning processes (e.g. LLM).

WHY DID THE EU LEGISLATOR DRAFT SUCH DEFINITION OF AI SYSTEM?

- Simple, broad and flexible definition, aligned with the definition adopted at the OECD (see Explanatory Memorandum No. 8, March 2024) to ensure legal certainty and facilitate international convergence.

- Focus on functional characteristics of the system, not on technical specifications and development methodologies, to ensure flexibility to facilitate rapid technological developments. This does not include traditional software, simpler programming approaches, systems that automatically perform operations according to predefined human rules (i.e. static or deterministic ‘if-then’ programming, as opposed to dynamic-probabilistic programming).

- The Commission will develop guidelines on the application of the definition.

THE RISK-BASED APPROACH (RECITAL 26)

- Unacceptable risk: prohibited AI practices (Art. 5). Example: social scoring, biometric recognition, emotion assessment, behaviour prediction, trawling of people's images.

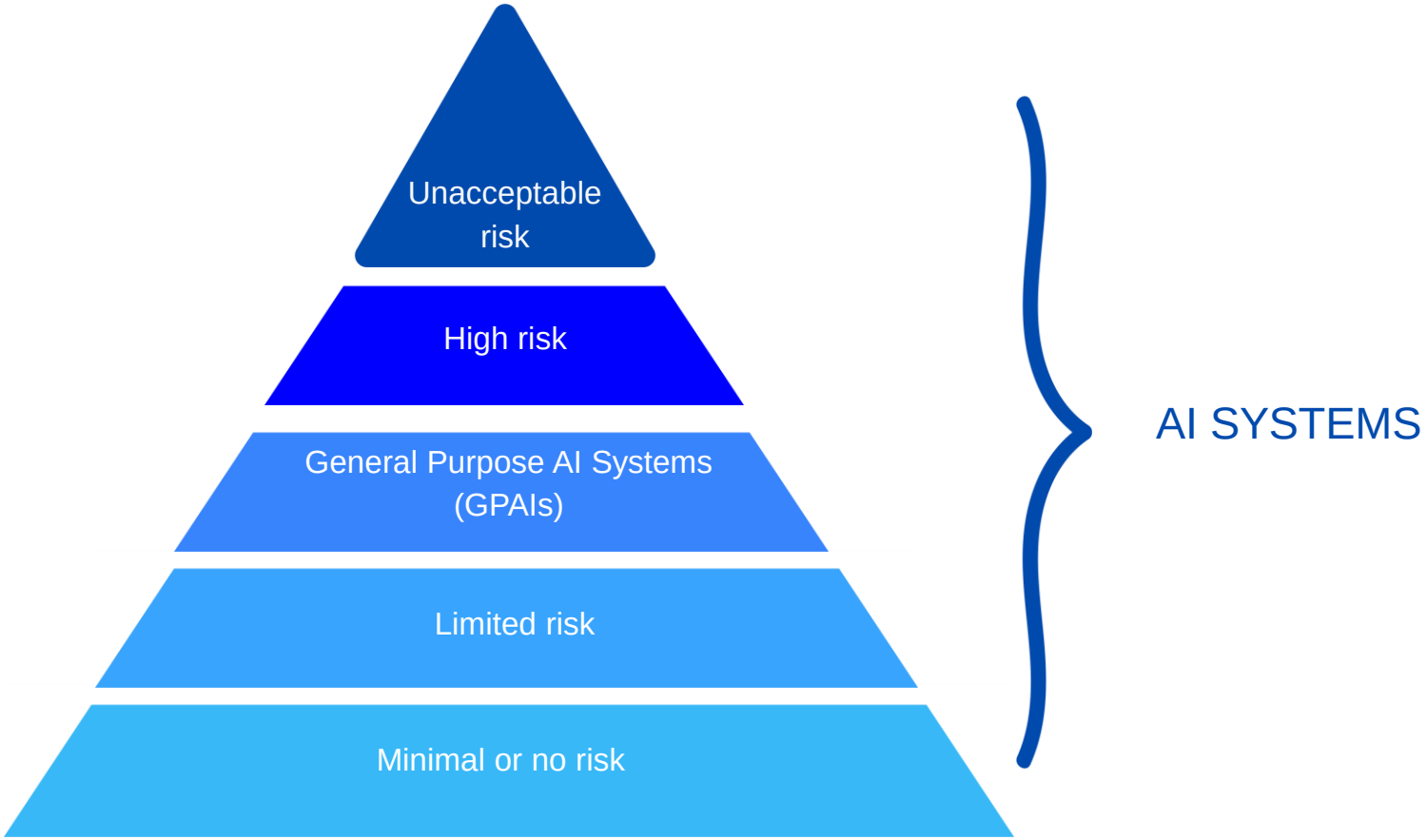
- High risk: compliance requirements, ex ante compliance assessment and obligations for operators (Art. 6-49). Example: AI systems used in medical devices, recruitment tools, human resources and workers management and critical infrastructure management.

- Specific risks related to deception or impersonation: transparency obligations for operators, possibly in addition to those for high-risk systems (Art. 50). Example: chatbots, deepfakes, AI-enabled video games, inventory management systems, market segmentation systems.

- Minimal or no risk: no specific obligation, but duty of literacy (Art. 4) and voluntary adherence to codes of conduct (Art. 95). The codes provide the same obligations for providers of general purpose AI models.

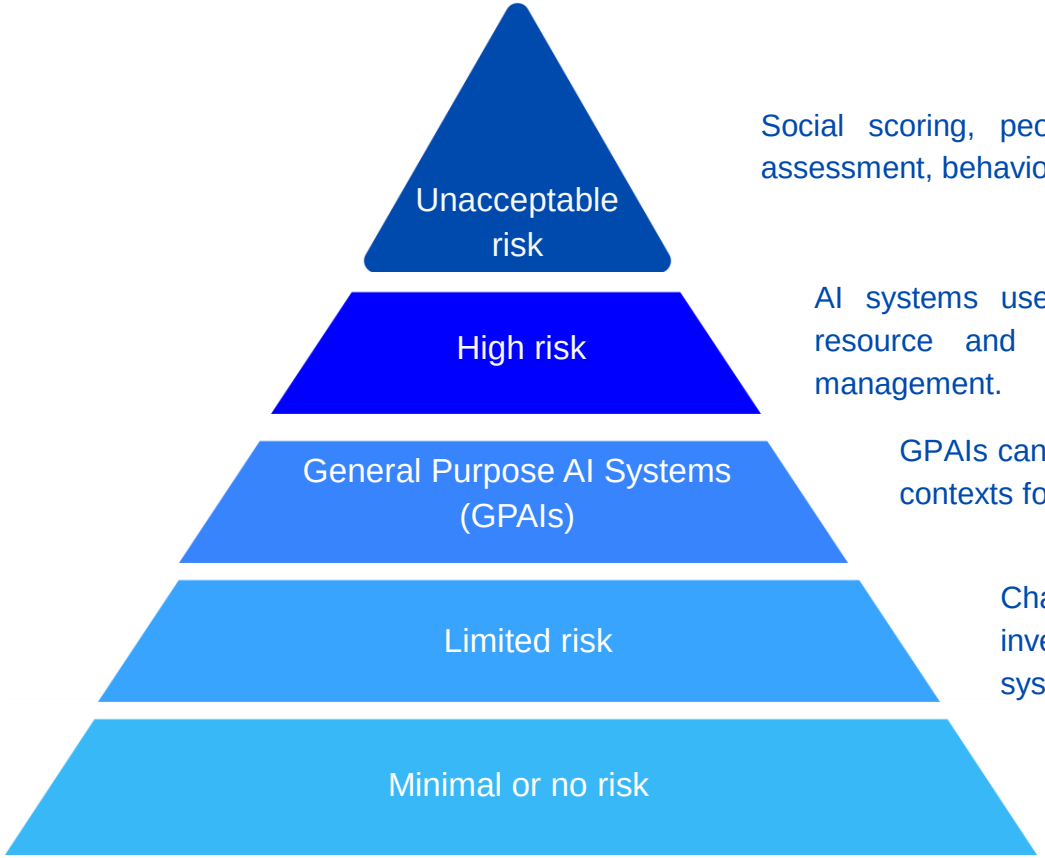
Co-funded by the European Union

THE EUROPEAN REGULATION ON ARTIFICIAL INTELLIGENCE: THE AI ACT



Co-funded by the European Union

THE EUROPEAN REGULATION ON ARTIFICIAL INTELLIGENCE: THE AI ACT



Social scoring, people manipulation, biometric recognition, emotion assessment, behaviour prediction, image trawling of people.

AI systems used in medical devices, recruitment tools, human resource and worker management and critical infrastructure management.

GPAIs can be used or adapted to a wide number of application contexts for which they were not specifically designed.

