Gatt L, Lo Conte MT, Mazzarella ME, 'Analysis and perspectives on the AI Act: scope and impact assessment' (2025) 2 EJPLT.
DOI: https://doi.org/10.57230/EJPLT252LGMTLCMEM

EJPLT
ISSN 2704-8012



Analysis and perspectives on the AI Act: scope and impact assessment

Analisi e prospettive dell'AI Act: ambito di applicazione e valutazione d'impatto

LUCILLA GATT
Full Professor of Private Law, Suor Orsola Benincasa University of Naples
Director, Research Centre of European Private Law (ReCEPL)

MARIA TERESA LO CONTE
Junior Researcher, Research Centre of European Private Law (ReCEPL)
Ph.D. student, Università degli Studi di Roma "Tor Vergata"

MARIA EUGENIA MAZZARELLA
Ph.D. student, Scuola Superiore Meridionale

Abstract

The article specifically examines the AI Act by analysing its objective, subjective and territorial scope. After an examination of the actors involved, dwelling on all the actors mentioned in the Regulation, the risk-based approach of AI systems and models is analysed, and the specific obligations of the actors along the AI chain of responsibility are outlined. Particular emphasis is then placed on the Fundamental Rights Impact Assessment (FRIA) and the ethical and legal challenges posed by generative AI, emphasising the need for an anthropocentric approach and governance that safeguards human rights in the digital age. Finally, the contribution also reflects on the role of AI in democratic life and its influence on information and political power.

* This article is the updated version, translated in English of the contribution L Gatt, MT Lo Conte, ME Mazzarella, 'L'ambito di applicazione, soggettivo e oggettivo, dell'AI Act' in U. Ruffolo (ed.), *AI Act - La regolamentazione europea dell'Intelligenza artificiale* (Luiss University Press, 2025), 49 ff.

© The author(s) 2025, published by Suor Orsola Benincasa Università Editrice.
This contribution is licensed under a Creative Common Attribution 4.0 International Licence CC-BY-NC-ND, all the details on the license are available at:
https://creativecommons.org/licenses/by-nc-nd/4.0/

Abstract
L'articolo esamina in modo specifico il Regolamento dell'Unione Europea sull'Intelligenza Artificiale (AI Act) analizzandone l'ambito di applicazione oggettivo, soggettivo e territoriale. Dopo la disamina dei soggetti coinvolti, soffermandosi su tutti gli attori menzionati dal Regolamento, viene analizzato l'approccio basato sul richio dei sistemi e modelli di IA enucleando inoltre gli specifici obblighi degli attori lungo la catena di responsabilità dell'IA. Particolare enfasi viene poi posta sulla valutazione d'impatto dei diritti fondamentali (Fundamental Rights Impact Assessment - FRIA) e sulle sfide etiche e legali poste dall'IA generativa, sottolineando la necessità di un approccio antropocentrico e di una governance che salvaguardi i diritti umani nell'era digitale. Il contributo riflette in ultimo anche sul ruolo dell'IA nella vita democratica e sulla sua influenza sull'informazione e sul potere politico.

Keywords: AI Act; Application; AI Systems; AI Models; FRIA

Summary: 1. Introduction. – 2. Territorial and material scope. – 3. Subjective scope. – 3.1 Provider. – 3.1.1 Extra-Providers. – 3.1.1.a Prospective provider. – 3.1.1.b Downstream provider. – 3.2 Product Manufacturer. – 3.3 Deployer. – 3.4 Other Operators. – 3.5 Interposing Subjects. – 3.6 Institutional Actors. – 4. Objective Scope. – 4.1 AI Systems. – 4.1.1 Prohibited AI practices. – 4.1.2 High-risk AI systems. – 4.2 General Purpose AI Models. – 4.2.1 Systemic AI Models. – 4.2.2 Generic AI Models. – 4.2.3 Certain AI systems. – 5. FRIA - Fundamental Rights Impact Assessment – 5.1 Possible new developments. – 6. Conclusions.

1. Introduction.

In the long process of European Union standardisation, the subject of Artificial Intelligence has gone through a tortuous path of proposals, opinions and corrigenda, but 21 May 2024 saw the final decision by the Council of the European Union:[1] the Artificial Intelligence Act (now and henceforth 'AI Act') is law (*rectius* Regulation) and in its final version aims to regulate the use of Artificial Intelligence (now and henceforth 'AI') in the territory of the Union.

The AI that is the subject of EU regulation is understood in a broad and dynamic sense in order to remain relevant in the rapid course of technological innovation. Pursuant to Article 3(1)(1), in going on to specify Recitals 12 and 4, the 'AI system' that is the subject of this regulation is in fact defined as an automated system that, jointly:

- is designed to operate with varying levels of autonomy;
- can be adaptable after deployment;
- deduces from the input it receives how to generate output (in the guise of predictions, content, recommendations and decisions) that can influence both physical and virtual environments.

[1] Council of the EU (2024, 21 may) Artificial intelligence (AI) act: Council gives final green light to the first worldwide rules on AI [Press release]. https://www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/
Vote: eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CONSIL:ST_10199_2024_INIT

Having established a clear, comprehensive, and inclusive definition of AI and identified its fundamental characteristics – autonomy, adaptability, and generativity – it is widely recognized that AI refers to technologies designed to replicate human intellect and cognitive abilities, such as learning and problem-solving, with the goal of performing tasks in real-world contexts.[2]

To establish a uniform legal framework for the development, market placement, deployment, and use of safe, reliable, ethical, and human-centered AI systems, the EU legislator immediately recognized the necessity of harmonized protection across Member States, balancing the risks and benefits of AI. Indeed, given the exponential use of AI, it naturally brings economic, environmental, and social benefits, yet it also inevitably presents inherent risks. These include potential infringements on public interests and fundamental rights, as well as possible 'physical, psychological, social, or economic harm'.

To understand this ambivalence of new technologies, we must take a brief step back.

In the field of AI, a longstanding debate exists between proponents of engineering-reproductive AI and semantic-productive AI. The former is characterized by AI systems that replicate, in the sense of copying, certain human behaviors through algorithms programmed by engineers and developers. The latter focuses on developing machines and digital systems capable of functioning entirely autonomously, learning from their experience (essentially their operational processes, i.e., machine learning), and potentially surpassing human capabilities. So far, the first category has produced the most concrete and operationally manageable results, though research into the second type is advancing rapidly. This fuels collective imagination, as these AI systems can manage complex situations and make autonomous, unforeseen, and unpredictable decisions – this so-called 'strong AI' is, unsurprisingly, the most challenging in terms of human control and regulatory oversight.

Since the Dartmouth College conference in 1956, where the term 'Artificial Intelligence' was first coined for a branch of computational informatics that had already been in development for some time, the field has experienced alternating periods of significant and modest investment (the so-called AI 'springs and winters').

In the last decade, investments in new technologies have surged dramatically, mainly due to two factors: a significant increase in computational power and the vast amount of data available to train AI algorithms, coupled with a reduction in sensor costs, which facilitate data collection.

Generative AI, in particular, has undergone exceptional development: it has demonstrated an unprecedented ability to generate digital content surpassing anything previously created by humans. However, although GenAI can produce new content, but cannot currently generate new ideas or solutions to real-world challenges, as it does not understand the social connections or relationships underlying the language or content provided to it. [3] These systems still lack autonomous semantic capability. Generative AI systems based on large language models (LLMs) operate purely on statistical reasoning and

[2] IBM, 'Cos'è l'intelligenza artificiale (AI)?' (*IBM*) <www.ibm.com/it-it/topics/artificial-intelligence>.

[3] S da Empoli, R Loiero, 'L'Intelligenza Artificiale tra rischi e opportunità' in R. Razzante (ed.), *Manuale sull'Intelligenza Artificiale* (Giappichelli Editore 2024) 33.

probability. This functional model has led some authors to describe AI as *agere sine intelligere* (acting without understanding).[4]

Artificial Intelligence, in general, can be considered a dual-faced concept. Just as a knife can be both a useful tool for nourishment and a weapon for harm, so-called intelligent technologies offer significant opportunities as well as risks. AI systems can replace humans in dangerous, exhausting, and repetitive tasks (e.g., applications in agriculture, industry, and military sectors). They can also predict human behavior, identify diseases early, and assess risks to people's interests and well-being.

Thus, if today we can discuss soft robotics[5] in healthcare and robotic surgery, if the latest robotic innovations are presented annually at the World FIRA in agriculture, and if, along with logistics, these fields have been identified by the International Federation of Robotics in World Robotics 2023[6] as three of the five key areas for the deployment of so-called Professional Service Robots, it is undoubtedly thanks to both the *homo faber* who has transformed into the *homo creator* – a producer of artificial nature – [7] and the continuous advancements in AI and its associated benefits.