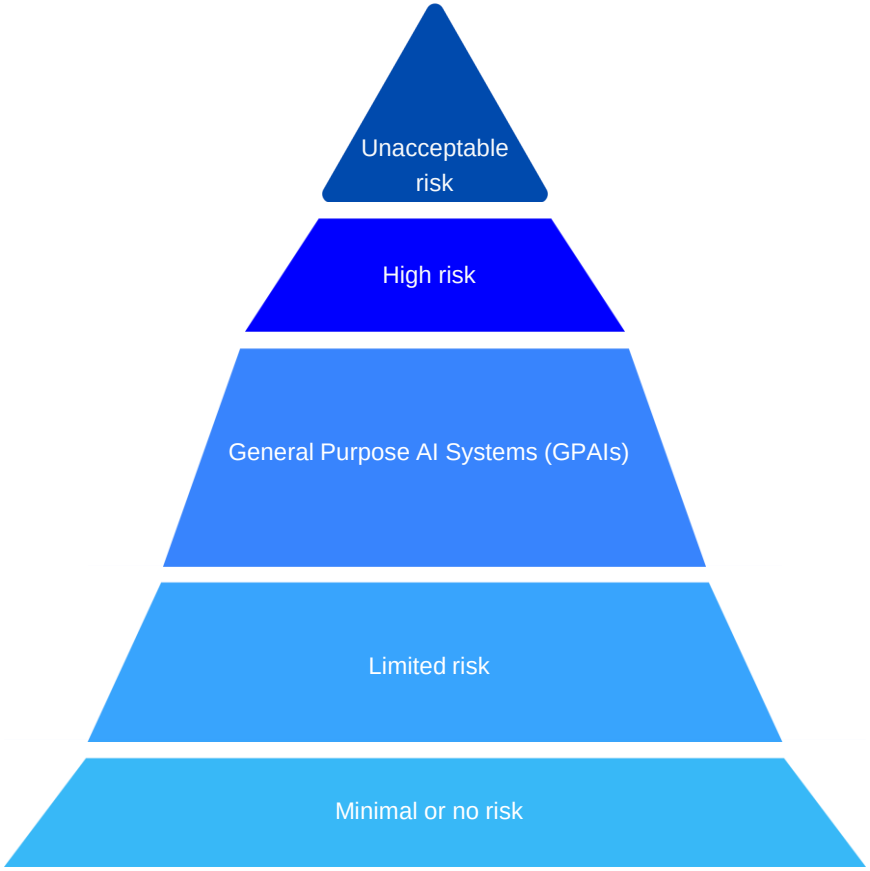
Chatbots, deepfakes, AI-enabled video games, inventory management systems, market segmentation systems.

All AI systems not included in the other categories.

Co-funded by the European Union

Università Europea di Roma

THE EUROPEAN REGULATION ON ARTIFICIAL INTELLIGENCE: THE AI ACT



An AI system that poses an unacceptable risk because it violates the fundamental rights of end users is prohibited throughout the EU.

The provider of a high-risk AI system must comply with requirements (Arts. 6 - 49), including subjecting the system to a conformity assessment before placing it on the market.

The provider of GPAIs is required to comply with transparency requirements (Art. 53), including the disclosure of certain information to downstream system providers. Additional obligations exist for GPAI systems that pose 'systemic risks', including GPAIs trained using computing power exceeding 10^25 FLOPs, such as GPT-4.

The provider of a low-risk AI system, including AI systems for general low-impact purposes (such as chatbots and deepfakes), must comply with specific transparency obligations (Art. 50), which include, for example, ensuring that users are aware that they are interacting with an AI.

Providers of AI systems that present a low or minimal risk to the security and fundamental human rights of end users are encouraged to voluntarily comply with mandatory requirements for high-risk AI systems through voluntary codes of practice.

AI MODELS FOR GENERAL PURPOSES OR GPAI (ART. 3(63); RECITALS 97-99)

Usually trained on large amounts of data by various methods, such as self-supervised, unsupervised or reinforcement-based learning, it is characterised by:

- significant generality;
- ability to competently perform a wide range of distinct tasks;
- suitability to be integrated into a variety of downstream systems or applications.

GPAI models are certainly those with at least one billion parameters and trained by means of large-scale self-supervision (recital 98), especially large generative AI models (recital 99).

The regulation applies to GPAI models once they have been placed on the market (regardless of the mode), not to those used before they are placed on the market for research, development and prototyping purposes only.

Providers and deployers of AI systems with limited risk, including general purpose AI systems with low impact, must comply with a number of transparency obligations regulated in Art. 50.

SYSTEMIC GPAI MODELS (ART. 3(65), ART. 51; RECITAL 110)

Due to their high impact capacity, they may pose a systemic risk that significantly affects the EU market due to their scale and with actual or reasonably foreseeable negative effects on public health, security, fundamental rights or society as a whole, which may propagate along the entire value chain.

According to Art. 51 and recitals 111-113, systemic GPAIs are classified as those that:

- have high impact capabilities assessed on the basis of appropriate technical tools and methodologies (notification procedure) or;
- are designated as such by an individual decision of the Commission, based on the criteria set out in an annex to the AI Act.

High impact capacity presumed if FLOP greater than 10^25. This threshold will be reviewed by the Commission in the light of technological developments.

GPAI SYSTEMS (ART. 3, PARA. 66; RECITAL 100)

- based on a GPAI model;
- because of this integration, it has the capacity to serve various purposes, either for direct use or for integration into other AI systems.

Recital 85

"General-purpose AI systems may be used as high-risk AI systems by themselves or be components of other high-risk AI systems. Therefore, due to their particular nature and in order to ensure a fair sharing of responsibilities along the AI value chain, the providers of such systems should, irrespective of whether they may be used as high-risk AI systems as such by other providers or as components of high-risk AI systems and unless provided otherwise under this Regulation, closely cooperate with the providers of the relevant high-risk AI systems to enable their compliance with the relevant obligations under this Regulation and with the competent authorities established under this Regulation."

TRANSPARENCY OBLIGATIONS FOR GPAI SYSTEMS (ART. 53; RECITAL 101)

In addition to the transparency obligations set out in Art. 50, providers of general purpose AI systems, general purpose AI models and generative AI must comply with a number of result obligations set out in Art. 53.

TRANSPARENCY OBLIGATIONS GPAI SYSTEMS, GPAI MODELS AND GENERATIVE AI (ART. 53; RECITAL 101)

In addition to the transparency obligations set out in Art. 50, providers of general purpose AI systems, general purpose AI models and generative AI must comply with a number of result obligations set out in Art. 53.

Recital 101
“Providers of general-purpose AI models have a particular role and responsibility along the AI value chain, as the models they provide may form the basis for a range of downstream systems, often provided by downstream providers that necessitate a good understanding of the models and their capabilities, both to enable the integration of such models into their products, and to fulfil their obligations under this or other regulations. Therefore, proportionate transparency measures should be laid down, including the drawing up and keeping up to date of documentation, and the provision of information on the general-purpose AI model for its usage by the downstream providers. Technical documentation should be prepared and kept up to date by the general-purpose AI model provider for the purpose of making it available, upon request, to the AI Office and the national competent authorities. The minimal set of elements to be included in such documentation should be set out in specific annexes to this Regulation. The Commission should be empowered to amend those annexes by means of delegated acts in light of evolving technological developments.”

PROHIBITIONS: PROTECTED VALUES, INCIDENCE OF RISK

- Freedom of choice, self-determination: subliminal, manipulative and vulnerability-exploiting techniques capable of significantly altering the decision-making capacity and distorting the behaviour of individuals or groups (with actual or potential serious harm).

- Non-discrimination: social scoring systems with disproportionate prejudicial effect (and/or based on data acquired in other contexts); biometric categorisation to infer or deduce ‘sensitive’ characteristics of individuals; recognition of emotions in the context of work or education (characterised by power imbalance).

- Rule of law and presumption of innocence: predictive systems of criminal risk based on profiling of individuals.

- Privacy, personal data protection: image scraping to create facial recognition databases (increasing the sense of mass surveillance); real-time remote biometric identification in publicly accessible spaces for law enforcement purposes.

It is not necessary for the provider or deployer to have the intent to cause significant harm, as long as the harm results from the manipulation/exploitation made possible by the AI.

THE PROHIBITIONS IN DETAIL (ART. 5; RECITALS 29, 30, 31, 42, 43, 44)

It is prohibited to place on the market, putting into service or use:

(a) an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm;

(b) an AI system that exploits any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective, or the effect, of materially distorting the behaviour of that person or a person belonging to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm;

THE PROHIBITIONS IN DETAIL (ART. 5; RECITALS 29, 30, 31, 42, 43, 44)

It is prohibited to place on the market, putting into service or use:

(c) AI systems for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, with the social score leading to either or both of the following:

(i) detrimental or unfavourable treatment of certain natural persons or groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected;

(ii) detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity.

THE PROHIBITIONS IN DETAIL (ART. 5; RECITALS 29, 30, 31, 42, 43, 44)

It is prohibited to place on the market, putting into service or use:

(d) an AI system for making risk assessments of natural persons in order to assess or predict the risk of a natural person committing a criminal offence, based solely on the profiling of a natural person or on assessing their personality traits and characteristics; this prohibition shall not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity;

(Tools for analysing the risks of financial fraud by companies on the basis of suspicious transactions or aimed at locating narcotic drugs or illicit goods by customs authorities are not affected by the ban).

(e) AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage;

.

THE PROHIBITIONS IN DETAIL (ART. 5; RECITALS 29, 30, 31, 42, 43, 44)

It is prohibited to place on the market, putting into service or use:

(f) AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons;

(g) biometric categorisation systems that categorise individually natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation; this prohibition does not cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data or categorizing of biometric data in the area of law enforcement;

(The labelling/filtering, on the basis of biometric data, of legally acquired datasets and the categorisation of biometric data in the field of law enforcement are excluded from the prohibition).

THE PROHIBITIONS IN DETAIL (ART. 5; RECITALS 29, 30, 31, 42, 43, 44)

It is prohibited to place on the market, putting into service or use:

(h) 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement, unless and in so far as such use is strictly necessary for one of the following objectives:

(i) the targeted search for specific victims of abduction, trafficking in human beings or sexual exploitation of human beings, as well as the search for missing persons;

(ii) the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or a genuine and present or genuine and foreseeable threat of a terrorist attack;

(iii) the localisation or identification of a person suspected of having committed a criminal offence, for the purpose of conducting a criminal investigation or prosecution or executing a criminal penalty for offences referred to in Annex II and punishable in the Member State concerned by a custodial sentence or a detention order for a maximum period of at least four years.

Point (h) of the first subparagraph is without prejudice to Article 9 of Regulation (EU) 2016/679 for the processing of biometric data for purposes other than law enforcement.