Beyond these advantages, however, generative AI systems can have ambivalent effects on productivity, economic growth, and especially employment. Other risks concern potential violations of individual privacy and fundamental rights in general, with the possibility of perpetuating large-scale discriminatory patterns.

This broad range of opportunities and risks stems from the intrinsic characteristics of generative AI, which many experts classify as a so-called general-purpose technology (GPT) due to its pervasiveness (it can be applied across multiple sectors), its high potential for technological development, and its complementarity with other technological innovations.[8]

Although previous predictions of widespread mechanization of the production sector have not materialized, generative AI could still enhance productivity in several ways. Firstly, its ability to replace humans in certain tasks – whether manual or intellectual – allows individuals to focus on more productive activities. Moreover, those forced to change jobs may find new opportunities, potentially created by future AI implementations. This could provide greater opportunities for younger generations, who are more inclined to develop digital skills, but at the same time, it might disadvantage those in advanced career stages who may struggle to acquire new competencies (digital divide).

The public and academic debate of the past decade regarding AI's effects on employment has been divided between a positive perspective – arguing that AI systems would boost productivity, significantly optimize economic processes, and negatively impact only a limited number of job roles – and a negative view,

---

[4] L Floridi, F Cabitza, *Intelligenza Artificiale. L'uso delle nuove macchine* (Bompiani 2021).

[5] M Yip and others 'Artificial intelligence meets medical robotics' (2023) 381(6654) Science 141 – 146.

[6] See International Federation of Robotics, *World Robotics* (2023) 36. <https://ifr.org/img/worldrobotics/2023_WR_extended_version.pdf>

[7] PP Portinaro, 'Introduzione' in PP Portinaro (ed.), *Il principio di responsabilità. Un'etica per la civiltà tecnologica* (Einaudi 2009) XXIV.

[8] TH Bresnahan, M Trajtenberg, 'General purpose technologies 'Engines of growth'?' (1995) 65(1) Journal of Econometrics 83-108.

predicting a near-total replacement of human workers in both manual and intellectual tasks, leading to widespread unemployment.

As is often the case, the truth lies somewhere in the middle. Recently, two studies – conducted by the OECD (2023) and the International Monetary Fund (2024) – have analyzed AI's potential impact on the labor market.

The OECD report focused on AI in general, highlighting that AI will replace certain human jobs (displacement effect), but it can also increase labor demand due to productivity growth (productivity effect) and AI may also create new job opportunities (reintegration effect).[9] The most at-risk workers are those with low qualifications: according to the OECD, occupations at risk of automation account for 27% of total employment, rising to 30% in Italy.

The International Monetary Fund study focused on generative AI's impact on employment by applying an index measuring the exposure and complementarity of human labor concerning AI. According to this index, high exposure reduces labor demand, while high complementarity can increase it by boosting productivity through human-machine collaboration, provided workers possess adequate skills.[10] The IMF estimates that 40% of jobs are at risk of automation, increasing to 60% in advanced economies.

Other negative impacts on fundamental rights such as privacy are due to the structural flaws of new technologies: new generative AI chat-bots can produce so-called hallucinations, i.e. false but truthful content. These incorrect out-puts can influence public opinion and through the improper treatment of personal data damage the privacy and potentially the dignity and reputation of those involved, as was the case with Brian Hood, now mayor of an Australian city, who had denounced the wrongdoings of the company he worked for, but whom Chat-GPT identified as one of those actively involved in the scandal.

The exceptional potential accompanied by equal risks of the new technologies must push those in charge of their governance to lay the foundations for a sustainable digital transition by setting up a clear regulatory framework, of which the regulation that is the subject of this contribution is a commendable starting point.

With regard to the purpose and subject matter of the AI Act, the legislator, in Article 1, specifying Recitals 1-6 and 8 and briefly anticipating the content of the legislation, states that the Regulation introduces a real set of harmonised rules for the use and market of artificial intelligence systems in the EU, providing both prohibitions for certain types of practices and specific requirements for certain systems and models, but at the same time also providing for monitoring, supervision and support measures for innovation, with particular attention to SMEs and start-ups.

2. Territorial and material scope.

In defining this scope of application, it is essential to clarify from the outset that the actors of the AI Act must comply with the obligations of the Regulation in a non-discriminatory manner, regardless of their place of establishment,

[9] da Empoli, Loiero (n 2) 42.
[10] da Empoli, Loiero (n 2) 42.

location, or physical presence. In this regard, the Regulation has intentionally introduced the so-called Brussels effect to indirectly extend its provisions beyond the geographical boundaries of the European Union. This intent is further reinforced in Article 2 of the Regulation, which explicitly states that it applies to providers [and deployers] established or located in a third country outside the EU if the outputs of the AI system are used within the Union, as well as to Authorised Representatives of providers not established in the Union.

In light of this strategy, the European legislator – drawing from the experience of the GDPR and the EDPB Guidelines[11] – has chosen to combine the principle of establishment with an approach that more effectively considers the influence (and legal authority) of the European Union. This applies whether the providers are located within the Union or in a third country, whether the deployers are situated within the Union, or whether providers or deployers of AI systems from non-EU countries intend to use them within the EU. On this point, under the AI Act, territorial considerations become irrelevant due to the introduction of a Union-centric approach, taking into account the fact that the world's leading AI powers are not European.

In this regard, the European legislator has defined a precise characteristic of this regulation: whenever an AI system is placed on the market or put into service in Europe, or whenever the output of an AI system has an impact on the territory, market, or individuals within the Union, it must necessarily be subject to this Regulation in order to safeguard the internal market and the fundamental rights and freedoms of EU citizens.

However, it is important to note that this extension – sectorally speaking – is not unlimited: the Regulation applies only to sectors falling within the scope of EU law. In this sense, it has struck the right balance in strengthening the European Single Market while simultaneously ensuring the outright exclusion of the potentially harmful impacts of a – albeit utopian – Korean-style Data Dam[12], which envisions total data openness for AI training. By asserting the supremacy of EU law in the face of AI economic giants, the Regulation has imposed its unequivocal compliance as a condicio sine qua non for any international entity seeking to operate in the AI market within the Union.

3. Subjective scope.

While a first reading of the Regulation may suggest that its scope of application is outlined in the aforementioned Article 2, a thorough examination of the entire text is essential to understand that the actors involved are numerous and diverse. The entities concerned can be categorized into two main groups – Operators and Institutional Actors – which, despite comprising distinct and well-defined subjects, maintain multiple points of interaction and

[11] European Data Protection Board, 'Guidelines 3/2018 on the territorial scope of the GDPR (Article 3)' (2018) <https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_3_2018_territorial_scope_en.pdf>.

[12] Republic of Korea Ministry of Science and ICT (2020, 2 Sept), Data Dam project begins, being key to Digital New Deal [Press Release] <https://www.msit.go.kr/eng/bbs/view.do?sCode=eng&mId=4&mPid=2&bbsSeqNo=42&nttSeqNo=453>.

reciprocal implications. The Provider, the product manufacturer, the deployer, the Authorised Representative, the importer, and the distributor all belong to the category of Operators, as each operates within the AI value chain, and the Regulation assigns each of them specific obligations, rights, duties, and prerogatives. At the same time, Institutional Actors are manifold, some newly introduced, others long established: the AI Office, the Market Surveillance Authority, the Enforcement Authority, and the European Commission are some examples. Finally, there are additional entities that, acting as Intermediary Subjects, are either involved in the activities of Operators or necessary for the practical implementation of the AI Act (Fig. 1). Moreover, while the notion of Operator is explicitly provided by Article 3, paragraph 1, point 7 of the AI Act, and other entities are regulated either explicitly or implicitly, it becomes evident that a true understanding of the AI Act often requires looking beyond the formal qualification assigned to them by the Regulation.