‘REAL-TIME’ REMOTE BIOMETRIC IDENTIFICATION SYSTEMS IN DETAIL (ART. 5, LETTER H; RECITALS 33-34)

Real-time remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement are prohibited except as necessary to search for victims of certain crimes or missing persons, prevent imminent threats to life or limb or terrorist attacks, locate or identify suspected perpetrators of specific serious crimes, and provided that:

- the use is intended only to confirm the identity of a specific person;
- the conditions and safeguards provided for by national law are respected;
- the law enforcement authority has carried out a fundamental rights impact assessment and registered the system in the EU database;
- the use is authorised in advance by a court or an independent administrative authority (except in cases of urgency), expressly provided for by national rules, notified to the market surveillance authority and the data protection authority.

CLASSIFICATION RULES FOR HIGH-RISK AI SYSTEMS (ART. 6(1-2); RECITALS 46-52)

- Systems intended to be used as “a safety component of a product, or the AI system is itself a product” subject to harmonised EU standards (including machinery, toys, lifts, radio equipment, medical and safety devices, motor vehicles, unmanned aircraft) and subject to related ex ante conformity assessment by third parties.

- Safety component that performs a safety function for the product or whose failure or malfunction endangers the health and safety of persons or property.

- “Stand-alone” systems identified in Annex III with reference to specific sectors: biometrics; critical infrastructure; education and vocational training; employment, management of workers and access to self-employment; access to and use of essential private services and public services; law enforcement; migration, asylum and border control management; administration of justice and democratic processes.

EXCEPTIONS (ART. 6(3); RECITAL 53)

AI systems listed in Annex III that do not pose a significant risk of harm to health, safety or fundamental rights of natural persons are not considered high-risk (unless they involve profiling) because they are intended to:

- perform only a “narrow procedural task” (e.g. categorisation of documents);
- “improve the result of a previously completed human activity” (e.g. improve the language of already drafted documents);
- “detect decision-making patterns or deviations from prior decision-making patterns and is not meant to replace or influence the previously completed human assessment, without proper human review”;
- “perform a preparatory task for an assessment relevant for the purposes of the use cases listed in Annex III” (e.g. intelligent file management solutions, translation of documents).

The provider shall carry out and document the assessment prior to placing on the market/commissioning and provide documentation to the competent authorities upon request.

SOME HIGH-RISK STAND-ALONE AI SYSTEMS

Biometrics (provided use is permitted under EU or national law)

- remote biometric identification systems, if not prohibited under Art. 5. Not high-risk those used for biometric verification or authentication (e.g. allowing access to a location or unlocking a device);
- systems for biometric categorisation based on sensitive data, if not prohibited under Art. 5 (i.e. not intended to infer or deduce race, political opinions, etc.);
- systems for emotion recognition, if not prohibited under Art. 5 (i.e. used in contexts other than work and education).

Critical infrastructure

- Systems operating as security components in the management and operation of critical digital infrastructure, road traffic, water/gas/heating/electricity supply (e.g., water pressure monitoring or fire control in cloud computing centres) Security components are not those used for cybersecurity purposes only.

SOME HIGH-RISK STAND-ALONE AI SYSTEMS

Jobs

- systems for recruiting or selecting individuals, in particular for publishing targeted job advertisements, analysing or filtering applications and evaluating candidates;

- systems for making decisions concerning the conditions of employment relationships, the promotion or termination of employment relationships, for assigning tasks on the basis of individual behaviour or personal traits and characteristics, or for monitoring and evaluating people's performance and behaviour in the context of such employment relationships.

They can have a significant impact on the future of individuals in terms of career and livelihood prospects and workers' rights, perpetuate historical patterns of discrimination, and undermine fundamental rights to data protection and privacy.

SOME HIGH-RISK STAND-ALONE AI SYSTEMS

Essential public and private services and benefits

- systems for assessing, by or on behalf of public authorities, the eligibility of natural persons for essential public assistance benefits and services and for granting, reducing, withdrawing or recovering such benefits and services;
- systems to assess the creditworthiness of natural persons or to establish their credit score (excluding systems used to detect financial fraud and for prudential purposes to calculate the capital requirements of banks and insurance companies);
- systems to assess risks and determine prices in relation to natural persons in the case of life and health insurance;
- systems for assessing and classifying emergency calls made by natural persons, dispatching or prioritising emergency first aid services or triaging patients in emergency health care.

SOME HIGH-RISK STAND-ALONE AI SYSTEMS

Risk Management (Art. 9, para. 65)

Establishment, implementation, documentation and maintenance throughout the life cycle of the system, with constant and systematic updating, of a risk management system that includes:

- identification and analysis of risks a) known and reasonably foreseeable arising from use in accordance with the intended purpose, b) that may arise from reasonably foreseeable misuse (human behaviour, recital 65) of the system, c) that emerge from post-market monitoring (also based on data provided by the deployer).

- adoption of appropriate and targeted risk management measures, such as to ensure, as appropriate, the elimination, reduction, mitigation or control of risks (if not eliminable, they must become ‘acceptable’) and to ensure that the deployer has the necessary information/instructions for use and training to understand the operation of the system.

MANDATORY REQUIREMENTS FOR HIGH-RISK AI SYSTEMS

Data and data governance (Art. 10, recitals 67-70)

- "Training, validation and testing data sets shall be subject to data governance and management practices appropriate for the intended purpose of the high-risk AI system". This covers in particular: design choices, data collection processes, data preparation operations, assessment of the adequacy of available datasets, evaluation of possible biases and measures to mitigate them), to ensure the high quality of training, validation and testing datasets.

- "Training, validation and testing data sets shall be relevant, sufficiently representative, and to the best extent possible, free of errors and complete in view of the intended purpose".

If personal data are involved, minimisation, privacy by design and privacy by default must be ensured, in particular by anonymisation and encryption techniques (recital 69). Exceptionally, if strictly necessary to detect and correct bias, the processing of special categories of personal data is allowed, with stringent security measures.

MANDATORY REQUIREMENTS FOR HIGH-RISK SYSTEMS

Technical documentation and record-keeping (Artt. 11-12; recital 71)

- Preparation (prior to placing on the market or putting into service) and updating of clear and comprehensible technical documentation necessary to demonstrate the conformity of the system with the requirements, to be made available to competent authorities and notified bodies. This implies a high level of competence within companies.

- SMEs, including start-ups, can provide in a simplified manner the elements of the technical documentation specified in Annex IV.

- Minimum content in Annex IV: general description of the system, detailed description of the development process (algorithms, data training, validation and testing procedures, cybersecurity measures, etc.), information on monitoring, operation and control, description of the risk management system, etc. The Commission will develop a simplified technical documentation form for SMEs.

- Design to ensure at technical level the automatic logging of events (logs) for the entire life cycle of the system (and thus traceability of operation and use by the deployer).

MANDATORY REQUIREMENTS FOR HIGH-RISK AI SYSTEMS

Transparency and provision of information to deployers (Art. 13; recital 72)

- Design and development to ensure transparency of operation and to help deployers interpret the system output and use it properly;
- Provision of instructions “for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to deployers”.
- Information should include system characteristics, capabilities and performance limitations (including known or foreseeable circumstances that may entail risks, including the action of the deployer that may influence system behaviour and performance), planned human oversight measures, computational and hardware resources required for the proper functioning of the system. Where appropriate, include illustrative examples in the instructions, e.g. on limitations and intended and prohibited uses of the AI system.

MANDATORY REQUIREMENTS FOR HIGH-RISK AI SYSTEMS

Human oversight (Art. 14; recital 73)

- Design and development conducted so as to ensure human supervision during use/operation and to prevent or minimise risks.
- Ensure inherent operational constraints that the system cannot override and that the system is responsive to the human supervisor.
- Measures should be identified by the provider prior to marketing or commissioning and either integrated upstream into the system or deferred for implementation by the deployer.

MANDATORY REQUIREMENTS FOR HIGH-RISK AI SYSTEMS

Human oversight details (Art. 14, par. 4)

Supervisors must be able to: (a) properly understand the relevant capacities and limitations of the high-risk AI system and be able to duly monitor its operation, including in view of detecting and addressing anomalies, dysfunctions and unexpected performance; (b) to remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system (automation bias), in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons; (c) to correctly interpret the high-risk AI system's output, taking into account, for example, the interpretation tools and methods available; (d) to decide, in any particular situation, not to use the high-risk AI system or to otherwise disregard, override or reverse the output of the high-risk AI system; (e) to intervene in the operation of the high-risk AI system or interrupt the system through a 'stop' button or a similar procedure that allows the system to come to a halt in a safe state.

MANDATORY REQUIREMENTS FOR HIGH-RISK AI SYSTEMS

Human oversight (Art. 14; recital 73)

For high-risk AI systems referred to in point 1(a) of Annex III, the measures referred to in paragraph 3 of this Article shall be such as to ensure that, in addition, no action or decision is taken by the deployer on the basis of the identification resulting from the system unless that identification has been separately verified and confirmed by at least two natural persons with the necessary competence, training and authority. The requirement for a separate verification by at least two natural persons shall not apply to high-risk AI systems used for the purposes of law enforcement, migration, border control or asylum, where Union or national law considers the application of this requirement to be disproportionate.

MANDATORY REQUIREMENTS FOR HIGH-RISK AI SYSTEMS

Accuracy, robustness and cybersecurity (Art. 15; recitals 74-78)

- Design and development conducted so as to achieve an adequate level of accuracy, robustness (resilience against errors, failures, inconsistencies) and cybersecurity (resilience to malicious attacks by unauthorised third parties). The Commission will promote the development of benchmarks and measurement methodologies. The instructions for use will specify accuracy levels and metrics.

- Adoption of technical and organisational measures and technical redundancy solutions (back-up or fail-safe plans). For continuously learning systems, measures to avoid feedback loops.