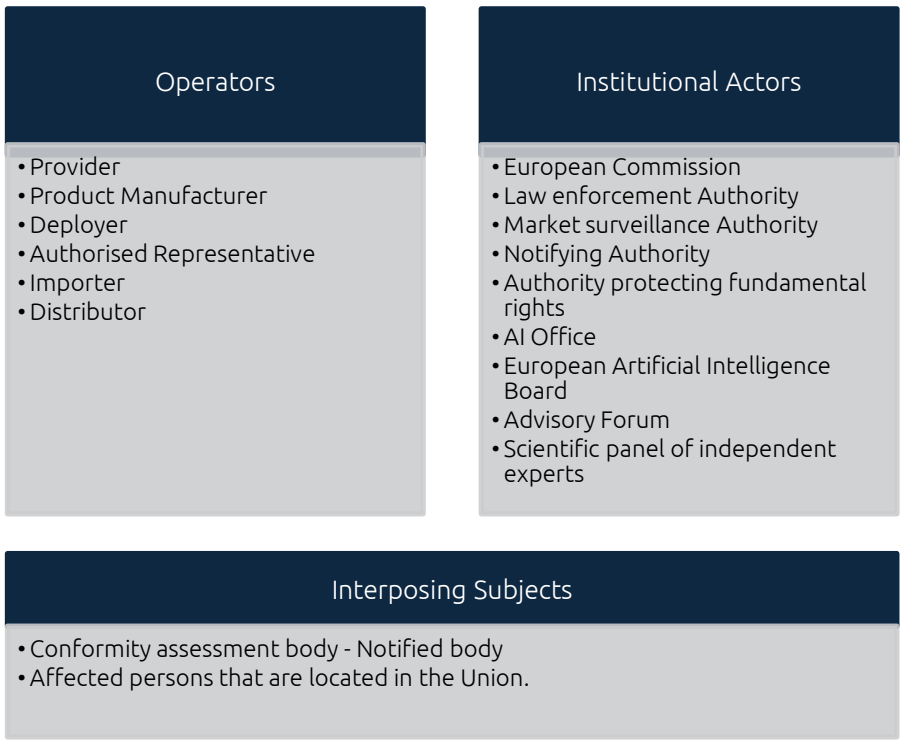


Fig. 1

3.1 Provider.

Among the Operators, the first entities to be highlighted are Providers, namely all those – whether natural persons, legal entities, public authorities, agencies, or other organizations – that develop or have third parties develop an AI system or a general-purpose AI model and subsequently, separately:

- place the AI system or general-purpose AI model on the market, meaning its initial availability on the Union market;

- put the AI system into service under their own name or brand, for any purpose (whether for consideration or free of charge), meaning the provision of such a system by the Provider to the deployer for first use or for their own use in the Union market.

The definition of Provider, as established in Article 3, paragraph 1, point 3 of the AI Act, appears to be all-encompassing, as it does not make formal distinctions between public, private, institutional, or non-institutional entities but instead assesses their actual activities and thus their substantive qualification. This entity is responsible for the placement on the market or the putting into service of an AI system, particularly if it is high-risk. Furthermore, while AI literacy remains necessary to equip Providers with the knowledge required to make informed decisions about AI systems, they are subject to various highly detailed obligations, depending on the specific AI system or model involved.

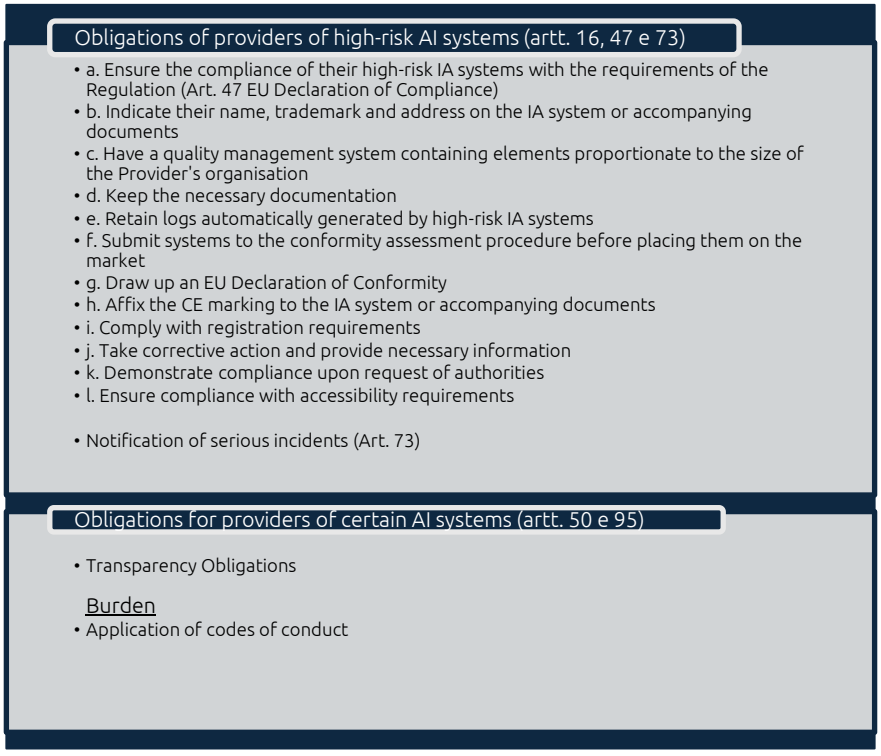


Fig. 2

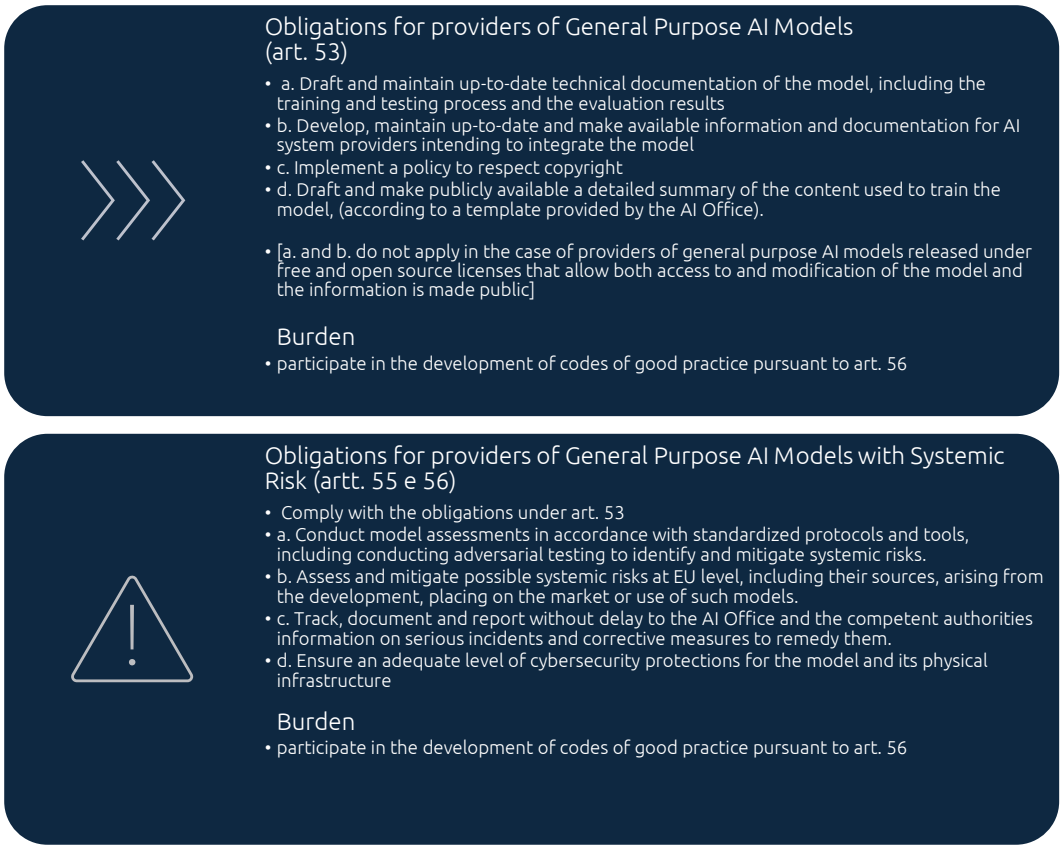


Fig. 3

3.1.1 Extra-Providers.

Alongside the figure of the so-called pure provider, it appears necessary to identify further figures: that of the prospective provider and the downstream provider.

3.1.1.a Prospective provider.

With regard to the Prospective provider, the AI Act does not provide a specific definition. Therefore, it is undoubtedly necessary to refer to the general definition set out in Article 3 of the AI Act, but understanding it in the merely potential form of its typical activities: development and putting into service or placing on the market.

This definition appears necessary, as it is not possible to include this figure among the 'affected persons' under Article 2, paragraph 1, letter g) of the Regulation; this concept only includes individuals who 'are located in the Union', whereas a prospective provider could very well be located outside EU territory and engage in practices that circumvent the Regulation.

This extension of meaning can be traced to the rationale of the AI Act and the Union legislator's intention to encompass as broad a subjective scope as possible, provided that the AI output affects the European market. Confirmation of this interpretation can primarily be found in Recitals 140 and

141, but above all within the main text, where prospective providers are explicitly included among the entities authorized to develop, train, validate, and test an innovative AI system in a regulatory sandbox.

At the same time, a specific procedure is established: they must agree on a detailed experimentation plan with the competent authority and be monitored by the authorities to ensure compliance – both regulatory and procedural – with the AI Act.

3.1.1.b The Downstream provider.

The Downstream provider perfectly fits into the so-called AI supply chain, representing the first key entity in the most complex scenario of an AI system integrating an AI model. As highlighted in Article 3, paragraph 1, number 68) of the AI Act, it is essential to specify that such a model can either be supplied directly by the same downstream Provider (and then vertically integrated by them) or provided by another external entity based on contractual agreements, followed by the subsequent integration of the model into an AI system by the downstream Provider. In the second scenario, specific complexities arise when an AI system is integrated with a general-purpose AI model, as general-purpose AI models – supplied by their respective providers – often form the foundation of downstream systems provided by downstream providers. (Fig. 4)

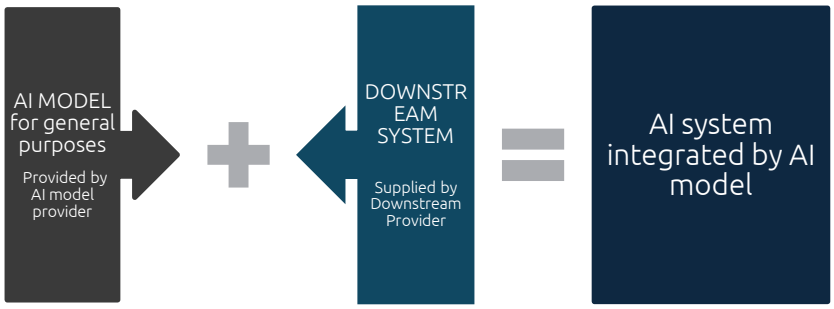


Fig. 4

In light of this significant connection between the two entities – both 'providers' but with different qualifications – the Regulation establishes specific transparency measures for the former when supplying general-purpose AI models. These measures ensure that the downstream Provider wishing to use a given AI model for its system is aware of its characteristics and can, in turn, comply with the obligations set out in the Regulation.

To foster a proper understanding of general-purpose models and their actual capabilities, the Regulation, in Recital 101, considers it appropriate to impose proportionate transparency measures on the model Provider. These include drafting and updating documentation and the general 'provision of information' on the general-purpose AI model to facilitate its use by downstream providers, incorporating all the indications listed in Annex XII of the AI Act.

At the same time, regarding obligations, reference is undoubtedly made to the requirement to label non-human outputs created by AI, the so-called

synthetic content. Recital 133 explicitly supports this, facilitating compliance for the downstream Provider of the AI system. If this operation is performed upstream by the Provider of the general-purpose AI model, it undoubtedly simplifies the work of the downstream Provider, who, in this case, supplies a general-purpose AI system integrating a general-purpose AI model.

It is evident that the EU legislator assigns particular importance to the figure of the downstream Provider. This is confirmed from Recital 162, which anticipates Article 89 of the AI Act and foresees its involvement in effective monitoring activities carried out by the AI Office. In fact, within the monitoring actions for enforcing obligations imposed on providers of general-purpose AI models, entrusted to the AI Office by the European Commission, the downstream Provider serves as the trait d'union between economic and institutional actors. It is formally granted, under Article 89(2) of the AI Act, the right to file complaints regarding potential violations of the Regulation by providers of AI models and general-purpose AI systems.

Due to its monitoring responsibilities, the downstream Provider benefits from the protections that the Regulation grants to so-called informants or whistleblowers, particularly those outlined in Directive (EU) 2019/1937[13], as stipulated in Recital 172 of the Regulation.

Finally, in light of the strong emphasis placed on soft law measures, especially codes of good practice, the downstream Provider is explicitly included in Article 56(3) among those entities eligible to support the development of such measures.

In conclusion, it is possible to affirm that the downstream Provider, beyond acting as a warning signal regarding the conduct of the entities from which it sources materials or models for its system, also plays a primary role in regulating them.

3.2 Product Manufacturer.

The figure of the Product Manufacturer represents another Operator under the AI Act when it proceeds to place on the market (i.e., the first making available in the Union Market) or put into service (i.e., direct supply to the deployer for first use or own use in the Union for the intended purpose) an AI system together with its product. This makes it a figure very close to that of the Provider.

It is worth noting that this figure is understood as the one who manufactures a product in which the safety component is represented by an AI system, although it is generally referred to as a commercial entity responsible for the production of final products – falling within the scope of Union harmonization legislation – that incorporate a high-risk AI system as a safety component. Considering that such high-risk systems used as safety components could negatively impact people's health and safety, even while aiming to facilitate the free movement of products in the internal market, there remained the need to prevent and mitigate the safety risks of such products. In light of these needs, on the one hand, the Regulation has classified all AI systems that serve

[13] Parliament and Council Directive 2019/1937/EU of 23 October 2019 on the protection of persons who report breaches of Union law [2019] OJ L305/17. <https://eur-lex.europa.eu/eli/dir/2019/1937/oj/eng>

as safety components of products (or that are products themselves) as 'high-risk systems' when the affected product (such as machinery, toys, pressure equipment, or medical devices) is subject to a conformity assessment by a third-party conformity assessment body, as provided in Recital 50. On the other hand, Recital 51 states that products based on the same system classified as 'high-risk' are not necessarily required to be classified as 'high-risk' under the harmonization legislation applicable to products. At the same time, AI systems linked to non-high-risk products, while not required to comply with the requirements set for a high-risk AI system, must still be safe when placed on the market or put into service.

The Regulation further stipulates that a manufacturer may assume the role of Provider and thus be subject to the corresponding obligations imposed on the latter, as a result of the responsibility along the AI value chain anticipated in Recital 87 and detailed in Article 25(3) of the AI Act. This occurs when, in the case of high-risk AI systems serving as safety components of a product, the AI system is placed on the market together with the product under the manufacturer's name or trademark, or when the product is initially put into service, followed later by the AI system under the same manufacturer's name or trademark.

3.3 Deployer.

The deployer represents a newly defined figure, described as the 'natural or legal person, public authority, agency or other body using an AI system under its authority', according to Article 3, paragraph 1, point 4, AI Act. For this subject, the centrality of professional activity is of particular relevance, as the Regulation explicitly excludes from this subjective category those who use the AI system «in the course of a non-professional personal activity».

Once again, the need to provide AI literacy remains essential to equip the deployer with the necessary knowledge to make informed decisions regarding AI systems. Various obligations are placed on the deployer, especially in the case of high-risk systems, as outlined in Article 26. (Fig. 5)

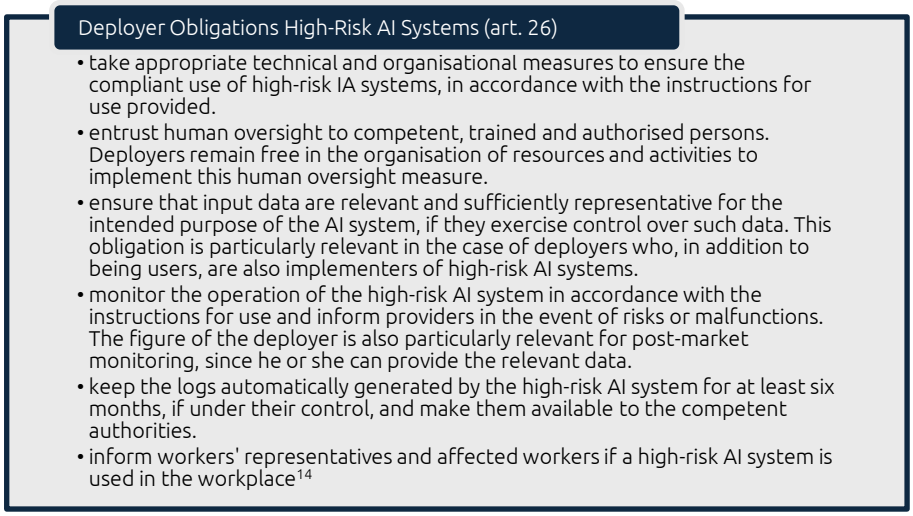


Fig. 5

Additionally, for certain subjects, further obligations are added, including the requirement for an impact assessment on fundamental rights of high-risk AI systems. This activity, as will be discussed later, is limited to the subjects indicated in paragraph 1 of Article 27 AI Act, and these subjects, at the time of the first use of a high-risk AI system, are required to describe the process and frequency of use, the categories of individuals involved, and a specific description of the risks to which these individuals will be exposed, with particular attention to monitoring measures and those aimed at mitigating them. It is clear that in this case, the European legislator, not being in a position to know the specific risks associated with a given AI system, leaves the individual deployer with the responsibility to assess them through a self-assessment of their own safety framework and requires communication of the results to the market surveillance authority.