- Cybersecurity solutions include measures to prevent and control data poisoning or model poisoning, confidentiality attacks, etc. A system that complies with the essential requirements of the EU cybersecurity regulation is considered adequate from a cybersecurity perspective.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Provider (Art. 3 (3): a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge.

The following are subject to the AI Act (Art. 2, para. 1 a-c):

- providers placing on the market or putting into service AI systems or placing on the market general-purpose AI models in the Union, irrespective of whether those providers are established or located within the Union or in a third country;
- providers and deployers of AI systems that have their place of establishment or are located in a third country, where the output produced by the AI system is used in the Union;

VALUE CHAIN AND OPERATORS OBLIGATIONS

Obligations of providers of high-risk AI systems (Arts. 16-22; recital 81)

(a) ensure that their high-risk AI systems are compliant with the requirements set out in Section 2;
(b) indicate on the high-risk AI system or, where that is not possible, on its packaging or its accompanying documentation, as applicable, their name, registered trade name or registered trade mark, the address at which they can be contacted;
(c) have a quality management system in place which complies with Article 17;
(d) keep the documentation referred to in Article 18;
(e) when under their control, keep the logs automatically generated by their high-risk AI systems as referred to in Article 19;
(f) ensure that the high-risk AI system undergoes the relevant conformity assessment procedure as referred to in Article 43, prior to its being placed on the market or put into service;
(g) draw up an EU declaration of conformity in accordance with Article 47;
(h) affix the CE marking to the high-risk AI system or, where that is not possible, on its packaging or its accompanying documentation, to indicate conformity with this Regulation, in accordance with Article 48;
(i) comply with the registration obligations referred to in Article 49(1);
(j) take the necessary corrective actions and provide information as required in Article 20;
(k) upon a reasoned request of a national competent authority, demonstrate the conformity of the high-risk AI system with the requirements set out in Section 2;
(l) ensure that the high-risk AI system complies with accessibility requirements in accordance with Directives (EU) 2016/2102 and (EU) 2019/882.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Main obligations on provider of high-risk AI systems (Artt. 16-22; recital 81)

- keep - for 10 years after placing on the market or putting into service - and make available to the authorities all technical documentation relating to conformity with requirements, the quality management system and the EU declaration of conformity, as well as any documents issued by notified bodies
- keep (for a period appropriate to the purpose of the system, not less than six months) the logs automatically generated by the system, if under their control, and give access to them to the national authority upon request;
- ensure that before the system is placed on the market/commissioned, it undergoes a conformity assessment procedure (Art. 43), based on internal control by the provider or the involvement of notified bodies (see below);
- draw up an EU declaration of conformity (Art. 47), attesting the fulfilment of the mandatory requirements and by which the provider assumes responsibility for the conformity of the system.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Main obligations on provider of high-risk AI systems (Artt. 16-22; recital 81)

- affix the CE marking on the system (or packaging/accompanying documents), which allows free circulation in the internal market (Art. 48)
- register the system in the EU database of systems aiAnnex III;
- take the necessary corrective measures immediately (and inform distributors, importers and deployers) if they consider that the system is not in conformity, investigate the causes and inform the authorities;
- demonstrate the conformity of the system upon reasoned request by a national authority and cooperate by providing information and documentation;
- ensure that the system complies with the accessibility requirements of EU regulations for the protection of persons with disabilities;
- if established in third countries, appoint an authorised representative and specify their tasks in a written mandate.

POST-MARKETING MONITORING

Post-market monitoring by providers and post-market monitoring plan for high-risk AI system (Art. 72)

- Providers shall establish and document a post-market monitoring system in a manner that is proportionate to the nature of the AI technologies and the risks of the high-risk AI system.
- The post-market monitoring system shall actively and systematically collect, document and analyse relevant data which may be provided by deployers or which may be collected through other sources on the performance of high-risk AI systems throughout their lifetime, and which allow the provider to evaluate the continuous compliance of AI systems with the requirements set out in Chapter III, Section 2. Where relevant, post-market monitoring shall include an analysis of the interaction with other AI systems. This obligation shall not cover sensitive operational data of deployers which are law-enforcement authorities.
- The post-market monitoring system shall be based on a post-market monitoring plan. The post-market monitoring plan shall be part of the technical documentation referred to in Annex IV. The Commission shall adopt an implementing act laying down detailed provisions establishing a template for the post-market monitoring plan and the list of elements to be included in the plan by 2 February 2026. That implementing act shall be adopted in accordance with the examination procedure referred to in Article 98(2).

POST-MARKETING MONITORING

Post-market monitoring by providers and post-market monitoring plan for high-risk AI system (Art. 72)

- For high-risk AI systems covered by the Union harmonisation legislation listed in Section A of Annex I, where a post-market monitoring system and plan are already established under that legislation, in order to ensure consistency, avoid duplications and minimise additional burdens, providers shall have a choice of integrating, as appropriate, the necessary elements described in paragraphs 1, 2 and 3 using the template referred in paragraph 3 into systems and plans already existing under that legislation, provided that it achieves an equivalent level of protection.

The first subparagraph of this paragraph shall also apply to high-risk AI systems referred to in point 5 of Annex III placed on the market or put into service by financial institutions that are subject to requirements under Union financial services law regarding their internal governance, arrangements or processes

VALUE CHAIN AND OBLIGATIONS OF DISTRIBUTORS, IMPORTERS, DEPLOYERS OR OTHER THIRD-PARTIES

Art. 25(3):

For high-risk AI systems as safety components of products subject to EU 'New Approach' harmonisation rules, the product manufacturer shall be considered to be the provider of the high-risk system and shall be subject to the obligations under Article 16 under either of the following circumstances:

- the high-risk AI system is placed on the market together with the product under the product manufacturer's name or trademark;
- the high-risk AI system is put into service under the product manufacturer's name or trademark after the product has been placed on the market.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Conformity assessment and European standards (Art. 40-49)

Compliance of a high-risk AI system with mandatory requirements is presumed if the provider applies harmonised standards established by European standardisation organisations or, in the absence of such standards and until their adoption, common specifications established by the Commission

- For products subject to EU ‘new approach’ standards: relevant conformity assessment procedure;
- For Annex III AI systems (except those used for biometrics): internal control (the Commission may, by means of delegated acts, impose the use of notified bodies).

Notified bodies issue certificates valid for 4 years (5 for products). The provider completes an EU declaration of conformity attesting that the mandatory requirements have been met and by which he assumes responsibility for conformity.

For exceptional reasons of protection of important public interests (security, health, etc.), or in the case of specific, substantial and imminent threat to the life or physical safety of natural persons, law-enforcement authorities or civil protection authorities may authorise the marketing of high-risk AI systems without a conformity assessment procedure (with Commission supervision).

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Importer (Art. 3(6)): A natural or legal person located or established in the Union that places on the market an AI system that bears the name or trademark of a natural or legal person established in a third country.

Obligations on importers of high-risk AI systems (Art. 23):

- before placing the system on the market, the importer must verify its conformity (1. conformity assessment procedure referred to in Art. 43; 2. technical documentation in accordance with Art. 11; 3. CE marking and EU declaration of conformity referred to in Art. 47; 4. appointment of an authorised representative).
- the importer must refrain from placing on the market systems deemed non-compliant/falsified or accompanied by falsified documentation; in the event of a risk, inform the provider and the supervisory authorities;
- the impoert must indicate their references on the packaging/accompanying document; ensure that transport/storage conditions do not jeopardise compliance;
- the importer must keep documentation; cooperate with competent authorities.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Distributor (Art. 3(7)): a 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.

Obligations on distributors of high-risk AI systems (Art. 24):

- before making the high-risk AI system available on the market, distributors must verify the presence of the required (1) CE marking, (2) copy of the EU declaration of conformity and (3) instructions for use;
- refrain from making a system available on the market that is considered non-compliant;
- inform the provider/importer of any risks;
- ensure that storage/transport conditions do not jeopardise the compliance of the system;
- if they consider that a system already made available on the market does not comply with the requirements, take the necessary corrective measures, or withdraw/recall the system; if the system presents a risk, inform the provider/importer and the authorities and cooperate with them.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Deployer (Art. 3(4); recital 13): a 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.

They are subject to the AI Act (Art. 2(1) b-c):

- deployers of AI systems that have their place of establishment or are located within the Union;
- providers and deployers of AI systems that have their place of establishment or are located in a third country, where the output produced by the AI system is used in the Union.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

In certain circumstances, the deployer (as well as the distributor, importer or other third party) is considered to be the provider of a high-risk AI system and assumes the corresponding obligations (Art. 25; recital 84):

(a) if it affixes its name or trademark to a high-risk AI system that has already been placed on the market or put into service (without prejudice to contractual agreements providing for a different division of obligations);

(b) if it makes a substantial change to a high-risk AI system already placed on the market or put into service so that it remains high-risk;

(c) if it changes the intended purpose of an AI system (including GPAIs) not classified as high risk so that it becomes high risk.

In such cases, the initial provider must cooperate closely with the new providers (information, reasonably expected technical access and any other assistance), unless it has clearly excluded the transformation of its system into a high-risk AI system.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Obligations on deployers of high-risk AI systems (Art. 26; recitals 91-95)

- take appropriate technical and organisational measures to ensure that high-risk AI systems are used in accordance with the instructions for use;
- entrust human supervision to natural persons who have the necessary competence, training and authority;
- ensure that input data (if under its control) are relevant and sufficiently representative in light of the intended purpose of the system;
- monitor the operation of the system in accordance with the instructions for use (transmitting the relevant information to the provider). If they consider that the use of the system may present a risk, they inform the provider/distributor and the supervisory authority without delay and suspend the use of the system. If they detect a serious incident, they inform the provider/importer/distributor and the supervisory authority.
- keep the logs automatically generated by the system for a period appropriate to the intended purpose of the system (not less than six months);
- if the deployer is an employer and the system is intended to be used in the workplace, inform workers' representatives and the workers concerned;
- for remote biometric identification systems to be used for the targeted search of a suspected or convicted offender, request prior judicial or administrative authorisation within 48 hours;
- for systems listed in Annex III that take decisions or assist in taking decisions concerning natural persons, inform them that they are subject to the use of the high-risk system;
- cooperate with the competent authorities.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Obligations for deployers of high-risk AI systems (Art. 27; recital 96)

- Deployers of Annex III systems (except those related to critical infrastructure security) that are public law bodies or private entities providing public services or entities that use credit scoring or risk assessment/pricing systems for life and health insurance: prior to first use of the system, they shall perform an assessment of the impact on fundamental rights that the use of such system may produce. Once the assessment has been performed, the deployer shall notify the market surveillance authority of its results, submitting the filled-out template referred to in paragraph 5 of Article 27 as part of the notification. In the case referred to in Article 46(1), deployers may be exempt from that obligation to notify.

The template includes (1) a description of the deployer's processes in which the system will be used according to its intended purpose; (2) period of time/frequency of use; (3) categories of natural persons and groups affected by the use and their specific risks of harm; (4) specific risks of harm likely to have an impact on the categories of natural persons or groups of persons; (5) human oversight measures implemented; (6) measures to be taken if risks materialise, including internal governance arrangements and grievance mechanisms.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Obligations on deployers of high-risk AI systems (Art. 86; recital 171)

Every data subject who is the subject of a decision taken by the deployer on the basis of the output of a high-risk AI system referred to in Annex III and which produces legal effects or similarly significantly affects him/her in a way that he/she considers to have an adverse impact on his/her health/safety/basic rights shall have the right to obtain clear and meaningful explanations from the deployer on the role of the AI system in the decision-making process and on the main elements of the decision taken.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Obligations on third parties providing elements of a high-risk AI system (Art. 25.4; recitals 88-90)

- Third parties that provide AI systems, tools, services, components or processes used or integrated into a high-risk AI system are required to provide the provider of the high-risk system, by written agreement, with the information, capabilities, technical access and any other assistance necessary to enable the provider to fully perform its obligations.

- Third parties who make tools, services, processes or components, other than GPAIs, publicly available under a free and open source licence are excluded.

- Voluntary standard contractual clauses will be developed by the AI Office, which will take into account the possible contractual requirements applicable in certain sectors and business cases.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Providers of certain AI systems (Art. 50(1-2); recitals 132-133) regardless of whether they are considered high-risk or not

- Providers of AI systems intended to interact directly with natural persons must design and develop the system in such a way that persons are informed (notified) that they are interacting with an AI system, unless this would not be apparent to a reasonably informed, observant and circumspect person, taking into account the circumstances and context of use (e.g., interaction with persons vulnerable by age or disability).

- Providers of AI systems, including GPAIs, that generate audio, image, video or synthetic text content, must be marked in a machine-readable format (watermarks, cryptographic methods, etc.) and detectable as artificially generated or manipulated.

- Exception for AI systems with a standard editing assistance function or which do not substantially alter the input data provided by the deployer or the respective semantics.

VALUE CHAIN AND OPERATORS' OBLIGATIONS

Deployer of certain AI systems (Art. 50(3)-(4); recital 132-134) regardless of whether they are considered high-risk or not

- Deployers of emotion-recognition or biometric categorisation systems must inform exposed natural persons about the operation of the system and complying with data protection regulations.

- Deployers of AI systems that generate or manipulate images or audio or video content that constitutes a deep fake must disclose that the content has been artificially generated or manipulated (if such content is part of a manifestly artistic or creative work or programme, etc., disclose the existence of the generated/manipulated content without hindering the exhibition or enjoyment of the work).

- Deployers of AI systems that generate/manipulate published text for the purpose of informing the public about matters of public interest must disclose that the text has been artificially generated/manipulated (exception for content subject to human review or control by an editorial manager).

OBLIGATIONS FOR GPAI MODELS PROVIDERS (ART. 53-56)

- draw up and keep up-to-date technical documentation of the model, including the training and testing process and the results of its evaluation (see minimum elements in Annex XI, including known or estimated energy consumption), to be forwarded on request to the AI Office and the competent national authorities.

- prepare and make available information and documentation to downstream providers that intend to integrate the GPAI model into their AI system, enabling them to have a good understanding of the capabilities and limitations of the GPAI model and to fulfil their obligations (see minimum elements in Annex XII).

- implement a policy of compliance with the EU copyright rules (including conditions of operation of the ‘text and data mining’ exception).

- draft and publish a detailed summary of training content.

- if established in third countries, appoint an authorised representative.

OBLIGATIONS ON GPAIS MODEL PROVIDERS (ARTS. 53-56)

- Providers of GPAI models released under a free and open source licence (which can be freely accessed, used, modified and distributed) are exempted from the technical documentation/information requirements to downstream providers provided that the relevant parameters, including weights, information on model architecture and information on model use, are made public (the exception does not apply to GPAI models with systemic risk).

- Providers of GPAI models with systemic risk are subject to additional obligations: (1) perform an assessment of the models in accordance with standardised protocols and tools; to assess and mitigate possible systemic risks; (2) assess and mitigate possible systemic risks at Union level, including their sources, that may stem from the development, the placing on the market, or the use of GPAI models with systemic risk; (3) track document and report serious incidents and possible corrective measures to the AI Office and relevant national authorities; (4) ensure an adequate level of cybersecurity protection.

- Codes of good practice are envisaged (driving and monitoring role of the AI Office), to which the Commission may give general validity; failing this, the Commission will define common standards.

GOVERNANCE AND ENFORCEMENT: EU LEVEL (Artt. 56, 64, 75, 95; recital 116, 148, 161, 162, 164)

- The AI Office, within the administrative structure of DG CNECT (Decision C(2024) 390), works to support the Commission in the implementation of the AI Act, with specific tasks related mainly to GPAIs, including the development of tools, methodologies and benchmarks for assessing the capacity of GPAI models, in particular those with systemic risks, monitoring their functioning and the emergence of unforeseen risks, and conducting investigations into possible breaches of the rules.

- The AI Office also assists the Commission in the preparation of decisions, executive and delegated acts, guidelines, requests for standardisation and definition of common specifications, coordinates the establishment of the governance system for the application of the regulation, and promotes the adoption of codes of conduct at EU level (GPAIs, marking obligations for artificially generated or manipulated content). In implementing its tasks, the AI Office is called upon to ensure cooperation with stakeholders, through consultations and ad hoc fora.

GOVERNANCE AND ENFORCEMENT: EU LEVEL

- The European Artificial Intelligence Board (Art. 65 and 66), composed of one representative per Member State, provides advice and assistance to the Commission and the Member States to facilitate the consistent and effective implementation of the regulation (collection and sharing of best technical and regulatory practices, contribution to the harmonisation of administrative practices, recommendations and opinions on relevant issues, including evolving trends in AI value chains, support for Commission initiatives on literacy, etc.).

- An Advisory Forum (Art. 67), composed of stakeholder representatives, provides advice and technical expertise to the European Artificial Intelligence Board and the Commission.

- A Scientific panel of independent experts (Art. 68) selected by the Commission provides advice and support to the Office of AI for the implementation of the regulation, in particular with regard to the supervision of GPAI systems and models and cross-border investigative activities (if serious risks in two or more Member States).

GOVERNANCE AND ENFORCEMENT: NATIONAL LEVEL

Member States shall designate at least one notifying authority and one market surveillance authority (Art. 70) which:

- exercise their powers independently, impartially and without bias;
- have adequate technical, financial and human resources (sufficient staff to ensure in-depth understanding of AI technologies, data and computing, personal data protection, cybersecurity, fundamental rights, health and safety risks, and knowledge of existing standards and legal requirements), as well as the infrastructure needed to perform their tasks effectively;
- may provide advice and guidance on the implementation of the regulation, in particular to SMEs including start-ups; when ruling on AI systems in areas covered by EU regulations, they consult the relevant sectoral authorities at national level.

A market surveillance authority is designated as single point of contact.

GOVERNANCE AND ENFORCEMENT: NATIONAL LEVEL

The notifying authority is responsible for the notification procedures and subsequent monitoring of conformity assessment bodies for high-risk AI systems (Art. 28-39).

- once it receives the application from the body concerned, it verifies the requirements laid down in the regulation, relating to the independence of the body (as well as any subcontractors or subsidiaries) from providers of the systems subject to conformity assessment and their competitors, and to internal organisation and management measures, which must guarantee the impartiality of assessment activities and the protection of confidentiality of information.

- notifies the Commission and the other Member States (which may raise objections within a given period of time) of the bodies deemed to fulfil the requirements, which are placed on a public list.

- limits, suspends or withdraws the designation of a notified body that no longer meets the requirements or fails to fulfil its obligations (of information on certificates issued and subsequent events).

GOVERNANCE AND ENFORCEMENT: NATIONAL LEVEL

For specific areas, the market surveillance authority's choice is constrained:

- for high-risk AI systems linked to products subject to the Union harmonisation legislation, it is the one designated under the relevant legislation (Art. 74(3));
- for high-risk AI systems directly linked to the provision of financial services regulated by EU law, it is the one responsible for the financial supervision of the institutions that market/service/use the AI system (Art. 74(6));
- for some of the high-risk AI systems listed in Annex III (biometrics; law enforcement; migration/asylum/border control; administration of justice and democratic processes), Member States designate as market surveillance authorities the competent data protection authorities under the GDPR or Dir. (EU) 2016/680 (Art. 70(8));

Outside of these areas, Member States enjoy autonomy in their choice.