Moreover, still within the scope of high-risk systems, in the case of deployers of emotion recognition systems or biometric categorization systems, they are required to inform individuals exposed to the functioning of the system, unless these systems have been authorized by law for the purpose of investigating, preventing, or detecting crimes.

On the other hand, in the context of generative AI, with the creation of so-called deepfakes [15] and systems that generate or manipulate texts for information purposes, the deployer is required to disclose that the text was created or modified by AI. All these information obligations placed on the deployer also have a precise temporal requirement – 'at the latest at the time of the first interaction or exposure' – and must be accessible.[16] It is clear that these provisions were designed to protect individuals exposed to synthetic content and to make it clear to them that they are interacting with a machine, not a human being.

In conclusion, it should be emphasized that in this case, the mechanism protecting the market introduced by the European legislator is of particular importance: the system of liability along the AI value chain. If certain subjects, including deployers – but also importers, distributors, or other third parties – engage in particularly impactful actions on the system, such as making substantial changes to the system or modifying the purpose of a non-high-risk system, precise consequences are triggered: in such cases, the initial Provider will no longer be considered as such for that system (although their obligation to cooperate for information or support remains), and the deployer (or anyone else acting on their behalf) must also comply with the obligations placed on the Supplier. Through this solution, the AI Act has established a key principle:

[14] For an initial assessment of the organizational impacts of the new regulations concerning the obligations for the user (deployer) of high-risk AI systems when used by an employer, see L Barbieri, A Rossetti 'Sistemi d'intelligenza artificiale ad alto rischio e obblighi del datore di lavoro' 2024 Norme&TributiPlus. The authors, in illustrating in practical terms the provisions of both Article 26 and Recital 92 of the AI Act, not only analyze the obligations imposed on the 4.0 employer, but also focus on the principle of transparency and information in favor of the worker, as well as the need for human oversight and a dedicated management system to 'counter, prevent, or minimize risks to health, safety, or fundamental rights that may arise'.

[15] These are images, audio, or videos generated or manipulated by AI to resemble real persons, objects, places, events, or entities, which may appear deceptively authentic or true.

[16] Art. 50, par. 5, AI Act.

whenever a system is modified, the entity making the modification is required to ensure its compliance.

3.4 Other Operators.

Alongside the aforementioned subjects, additional figures emerge: these are figures that are, in some cases, optional, but in others, essential in the so-called AI supply chain. The Authorised Representative and the Importer are operators provided for by the regulation only when the Provider of a high-risk AI system or a general-purpose AI model is a non-EU Provider, established in a third country.

In order to ensure that, under any circumstances, a person established in the Union can provide the authorities with all the necessary information on the compliance of an AI system, the AI Act has established the figure of the Authorised Representative. The Supplier of high-risk AI systems established in a third country must appoint this Representative through a written mandate before making the system available. The Representative will then be authorized to perform the tasks specified in the Supplier's mandate and indicated in Article 22 of the Regulation for high-risk AI systems. Similarly, this obligation is placed on suppliers of general-purpose AI models, except for suppliers of general-purpose AI models released under a free and open-source license that do not present systemic risks under Article 54 of the Regulation, although with the provision of different specific tasks entrusted to the Representative.

The figure of the Importer becomes relevant when it places a high-risk AI system on the EU market that was purchased from a non-EU Supplier. Given its position, the Importer is tasked with ensuring that the system complies with the Regulation, taking on the obligations set out in Article 23, verifying both the presence of the Authorised Representative appointed by the non-EU Proveder and the technical documentation, as well as the compliance with the conformity assessment procedures and the application of the CE marking. After a positive outcome from these checks, the Importer will proceed to indicate its name or trademark on the packaging or accompanying document of the system, ensuring that its transportation does not compromise its conformity.

The distributor, in turn, represents the intermediary between the Provider or Importer and the third party to whom it will resell the AI system. In an era of segmented production processes, the figure of the distributor is becoming increasingly important, which is why EU lawmakers have foreseen it as the subject that must carry out a series of checks before making a high-risk AI system available on the market – such as ensuring compliance with obligations by the Provider and the Importer – and must manage any non-conformities.

Lastly, it is important to note that, under certain conditions, any Distributor, Importer, Deployer, or other third-party subject could be considered a Provider of a high-risk AI system and, therefore, assume all the related obligations.

3.5 Interposing Subjects.

Among the Interposing Subjects, it is important to highlight primarily the conformity assessment body, which is the entity that, on behalf of third parties, performs conformity assessments, including testing, certification, and inspections. This body is the third party that verifies the conformity of high-risk AI systems to the Regulation before their market placement or deployment, with the aim of ensuring a high level of reliability of the systems. In order to assess the quality management system, the technical documentation, and the related monitoring to issue the corresponding certificates in accordance with Annex VII, the conformity assessment body must assume the role of a notified body. This qualification is granted by the notifying authority of each Member State upon the body's application. The AI Act further identifies sectoral groups for coordinating notified bodies and recognizes the conformity assessment bodies of third countries that must meet the requirements set by the AI Act.

Next, it is necessary to identify the so-called 'affected persons' that are located in the Union under Article 2, paragraph 1, letter g) of the Regulation. In this case, it is undoubtedly implied that this refers to individuals who participate in the real-world testing conditions specified, to whom specific provisions are dedicated under Articles 60 and 61, such as the requirement to obtain their informed consent before participating in the tests, after being 'duly informed' of the objectives of the tests, their duration, rights and guarantees, and many other essential details.

In conclusion, particular attention is paid to SMEs, including startups, in order to continue supporting the innovation process necessary in the field of AI.

3.6 Institutional Actors.

In identifying the subjective scope of application of the AI Act, it is not possible to leave out the Institutional Actors. These, involved in the first instance and in a more active way - since these actors must set up the organisational structures, must carry out the controls, and must ensure even more the performance of their functions - are numerous. Notwithstanding the presence of the European Commission, which approves, investigates, coordinates and modifies, and the Member States, which are called upon to implement, authorise and designate, there are many authorities involved. One refers to the typical law enforcement authorities, market surveillance authorities and authorities protecingt fundamental rights, as well as newly created authorities such as notifying authorities, which have been assigned the task of designating, notifying and monitoring conformity assessment bodies. In addition, there are additional actors:

- AI Office, i.e. that institution that enables the Commission to develop expertise and capacity in the field of AI;
- European Artificial Intelligence Board, i.e. a committee consisting of one representative from each Member State and with the non-voting participation of the EDPB and the AI Office;

- Advisory Forum, to provide technical advice to the Committee and the Commission;
- Scientific panel of independent experts, to support the implementing measures of the Regulation.

4. Objective Scope.

In light of the above, risk has been the evaluation parameter chosen by the legislator, as the need emerged to calibrate precautions and obligations for certain subjects along the AI value chain based on the type of risk that a given category of AI systems might pose. Although it has recently been criticized[17], the risk-based approach has allowed for the establishment of rules tailored to the scope of the risks generated by AI systems, with a broadly precautionary approach in mind. This has translated into a series of mechanisms for the review, supervision, and management of AI systems (continuous monitoring, audits, a multilayered governance system at the European and national levels with authorities and notified bodies, as well as provisions for holding entities providing critical infrastructures accountable) which should provide guarantees against non-compliance applications.