GOVERNANCE AND ENFORCEMENT: NATIONAL LEVEL

Market surveillance authorities operate according to the procedures and powers governed by Regulation (EU) 2019/1020 on market surveillance and conformity of products (Art. 74(1)). In order to perform their tasks, they have full access to the training, validation and test documentation and datasets used for the development of high-risk AI systems and, upon reasoned request, under certain conditions can access the source code (Art. 74(12)-(13)).

In the event of serious incidents, upon receipt of a report from the provider of the high-risk AI system, the market surveillance authorities inform the national authorities protecting fundamental rights, take appropriate measures (withdrawal, recall) and follows the notification procedures laid down in Reg. (EU) 2019/1010: Rapex rapid information system to the Commission (Art. 73).

Market surveillance authorities authorise and monitor the conduct of tests of AI systems under real conditions (both inside and outside sandboxes) and take any measures to modify, suspend or terminate the tests (Art. 60, 76).

Market surveillance authorities receive complaints about alleged breaches of the rules (Art. 85); whistleblowers benefit from whistleblower protection (Art. 87).

GOVERNANCE AND ENFORCEMENT: NATIONAL LEVEL

The market surveillance authority, in its market control/monitoring activities, may find that (Art. 79-83):

- an AI system poses a risk the market surveillance authority carries out an assessment of compliance of the AI system with the requirements/obligations of the regulation.

In the event of non-compliance, the market surveillance authority asks the operator concerned to take appropriate corrective action/withdraw/revocate the system from the market. If the operator fails to comply, the market surveillance authority takes provisional restrictive measures and notifies them to the Commission and the other Member States for possible objections. If no objections are raised, the measure is deemed justified and similar restrictive measures are taken in all Member States concerned.

If the market surveillance authority considers that, although compliant with the Regulation, a high-risk AI system nevertheless presents a risk, it requires the operator concerned to take appropriate measures to eliminate it and informs the Commission and the other Member States accordingly. The Commission decides whether the measure is justified and proposes any other appropriate measures.

GOVERNANCE AND ENFORCEMENT: NATIONAL LEVEL

- An AI system classified by the provider as not high risk under Art. 6(3) is in fact high riskà at the outcome of the assessment, the market surveillance authority shall require the provider to bring the system into compliance with the requirements and obligations of the Regulation, as well as to take appropriate corrective measures, and shall inform the Commission and the other Member States. The non-compliant provider shall be subject to financial penalties. The market surveillance authority adopts provisional restrictive measures, which may be objected to by the Commission or the other Member States; in the absence of objections, the measures are deemed justified and similar restrictive measures are adopted in the other Member States concerned.

- Safeguard procedure: if the Commission or other market surveillance authorities object to restrictive measures taken at national level (within 3 months/30 days for non-compliance with Article 5 prohibitions), the Commission decides whether the measure is justified. If so, all states take restrictive measures; if not, the measure must be withdrawn.

- If formal defects are present (absence of CE marking or EU declaration of conformity, etc.), the market surveillance authority takes restrictive measures if the provider does not remedy them.

SUPERVISION ON GPAIs

The Commission has exclusive competence and exercises it through the AI Office, which investigates possible breaches of the rules, either on its own initiative, based on its monitoring activities, or at the request of market surveillance authorities.

The AI Office:

- receives complaints from downstream providers concerning breaches of the regulation by GPAI model providers as well as reports from the Panel (concerning alleged concrete and identifiable risks at EU level or the classification of GPAI models as ‘systemic risk’) (Art. 89).
- may request documentation and information from the GPAI models provider (Art. 91).
- after consulting the European Artificial Intelligence Board, may conduct assessments of the GPAI to (1) assess compliance with obligations and (2) investigate systemic risks at EU level, including by requesting access to the GPAI in question (Art. 92).
- may require the adoption of measures (compliance/restrictive/systemic risk mitigation) by the GPAI provider.

COOPERATION MECHANISMS BETWEEN EU AND NATIONAL LEVELS

Member States are required to facilitate the tasks of the AI Office (Art. 64(1)) and to inform it of sandboxes and results (Art. 57(15)).

Notifying authorities notify the Commission and the other Member States of conformity assessment bodies and relevant changes to the notification (Artt. 30, 36). They inform the Commission and the other Member States of authorisations derogating from the conformity assessment procedure (Art. 46). In both cases, verification procedures at European level.

In the event of a serious incident, market surveillance authorities notify the measures taken through RAPEX (Art. 73(9)). They may propose joint activities/investigations with the Commission on categories of high-risk AI systems that present a serious risk in two or more Member States (Art. 74(2)). If they consider that high-risk AI systems are not in compliance with the RAPEX system, they may propose joint activities/investigations with the Commission on categories of high-risk AI systems that present a serious risk in two or more Member States (Art. 74(3)). 11). If they consider that GPAI systems that can be used directly by deployers for at least one high-risk purpose do not comply with the requirements of the Regulation, they cooperate with the AI Office to carry out compliance assessments and may request the AI Office for access to information on the AI model needed to conclude investigations on a high-risk AI system (Art. 75(2)-(3)).

PENALTIES

Artt. 99, 101

Member States shall lay down the rules on sanctions ('effective, proportionate and dissuasive') and other enforcement measures.

- Violation of Article 5 prohibitions: up to € 35 milions or 7% of worldwide turnover, whichever is higher.
- Violation of requirements for high risk AI systems and transparency obligations under Art. 50: up to € 15 millions or 3% total worldwide turnover whichever is higher.
- Provision of incorrect, incomplete or misleading information to notified bodies or competent authorities: up to € 7.5 millions or 1% total worldwide turnover, whichever is greater.
- Infringements committed by GPAI models providers (including failure to comply with requests for documents/information and failure to grant the Commission access to the model): up to 3% of total worldwide turnover or EUR 15 millions, whichever is higher. Penalties are imposed by the Commission. Judicial reviews are carried out by the Court of Justice.

AI REGULATORY SANDBOXES

Art. 57

Member States (including jointly) establish regulatory sandboxes for AI, providing a controlled environment (under the guidance of the competent authorities) that facilitates the development, training, testing and validation of innovative AI systems for a period of time prior to their placing on the market/commissioning; within the sandboxes, personal data may be processed under certain conditions and with appropriate measures.

The competent authorities may suspend the testing process if significant risks emerge that cannot be mitigated with appropriate measures. Providers and potential providers participating in sandboxes remain liable for damages to third parties but, if they have complied with the plan and terms of participation and followed the guidelines of the competent authorities, they are exempt from penalties.

The functioning of sandboxes will be defined by Commission implementing acts so as to ensure broad and equal access, flexibility, free of charge for SMEs, etc.

ENTRY INTO FORCE AND APPLICATION

The AI Act will enter into force 20 days after its publication in the Official Journal of the EU and will start to apply 24 months after its entry into force, except for:

- the prohibitions on prohibited practices, which will apply 6 months after entry into force;
- the codes of good practice (9 months after);
- the rules on AI systems for general purposes, including governance (12 months);
- the obligations for high-risk systems (36 months).

Without prejudice to the application of the prohibitions, exemptions and adaptation periods (3-6 years) are provided for GPAIs/high-risk AI systems placed on the market/commissioned before 12 months after entry into force).

IMPACT OF THE AI ACT ON THE ECOSYSTEM

GREATER RELIANCE ON AI	Increased adoption of AI by citizens and consumers
ALLOCATION OF RESOURCES	High investment needed by the public sector
REGULATORY BURDEN	High compliance costs and bureaucracy

The case of the AI Act

To sum up…

Objectives of the AI Act

The proposed regulatory framework on Artificial Intelligence has the following objectives:

1. ensure that AI systems placed on the Union market and used are safe and respect existing law on fundamental rights and Union values;

2. ensure legal certainty to facilitate investment and innovation in AI;

3. enhance governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems;

4. facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation.

Subject Matter of the AI Act

The scope of the AI Act is largely determined by the subject matter to which the rules apply. In that regard, Article 1 states that:

Article 1

Subject matter

This Regulation lays down:

(a) harmonised rules for the placing on the market, the putting into service and the use of artificial intelligence systems (‘AI systems’) in the Union;

(a) prohibitions of certain artificial intelligence practices;

(b) specific requirements for high-risk AI systems and obligations for operators of such systems;

(c) harmonised transparency rules for AI systems intended to interact with natural persons, emotion recognition systems and biometric categorisation systems, and AI systems used to generate or manipulate image, audio or video content;

(d) rules on market monitoring and surveillance.