From the framework identified by the legislator, four levels of risk have been identified, depending on the context of use, the potential impacts of the technologies, the application sector and the interested parties. In fact, all other AI applications fall within the zero-risk area not regulated by the AI Act (Fig. 4):

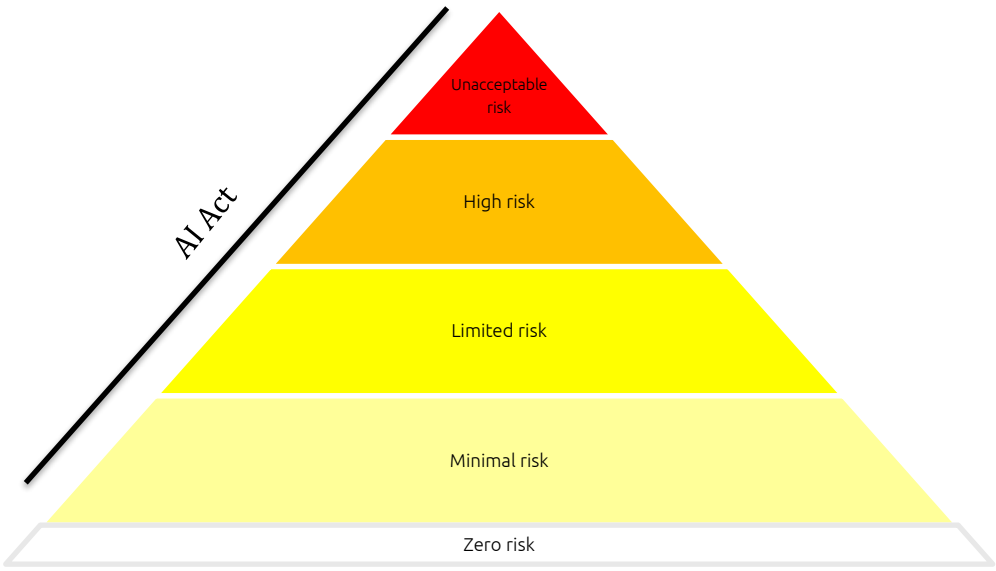


Fig. 4

The Regulation does not govern all AI applications, but only a minority of

[17] F Porzio, 'AI Act, il rischio di danneggiare l'innovazione in Ue' (*Agendadigitale*, 10 May 2024) <https://www.agendadigitale.eu/industry-4-0/ai-act-ecco-come-mette-a-rischio-linnovazione-in-ue-e-con-quali-rischi/>.

them, and from the text emerges a primary distinction between AI systems and AI models. This distinction also allows for overcoming the initial definition of the AI Act as a 'product regulation', as understood at the beginning of the legislative process. In fact, by regulating AI models – those fundamental elements related to the more intrinsic technology that enables the development of an AI system – the Regulation foresightedly governs even the technology underlying the product-system.

Technically, the Regulation, in addition to regulating Artificial Narrow Intelligence (ANI) – the so-called AI systems, or "weak" AI, trained on well-labeled data sets and capable of performing specific tasks in predefined environments – also provides provisions for what will soon be at the core of Artificial General Intelligence (AGI) – the so-called AI systems for general purposes, or "strong" AI, capable of performing a variety of intelligent tasks and adapting to a wide range of situations.

4.1 AI Systems.

With regard to AI systems, the Regulation provides for a differentiated regime of obligations corresponding to the level of risk of the individual system (Fig. 5):

Level of risk	Discipline
Unacceptable risk – Prohibited AI practices	Prohibition regime[except expressly exempted(*ex* art. 5 AI Act)]
High risk – High-risk AI systems	Specific requirements and obligations (*ex* artt. 6 e ss.)
Minimal risk – Certain AI systems	- Transparency Obligations (*ex* art. 50) - Burden of codes of good practice (Art. 56) - Burden of enforcing codes of conduct (*ex* art. 95)

Fig. 5

4.1.1 Prohibited AI practices.

Regarding AI systems with an unacceptable risk, the Regulation establishes a regime of prohibition with the so-called prohibited practices under Article 5 of the AI Act. These include the use of 'subliminal techniques' capable of compromising the ability to make informed decisions, and certain systems that impact individuals' behavior without their awareness, particularly concerning vulnerable groups such as minors, the elderly, and people with physical and mental disabilities. The use of AI systems by public authorities to qualitatively assess individuals is also prohibited, such as the so-called social ranking, which involves assigning scores that lead to measures or actions that impact individuals in an unjustified or disproportionate manner relative to the severity of their behavior. Additionally, the Regulation prohibits AI systems used to assess or predict the likelihood of committing crimes, those that create or expand facial recognition databases through non-targeted scraping of facial images from the internet or CCTV, systems used to recognize the emotions of

individuals in the workplace or educational settings, and those that categorize biometric data to infer race, political beliefs, and similar characteristics.

Furthermore, AI systems used for surveillance of individuals in public spaces via remote real-time biometric identification are prohibited, with a set of exceptions provided for the search of crime victims, missing minors, or for public safety and defense purposes, or for the search of individuals suspected of committing crimes punishable by at least four years of imprisonment. The exceptional use of such systems is subject to a motivated request and prior authorization by the relevant independent authority of the Member State concerned, except in urgent cases where authorization may be sought *ex post*. Each Member State is then responsible for establishing the rules, cases, and limits regarding facial recognition, given its direct impacts on the fundamental rights and freedoms of individuals (under Article 5, paragraph 1, letter d, paragraphs 2 and 4).

It is important to note that non-compliance with the prohibition of the practices outlined in Article 5 is subject to administrative financial penalties under Article 99.

Although the prohibitions set out in this article entered into force on 2 February, the Commission is empowered, pursuant to Article 96, to issue guidance on the practical implementation of the Regulation. To date, the Guidelines on Prohibited AI Practices[18] were published on 4 February 2025, and the Guidelines on the Definition of an AI System[19] under Article 3(1) on 6 February 2025.
The aim of the Guidelines on Prohibited AI Practices is to enhance legal certainty and provide clarity on the Commission's interpretation to ensure the consistent, effective, and uniform application of the prohibitions under Article 5 by competent authorities, providers, and deployers of AI systems.

These guidelines are not binding, and the application of Article 5 will in any case require an assessment based on the specific circumstances of each case. The legal basis of the European Regulation is provided by Articles 114 (internal market) and 16 TFEU (data protection). In particular, Article 114 concerns the adoption of measures relating to the approximation of the legislative, regulatory, and administrative provisions of the Member States that aim at the establishment and functioning of the internal market. Article 16 establishes the right of every person to the protection of personal data concerning them. It refers in particular to the prohibitions on the processing of personal data derived from the use of remote biometric identification systems, the use of biometric categorisation systems, and individual risk assessment for law enforcement purposes. All other prohibitions under Article 5 of the AI Act are

[18] Communication to the Commission "Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (*AI Act*)", Brussels, 4.2.2025 C (2025) 884 final. <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act>.

[19] (Annex to the) Communication to the Commission "Approval of the content of the draft Communication from the Commission - Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (*AI Act*)". <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application>.

legally based on Article 114.

As set out in the Commission's Guidelines, Article 5 of the AI Act prohibits the placing on the EU market, putting into service, or use of certain AI systems for practices of manipulation, exploitation, social control, or surveillance, which by their nature violate fundamental rights and the values of the Union. In particular, such systems infringe human dignity, freedom, equality, democracy, and the rule of law, as well as fundamental rights enshrined in the Charter of Fundamental Rights of the EU (CFREU), including, among others, the right to non-discrimination (Article 21), equality before the law (Article 20), data protection (Article 8), respect for private and family life (Article 7), the rights of the child (Article 24), freedom of expression (Article 11), the right to an effective remedy and to a fair trial (Article 47), and the presumption of innocence and the right of defence (Article 48).

The AI Act does not provide a precise definition of subliminal techniques, but these are understood as techniques that impair awareness and influence the behaviour of individuals interacting with the AI system. Practical examples include subliminal visual or auditory messages, hidden images embedded among other visual content that are unconsciously processed by individuals and influence their behaviour. Another example is diversion, where the AI system draws the user's attention to specific content in order to conceal other content, exploiting cognitive biases and vulnerabilities in attention (e.g., influencing consumer purchasing choices). The manipulation of time perception, which can induce anxiety and dependence, is another such example.

Deliberately manipulative techniques are designed to influence, alter, or control individual behaviour in ways that undermine autonomy and freedom of choice. A practical example is sensory manipulation, where AI systems provide an audiovisual background that alters mood, causing anxiety and stress, and influencing user behaviour to the point of causing significant harm. Such manipulative techniques can be personalised when AI systems generate persuasive messages based on individuals' personal data or exploit their specific vulnerabilities. Some AI systems may not have been developed with manipulative intent, yet may still apply these techniques based on the datasets on which they were trained. For this prohibition to apply, all the conditions set out in the provision must be met cumulatively.

Deceptive techniques may consist in the presentation of false or misleading information with the aim, or even the effect, of deceiving individuals and influencing their behaviour in ways that compromise their autonomy, decision-making process, and freedom of choice. Practical examples include intelligent chatbots that impersonate an acquaintance or relative using synthetic voice in order to carry out scams or cause significant harm; AI systems that learn to identify moments of evaluation and control and 'hide' dangerous behaviours for as long as necessary.

Significant harm may include stress, anxiety, insomnia, addiction, self-harm, exacerbation of stalking, cyberbullying, harassment, incitement to hatred and violence against certain minorities, or psychological harm resulting from identification with victimised groups through deepfake techniques.

With regard to the placing on the market, putting into service, or use of AI

systems that exploit the vulnerabilities of individuals or specific groups, the AI Act does not provide a precise definition of vulnerability – a concept that can encompass a wide range of types, including cognitive, emotional, physical, and other forms of fragility that may affect an individual's or a group's ability to make informed decisions or, more generally, their behaviour. However, the Regulation expressly prohibits AI practices that may exploit vulnerabilities arising from certain factors such as age, disability, or specific socio-economic situations, which generally make it more difficult to recognise or resist manipulative or distortive AI practices.