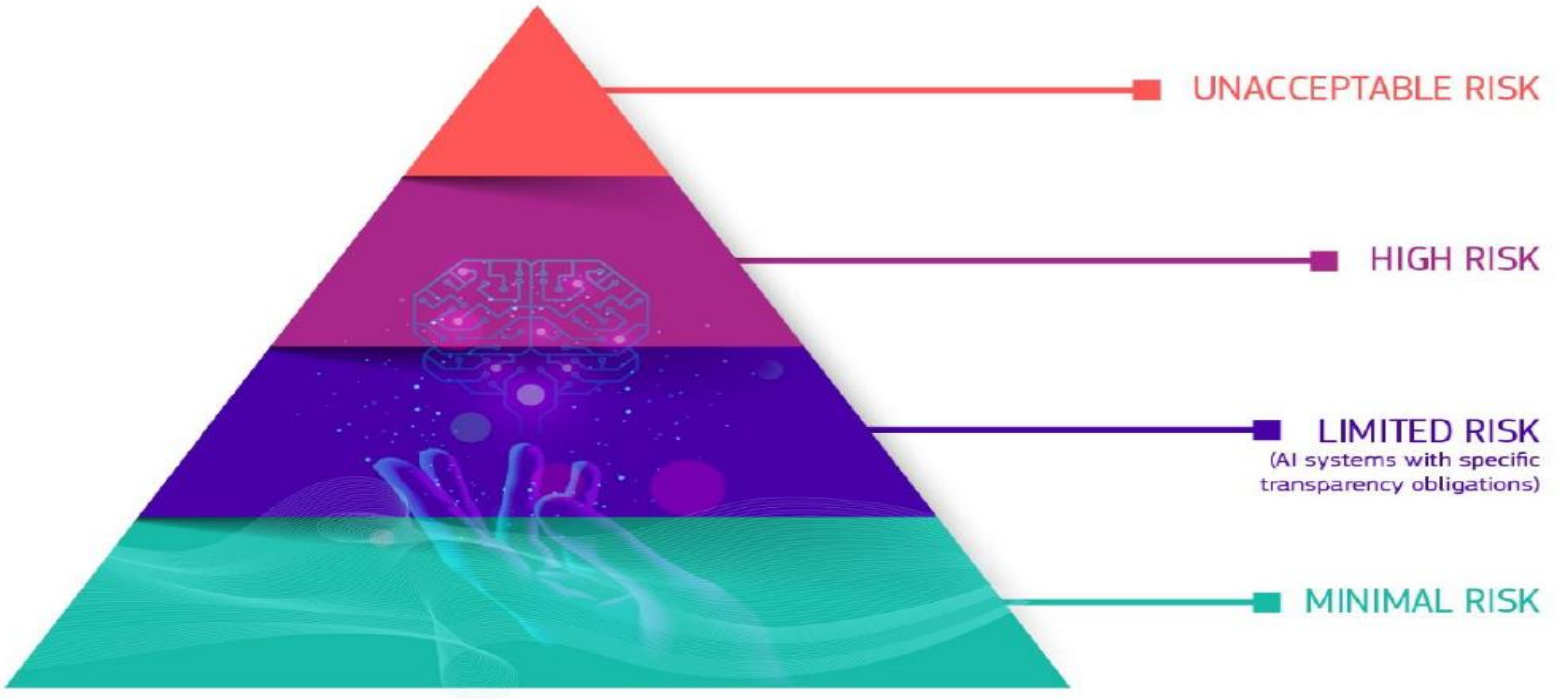
Pyramid of Criticality: Risk based approach

To achieve the goals outlined, the Artificial Intelligence Act draft combines a risk-based approach based on the pyramid of criticality, with a modern, layered enforcement mechanism.

This means, among other things, that a lighter legal regime applies to AI applications with a negligible risk, and that applications with an unacceptable risk are banned.

Between these extremes of the spectrum, stricter regulations apply as risk increases. These range from non-binding self-regulatory soft law impact assessments accompanied by codes of conduct, to heavy, externally audited compliance requirements throughout the life cycle of the application.

Pyramid of Criticality: Risk based approach



The Pyramid of Criticality for AI Systems

Unacceptable Risk AI systems

Unacceptable Risk AI systems can be divided into 4 categories: two of these concern cognitive behavioral manipulation of persons or specific vulnerable groups. The other 2 prohibited categories are social scoring and real-time and remote biometric identification systems. There are, however, exceptions to the main rule for each category. The criterion for qualification as an Unacceptable Risk AI system is the harm requirement.

Examples of High-Risk AI-Systems

Hi-Risk AI-systems will be carefully assessed before being put on the market and throughout their lifecycle. Some examples include:

- Critical infrastructures (e.g. transport), that could put the life and health of citizens at risk
- Educational or vocational training, that may determine the access to education and professional course of someone's life (e.g. scoring of exams)
- Safety components of products (e.g. AI application in robot-assisted surgery)

Unacceptable Risk AI systems

Unacceptable Risk AI systems

- Employment, workers management and access to self-employment (e.g. CV sorting software for recruitment procedures)
- Essential private and public services (e.g. credit scoring denying citizens opportunity to obtain a loan)
- Law enforcement that may interfere with people's fundamental rights (e.g. evaluation of the reliability of evidence)
- Migration, asylum and border control management (e.g. verification of authenticity of travel documents)
- Administration of justice and democratic processes (e.g. applying the law to a concrete set of facts)
- Surveillance systems (e.g. biometric monitoring for law enforcement, facial recognition systems)

Market Entrance of High-Risk AI-Systems: 4 Steps

- In a nutshell, these 4 steps should be followed prior to Hi-Risk AI-Systems market entrance. Note that these steps apply to components of such AI systems as well.

1. A High-Risk AI system is developed, preferably using internal ex ante AI Impact Assessments and Codes of Conduct overseen by inclusive, multidisciplinary teams.
2. The High-Risk AI system must undergo an approved conformity assessment and continuously comply with AI requirements as set forth in the EU AI Act, during its lifecycle. For certain systems an external notified body will be involved in the conformity assessment audit. This dynamic process ensures benchmarking, monitoring and validation. Moreover, in case of changes to the High-Risk AI system, step 2 has to be repeated.

3. Registration of the stand-alone Hi-Risk AI system will take place in a dedicated EU database.

4. A declaration of conformity must be signed and the Hi-Risk AI system must carry the CE marking (Conformité Européenne). Now the system is ready to enter the European markets.

Market Entrance of High-Risk AI-Systems: 4 Steps

But this is not the end of the story...

In the vision of the EC, after the Hi-Risk AI system haven obtained market approval, authorities on both Union and Member State level 'will be responsible for market surveillance, end users ensure monitoring and human oversight, while providers have a post-market monitoring system in place.

Providers and users will also report serious incidents and malfunctioning. In other words, continuous upstream and downstream monitoring.

Since people have the right to know if and when they are interacting with a machine's algorithm instead of a human being, the AI Act introduces specific transparency obligations for both users and providers of AI system, such as bot disclosure. Likewise, specific transparency obligations apply to automated emotion recognition systems, biometric categorization and deepfake/synthetics disclosure. Limited Risk AI Systems such as chatbots necessitate specific transparency obligations as well. The only category exempt from these transparency obligations can be found at the bottom of the pyramid of criticality: the Minimal Risk AI Systems.

In addition, natural persons should be able to oversee the Hi-Risk AI-System. This is termed the human oversight requirement.

Open Norms

The definition of high-risk AI applications is not yet set in stone. Article 6 does provide classification rules. Presumably, the qualification remains a somewhat open standard within the regulation, subject to changing societal views, and to be interpreted by the courts, ultimately by the EU Court of Justice. A standard that is open in terms of content and that needs to be fleshed out in more detail under different circumstances, for example using a catalog of viewpoints. Open standards entail the risk of differences of opinion about their interpretation. If the legislator does not offer sufficient guidance, the courts will ultimately have to make a decision about the interpretation of a standard.

This can be seen as a less desirable side of regulating with open standards. A clear risk taxonomy will contribute to legal certainty and offer stakeholders with appropriate answers to questions about liability and insurance.

Enforcement

The AI Act provides for the installation of a new enforcement body at Union level: the European Artificial Intelligence Board (AI Board). At Member State level, the AI Board will be flanked by national supervisors, similar to the GDPR's oversight mechanism. Fines for violation of the rules can be up to 6% of global turnover, or 30 million euros for private entities.

'The proposed rules will be enforced through a governance system at Member States level, building on already existing structures, and a cooperation mechanism at Union level with the establishment of a European Artificial Intelligence Board.'

AI Office

The Commission has established a new EU level regulator, the European AI Office, which will sit within the Directorate-General for Communication Networks, Content and Technology (DG CNECT) in the Commission.

The AI Office will monitor, supervise, and enforce the AI Act requirements on general purpose AI (GPAI) models and systems across the 27 EU Member States. This includes analysing emerging unforeseen systemic risks stemming from GPAI development and deployment, as well as developing capabilities evaluations, conducting model evaluations and investigating incidents of potential infringement and non-compliance. To facilitate the compliance of GPAI model providers and consider their perspectives, the AI Office will produce voluntary codes of practice, adherence to which would create a presumption of conformity.

The AI Office will also lead the EU in international cooperation on AI and strengthen bonds between the European Commission and the scientific community, including the forthcoming scientific panel of independent experts. The Office will help the 27 Member States cooperate on enforcement, including on joint investigations, and act as the Secretariat of the AI Board, the intergovernmental forum for coordination between national regulators. It will support the creation of regulatory sandboxes where companies can test AI systems in a controlled environment. It will also provide information and resources to small and medium businesses (SMEs) to aid in their compliance with rules

Università Europea di Roma

Co-funded by the European Union

AI & FUNDAMENTAL RIGHTS

AI & Fundamental Rights

Fundamental rights are mentioned throughout the AI Act as an overriding public interest that warrants legislative protection.

In particular, Article 65(1) AI Act extends the definition of product risks to include risks to fundamental rights. The result is a product safety instrument heavily couched in fundamental rights language.

The AI Act is not the first product safety instrument to cover fundamental rights.

AI & Fundamental rights

The EU regulation that lays down harmonized rules for medical devices (Medical Devices Regulation: Regulation (EU) 2017/745) explicitly refers to the protection of fundamental rights in general (Recital 89) and personal data more specifically (Recital 69) while including extra safeguards to two specific freedoms: freedom of expression and freedom of the press.

More generally, the EU is constitutionally required to protect fundamental rights as it exercises its powers, including in product safety.

AI & Fundamental Rights

However, the AI Act **displays a higher level of engagement with fundamental rights than other EU product safety instruments**. This can be seen in the practical requirements imposed on AI systems.

The segmentation of AI systems into various risk tiers **puts risks to fundamental rights on an equal footing with the risks to health and safety that are the bread and butter of product safety law.** Various essential requirements laid down for high-risk AI systems are formulated in terms of fundamental rights, such as the need to indicate circumstances in which the use of the AI system may impose risks or to design suitable mechanisms for human oversight of the AI system. Finally, conformity with essential requirements must be assessed, considering how well an AI system minimizes or eliminates risks to fundamental rights.

AI & Fundamental Rights

: Comparison between rationales in product safety law and constitutional reasoning

Product safety	Fundamental rights
Actuarial risks predominate	Actuarial, sociopolitical, and cultural risks
Risks stem from the technical object	Risks stem from the sociotechnical context
Small world: known and consistent problems	Multidimensional harm and wicked problems
Satisficing technical baselines	Constrained maximization of principles

Fundamental Rights' concerns

In 2019, the EU's High-Level Expert Group (HLEG) on AI published an **updated definition of AI, including its main capabilities and scientific disciplines** (High-Level Expert Group on Artificial Intelligence (HLEG), A Definition of AI: Main Capabilities and Disciplines, ec.europa.eu, p. 6).

According to this definition, AI systems **are designed by humans but can come in different forms, such as machine learning, machine reasoning, and robotics**.

Fundamental Rights' concerns

In all its forms but to varying degrees, AI is currently capable of acquiring, processing, and interpreting large amounts of data, making decisions based on the interpreted data, and translating these decisions into action.

Based on what AI is capable of, four specific characteristics become visible which, however, **do not only come with benefits but may also lead to fundamental rights concerns**.

Privacy concerns & deanonymization

First, AI is dependent on data, hence, it has enhanced capacities to collect and process large amounts of data. This gives AI an **increased power of human observation, for example, through biometric identification in public places, thus raising privacy concerns**.

Secondly, through the connectivity of many AI systems and by analyzing large amounts of data and identifying links among them, **AI may be used to deanonymise large data sets although such data** sets do not include personal data *per se*.

Black-box scenario & discrimination

Thirdly, based on the self-learning ability of AI and, hence, its increasing autonomy, coupled with the enhanced capacity of AI to learn quickly and explore decision paths that humans might not have thought about, AI is able to **find patterns of correlation within datasets without necessarily making a statement on causation**. Consequently, AI may produce new solutions that may be impossible for humans to grasp by making decisions without the reasons being known, potentially resulting in AI opaqueness. This opaqueness is **also known as the 'black-box phenomenon'** which drastically reduces the explainability of AI.

Fourthly, the training data of AI systems may be biased, leading to **AI systems producing discriminatory results**.

Fundamental Rights protection and EU Treaties

The EU Treaties provide for a **general guarantee of fundamental rights** protection.

Nonetheless, general principles of EU law have been constituting the principal source of fundamental rights protection in the EU whereby the **Charter of Fundamental Rights of the EU (the Charter) now codifies these fundamental rights**.

Specifically, **Arts 7, 8, and 21 lay down the rights to privacy, protection of personal data, and non-discrimination, respectively**. The European Commission has expressed **concerns regarding the limited scope of application of the EU Charter in the context of the AI discussion** (European Commission, Structure for the White Paper on artificial intelligence – a European approach).

AI systems and Charter scope of application

According to Art. 51 of the Charter and the case law of the CJEU, the Charter and general principles of EU law **apply to any action falling within the scope of EU law**.

Consequently, certain Member States' actions involving the development and/or use of AI systems may not fall within the Charter's field of application and may, thus, potentially lead to a compromised fundamental rights protection. For example, the **use of AI systems in the industry or the health sector is only partially or not covered at all by the Charter's scope of application because these fields fall primarily within the exclusive competences of the Member States**.

AI systems and Charter scope of application

Nevertheless, the EU often takes on an active supportive role to protect fundamental rights by adopting guidelines, even in areas that fall outside its main competences. For example, in the health sector, the Commission has adopted guidelines for Member States on the Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT) app, designed to help tackle the Covid-19 crisis by tracing infection chains, even across borders. The app is largely based on advanced algorithms and, hence, touches upon privacy and data protection concerns of interest by the Union.

AI systems and Charter scope of application

Another concern that was raised by the Commission was the **lack of horizontal direct effect of the Charter**. However, it must be noted that the Court has practically acknowledged the direct horizontal application of the Charter in specific situations, namely when EU secondary law gives expression to a general principle of EU law, such as the principles of privacy and protection of personal data and non-discrimination.

Hence, the use of AI systems must be in conformity with these principles, even in horizontal situations falling within the scope of EU law. For example, the observance of the principle of non-discrimination in situations covered by Directive 2000/78/EC on equal treatment in employment and occupation is particularly important when AI systems are used for recruitment purposes in employment matters, amongst others.

Fundamental Rights and GDPR

The GDPR is, amongst others, **specifically intended to apply to partly or fully automatic AI systems that process personal data forming part or intended to form part of a filing system**.

At the same time, the use of AI systems is limited under the GDPR. For example, while the GDPR applies to the processing of personal data by wholly automated means, Art. 22, para. 1, prohibits the use of fully autonomous AI systems for the processing of personal data which produces legal effects for individuals.

Hence, the GDPR limits the development and use of AI to systems that still function with some sort of meaningful human oversight.

Fundamental Rights and GDPR

Additionally, also functioning as one exception to the prohibition laid down in Art. 22, para. 1, the processing of personal data can only take place based on the specific consent of the data subject. The concept of specific consent entails informed consent, meaning that the **data subject must not only be informed that her personal data is being processed but also about how and for what purposes the processing takes place**.

While, in theory, the requirement of consent should provide for sufficient safeguards against fundamental rights violations by AI systems processing personal data, **it is difficult to obtain informed consent when AI systems make unpredictable decisions**.

Fundamental Rights and GDPR (AI and consent)

Moreover, the means of obtaining the specific consent of the data subject, such as “I have read and agree to the Terms”, is one of the biggest lies on the internet that poses the risk of rendering the protection offered by the concept of specific consent inefficient. To avoid this, it can be assumed that the use of fully, as well as partly automated AI systems, is further limited by the principle of controller responsibility under the GDPR.

For example, in Google Spain, the CJEU found that a search engine operator is a controller within the meaning of Art. 4, para. 7, GDPR when she processes personal data. This is when the activity of the search engine consists of finding information, indexing it automatically, storing it temporarily, and making it available to internet users, when that information consists of personal data. If this is the case, the controller has a responsibility to, under specific circumstances, remove searches based on a person’s name from the list of results. Although certain of these processing procedures by a search engine may be done by AI systems, it is the search engine operator who has the ultimate responsibility, thus limiting the use of AI systems in such circumstances.

Fundamental Rights and GDPR

Moreover, in GC and Others v. CNIL, the Court held that it is the responsibility of a search engine operator, when receiving a de-referencing request, to balance the right to personal data protection against other rights which may be affected by the de-referencing, for example, the right to freedom of information. Hence again, the use of AI systems for the operation of search engines is limited by the operator's responsibility to oversee and guarantee the necessary fundamental rights protection. In conclusion, this means that the full potential of AI can never be used in situations falling under the GDPR.

Considering this in the light of fundamental rights, the development and use of AI systems are generally limited by the concepts of specific consent and controller responsibility to safeguard the protection of the rights of the data subjects.

Fundamental Rights and GDPR (transparency and explainability)

As regards the opacity in AI decision-making, the GDPR requires the observance of the principles of transparency and explainability, including the data subject's rights to information and access to personal data. To uphold these principles, this also includes ex ante measures within the development phase of AI systems, such as conducting data protection impact assessments (DPIA) and implementing appropriate technical and organizational measures to help implement the data protection principles, also called data protection by design.

This means that developers of AI systems have a duty to build in safeguards that provide for a guarantee to uphold the data protection principles in the first place. In light thereof, three issues arise.

Fundamental Rights and GDPR

First, the concept of personal data in Art. 4, para. 1, of the GDPR is very broad and has been further expanded by the Court in cases like YS and Others, Nowak, and Breyer (Court of Justice: judgment of 29 June 2010, case C-28/08, Commission v. Bavarian Lager, paras 49-50; judgment of 20 December 2017, case C-434/16, Nowak, paras 54-55; joined cases C-141/12 and C-372/12, YS and Others, paras 45-47).

Hence, it is not exhaustively defined what personal data is which may make it difficult to determine the bounds of AI use for data processing purposes. This is problematic because AI systems cannot necessarily be simply aborted if they become independent, hence, the bounds of AI use should be determined in the development phase already. On the other hand, a broad concept of personal data guarantees to cover nearly all eventualities and thus reflects a technological reality. The very fact that a piece of information has been created or merely distributed by an individual may provide some clues about who that individual may be and AI is able to detect such correlations better than humans.

Fundamental Rights and GDPR (AI discrimination)

Lastly, regarding AI discrimination, the GDPR's prohibition of the processing of special categories of personal data – meaning data that also constitute potential grounds for discrimination – by solely automated means offers a concrete protection against AI discrimination. Unfortunately, the special categories of personal data laid down in Art. 9, para. 1, of the GPDR do not include the categories of colour, language, membership of a national minority, property, and birth which are, however, recognised as grounds of discrimination in Art. 21, para. 1, of the Charter. This constitutes a potential gap in the prevention of discriminatory results through personal data processing, both by AI systems and conventional means.

Fundamental Rights and GDPR (AI discrimination)

Moreover, Art. 22, para. 1, GDPR, further underlined by Art. 35, para. 3, prohibits profiling by fully automated means. Profiling is a form of processing carried out on personal data to evaluate personal aspects about a natural person and, as the name says, create profiles. This process places people in categories based on their personal traits and is thus likely to lead to discrimination. More specifically, data subjects are likely to be objectified because AI systems evaluate individuals by the probability of a group based on correlation and statistical models and thus do not regard individuals in light of their own rights. The prohibition in Art. 22, para. 1, GDPR provides for guarantees against such discrimination. However, the data subject's specific consent constitutes an exception to the prohibition whereby the same issues surrounding specific consent as explained above may arise, thus rendering the protection granted by Art. 22, para. 1, of the data subject's rights inefficient.