For this prohibition to apply, all the conditions set out must be met simultaneously, including, in particular, a plausible causal link between the exploitation, the distortion of the behaviour of the person interacting with the AI system, and the significant harm that has resulted or can reasonably be expected to result.

With regard to age, the vulnerable groups identified include both minors and older people. Minors are particularly sensitive to manipulative techniques because their stage of development limits their ability to critically assess what is real and the intentions behind AI-driven interactions. Owing to their general immaturity, minors are especially susceptible to the risk of developing forms of dependency on artificial agents and AI-based applications.

Examples include smart toys that encourage children to engage in dangerous actions by turning them into challenges to be won in exchange for digital rewards; certain video games capable of analysing children's behaviour and preferences and, on that basis, creating personalised digital rewards that may induce addiction, stress, and other significant and long-term consequences such as insomnia, myopia, concentration difficulties, social and relational difficulties, eating disorders, depression, and self-harm. AI systems may also be used for the generation of child sexual abuse material and the development of strategies for grooming minors. However, not all types of smart toys, video games, or educational applications are generally prohibited – only those that exploit minors' vulnerabilities and may cause them significant harm.

Older people may experience reduced cognitive abilities (not necessarily due to dementia) and may have difficulty using and interacting with AI systems because of limited digital skills, making them particularly susceptible to scams and manipulative marketing strategies. The harm is often financial and psychological, and is frequently exacerbated by conditions of isolation. Examples include AI systems that exploit the vulnerabilities of older people by advertising expensive medical treatments, unnecessary insurance, or deceptive investment schemes that lead to loss of savings, increased debt, and stress. These systems often take advantage of the specific socio-economic situations in which older people find themselves. Other examples include caregiver robots that may exploit the fragility of those they care for and coerce them into activities against their will, worsening their mental and physical health. Not all such applications are prohibited – only those that exploit vulnerabilities and cause or are likely to cause significant harm.

Article 5 of the AI Act also aims to prevent the risk that AI systems may exploit the limitations and vulnerabilities of persons with disabilities and to protect them from harmful influences and manipulative practices. The concept

of disability encompasses a wide range of long-term physical, mental, intellectual, and sensory impairments which, in interaction with various barriers, hinder the full and effective participation of individuals in society on an equal basis with others (see UN Convention on the Rights of Persons with Disabilities). Persons with disabilities may be more easily influenced and exploited than others. Examples include: a therapeutic chatbot intended to provide daily support and strategies for persons with mental disabilities but which may exploit their limited intellectual abilities to influence them into purchasing expensive devices and medical products that promise unrealistic benefits, or condition them to the point of adopting behaviours that are dangerous to themselves or others. Certain AI systems may identify women and girls with disabilities online in order to use them to generate pornographic content or to develop more effective grooming practices. With regard to specific social and economic situations, the term 'specific' does not refer so much to an individual characteristic as to a particular legal status or membership of a specific vulnerable social or economic group.

According to Recital 29, this group may include persons living in conditions of extreme poverty, as well as members of ethnic and religious minorities. The category covers not only relatively stable and long-term situations, but also temporary circumstances such as short-term unemployment, over-indebtedness, and migrant status. Persons in such conditions are typically particularly vulnerable, have few resources, and possess limited digital skills compared to the rest of the population, which makes it more difficult for them to recognise and resist manipulative AI practices. An example would be a predictive AI algorithm used in targeted advertising strategies for predatory financial products aimed at people living in low-income areas who are in precarious financial situations, exploiting their desperation.

AI systems that exploit socio-economic data could disproportionately affect ethnic minorities or individuals belonging to a particular ethnicity, exacerbating existing disparities and contributing to forms of systemic discrimination and even to the exclusion of individuals from those groups. However, the AI Act does not apply to systems that target consumers based on a wide range of variables that are not directly linked to vulnerable groups in specific socio-economic situations (such as smartphone use, the size of the city in which they live, or their travel destinations and frequency). Certain AI chatbots, for example, could push migrants or refugees – who often lack a clearly defined legal status and stable socio-economic conditions – towards extremist positions and incite hatred and violence against certain population groups by exploiting their particular sensitivity to specific content and fear-based narratives.

4.1.2 High-risk AI systems.

The type of systems most thoroughly regulated are those categorized as "high-risk". A system is classified as such under Article 6 of the AI Act if two conditions are met:

- It is intended to be used as a component of safety in a product, or it is itself a product, in both cases governed by the regulations in Annex I.
- The AI system (whether as a safety component or as a product) is subject

to third-party conformity assessment for market placement or deployment, according to the regulations in Annex I.

The conformity assessment (under Article 43) referred to in Article 6, paragraph 1, letter b, relates to the system's compliance with the requirements set out in Articles 8 and following: the presence of a risk management system (under Article 9) and appropriate technical documentation (under Article 11), the need for human oversight (under Article 14) and specific transparency obligations for deployers (under Article 13), as well as inherent characteristics of the system such as its accuracy, robustness, and cybersecurity (under Article 15), data quality, and data governance (under Article 10).

At the same time, AI systems listed in Annex III are also considered high-risk, namely those operating in the areas of:

- Biometric identification
- Management and maintenance of critical infrastructure
- Access to education or evaluation of educational pathways
- Selection of candidates or management of employment relationships
- Access to essential public or private services
- Systems used by law enforcement, if permitted under Union and national law
- Control of migration phenomena, asylum requests, and border control, if authorized by Union and national law
- Administration of justice and democratic processes

However, AI systems listed in Annex III may be considered not high-risk if they do not present a significant risk of harm to health, safety, or fundamental rights, even in the sense of not materially influencing the outcome of the decision-making process. It should be noted that if a system in Annex III involves the profiling of individuals, it is always considered high-risk.

4.2 General Purpose AI Models.

Regarding AGI, the AI Act regulates GenAI starting from its foundations, namely the so-called general-purpose models (GPAI - General Purpose AI models). Given their ability to serve multiple purposes, fundamental models have become one of the key topics in European AI regulation, particularly following the advent of Chat-GPT, and have been categorized into two groups:

- Systemic models
- Generic models

Both types of models are capable of performing a range of operations, but while the latter are considered to have greater safety, the former are characterized by their so-called high impact capability. For this reason, increasing obligations have been foreseen between generic and systemic GPAI models, with more stringent requirements for general-purpose AI models with systemic risk, as these may expose the European market and the foundational values of the EU to so-called systemic risks (Fig. 6).

There are numerous references to the regulation of the DSA (Digital Services Act), both in terms of the scale of AI and the definition of systemic risk. Indeed, there is a presumption of "high impact capability" when the total computing power used for training the AI model exceeds 10^25 FLOPs, although this can

be surpassed by the European Commission. The adopted approach grants the Commission significant discretionary power, both within the system for defining and governing Artificial Intelligence, including the designation of models as "systemically risky" and modifying the definition under changing conditions, as well as overseeing all general-purpose AI models, as the Commission holds exclusive competence under Article 88 of the AI Act.

Level of risk	Discipline
Systemic risk - AI models for general purpose systemic risk (Systemic models)	- Particularly detailed transparency obligations (Art. 53) - More intrusive obligations (Art. 55) - Burden of codes of good practice (Art. 56)
Limited risk – General purpose AI models (Generic models)	- Particularly detailed transparency obligations (Art. 53) - Burden of codes of good practice (Art. 56)
Minimal Risk - Certain General Purpose AI Systems	- Transparency obligations (Art. 50) - Burden of codes of good practice (Art. 56) - Burden of codes of conduct (Art. 95)

Fig. 6

4.2.1 Systemic AI Models.

For systemic GPAI models, not only are the same detailed transparency obligations applied as for generic models, but also an additional set of requirements due to the identified criticalities. These further obligations are set forth in Article 55 and carry a more intrusive nature:

- Compliance evaluation of the model with standardized protocols reflecting the state of the art, including the documentation of adversarial tests to identify and mitigate systemic risks;
- Assessment and mitigation of potential systemic risks at the Union level, including related causes such as the development, market placement, or use of general-purpose AI models with systemic risk;
- Tracking, documenting, and reporting – without undue delay – to the AI Office and, when appropriate, to the relevant national authorities – pertinent information about serious incidents and potential corrective measures to address them;
- Ensuring an adequate level of cybersecurity protection regarding the model and its physical infrastructure.

To demonstrate compliance with the above obligations, suppliers of GPAI models can rely on codes of good practices under Article 56.

4.2.2 Generic AI Models.

In regard to the regulation of generic GPAI models, detailed transparency obligations are established under Article 53, such as the requirement to prepare technical documentation that makes the model's functioning understandable, including the training and testing process. This documentation must be provided to those intending to integrate the model into their AI systems and, upon request, to the AI Office. At the same time, a policy must be in place to protect intellectual property rights, and a summary

of the content used during training must be made publicly available. The adoption of codes of good practices under Article 56 is also required, just as for systemic models.

Particular attention is given to GPAI models released under open-source licenses, which allow access, use, modification, and distribution of the model, with parameters – including weights, model architecture, and usage information – made public. In this case, providers are subject to fewer obligations unless the models present systemic risks.

Generative Artificial Intelligences (GenAI) are revolutionizing how people live and work. While human oversight has always been deemed essential in the AI field, it takes on a primary role in the context of information: technology renews information, but risks multiply, making it crucial that automatically generated content is clearly recognisable and not confused with conventionally generated content.[20]

In the legal field, this shift has led legal practitioners – including judges, lawyers, and legal scholars – to acquire digital skills for using new tools. These tools can now be delegated not only for drafting legal documents but even part of decision-making processes (Decision Support Systems, DSS). One example of a GenAI system is Microsoft's CoPilot chatbot, which can be used to draft legal act drafts using document models from the personal computer's memory or by retrieving large amounts of text from the web. While this is not a system specialized in the legal field, it reflects a new configuration of legal information, necessitating a change in cultural paradigms. The true revolution of these new intelligent systems lies in their potential to support legal reasoning. An Italian experiment in this regard is Giurimatrix, an AI system specialized for the legal sector developed by Professor Filippelli and Lawyer Viola.

The use of these programs, however, must be accompanied by an awareness of their fallibility and the need for digital literacy among professionals in all fields, not just in law. This awareness was achieved in the legal sector during the G7 for Lawyers organized by the CNF on April 16 2024.[21] For ensuring AI reliability in these contexts, the following ten actions are deemed indispensable:

1. Create a task force to conduct a survey to collect data on the state of the art regarding AI applications in law firms and their purposes.
2. Assess compliance with privacy regulations and the level of data digitization.
3. Identify the normative principles (e.g., ethical code) considered essential for the legal profession's identity.
4. Develop an assessment tool for AI tools based on the above principles.
5. Certify AI tools used in legal practice under the EU AI Regulation, with particular attention to the FRIA (Fundamental Rights Impact Assessment) procedure (Article 27 and Annex III).
6. Adapt all procedural acts (including those of judges) to the criteria of Ministry of Justice Decree no. 110 of 2023, enabling their use in training AI tools.

[20] G Corasaniti, *Tecnologie intelligenti: Rischi e regole* (Mondadori Università 2023) 38.
[21] L Gatt, '10 azioni per sviluppare la IA in Italia' (2024) 2 La Nuova Procedura Civile.

7. Standardize legal procedural acts and information for AI training: configuring uniform legal data that can be understood by different systems.
8. Provide intensive and thorough training for lawyers on the conscious and critical use of AI tools.
9. Create a basic AI tool shared with various functions (both drafting and legal reasoning – "Algorithm 12").
10. Ensure interoperability of all AI tools in the legal field with those used by judicial bodies and public administrations.

On a broader scale, GenAI raises complex regulatory issues due to its transition from processing linguistic models to processing world models. Whereas GenAI previously only understood syntax and grammar, this new alien manufactured[22] is now making the leap toward understanding meaning. This introduces new demands for legal mediation, prompting legislators, particularly in the EU, to consider whether «regulation or liability»[23] is the most appropriate protective tool, with further training needed for operators downstream.[24] In this context, it becomes necessary to speak of the intelligent use of smart technologies as the best means to detect and counteract the drifts of planned marginalisation.[25]

4.2.3 Certain AI systems.

With the advent of general-purpose AI models, the AI Act also regulates cases where these models form the basis of AI systems. This involves the so-called 'general-purpose AI systems', which, being based on general-purpose AI models, have the capability to pursue various objectives.

In light of this, for certain general-purpose AI systems and basic AI systems (i.e., non 'general-purpose'), the Union legislator has identified transparency obligations. According to Article 50:

- When basic AI systems interact directly with natural persons, they must be designed and developed in such a way that individuals are informed they are interacting with AI. The only exception to this rule is when it is clearly evident, taking into account the moment and context, that the interaction involves an AI system.
- When AI systems – both "general-purpose" and basic – generate synthetic content (audio, video, images, or texts created or modified by AI), the outputs must be marked in a readable format and identified as artificially generated or manipulated. The exception applies to AI systems that perform standard editing functions or do not substantially alter the input data provided by the deployer or its semantics, or if authorized by law for detecting, preventing, investigating, or prosecuting crimes.

[22] U Ruffolo, 'Intelligenza Artificiale generativa e mediazione giuridica', conference speech at Stati Generali del Diritto di Internet e della Intelligenza Artificiale on the 28, 29, 30 november 2024. <https://www.youtube.com/watch?v=aKylH1Zm2Hw>.
[23] Ruffolo (n 22).
[24] Corasaniti (n 20) 42.
[25] Corasaniti (n 20) 39. The author makes a warning: it is necessary for law, economics and technology to come together in order to identify the best actions to exploit the full potential of information and to counter illegal activities.

- In reference to emotion recognition systems or biometric categorization systems, individuals must be informed of the system's functioning, and these systems must handle personal data in compliance with Regulation (EU) 2016/679, [26] Regulation (EU) 2018/1725, [27] and Directive (EU) 2016/680,[28] depending on the case.
- For AI systems generating or manipulating deep fakes, it is necessary to disclose that the content has been generated or manipulated artificially (digital watermark). The exception applies when the use is authorized by law for detecting, preventing, investigating, or prosecuting crimes.
- For AI systems generating or manipulating texts published to inform the public on matters of public interest, it must be disclosed that the text has been generated or manipulated artificially. The exception applies when the use is authorized by law for detecting, preventing, investigating, or prosecuting crimes, or if the AI-generated content has undergone human review or editorial oversight, and a natural or legal person holds editorial responsibility for the publication of the content.

Additionally, these systems have the burden of adopting codes of conduct (art. 95) and code of good practices (art. 56).

To conclude, the final risk pyramid of the AI Act, when assessed in its entirety, appears as follows (Fig. 7).

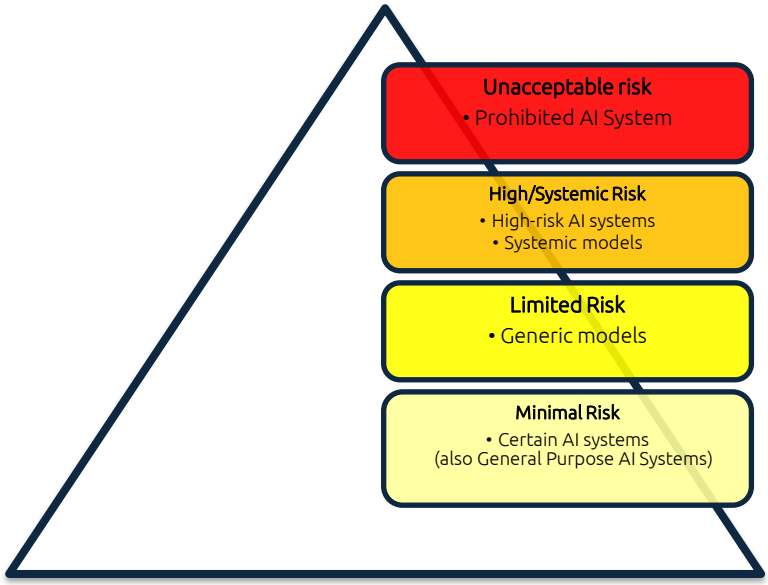


Fig. 7

[26] Parliament and Council Regulation 2016/679/UE of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation - GDPR) [2016] OJ L119/1. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679>.

[27] Parliament and Council Regulation 2018/1725/UE of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC [2018] OJ L295/39. <https://eur-lex.europa.eu/eli/reg/2018/1725/oj/eng>.

[28] Parliament and Council Directive 2016/680/UE of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA [2016] OJ L11/89. <https://eur-lex.europa.eu/eli/dir/2016/680/oj/eng>.

5. FRIA - Fundamental Rights Impact Assessment.

In the field of AI, particular importance is attached to the Fundamental Rights Impact Assessment (FRIA) for high-risk AI systems under Article 27 of the AI Act. This activity involves measuring the impact of high-risk AI systems on fundamental rights with a view to ex-ante protection, following a risk-based approach. By assessing the potential risks of harm arising from the application of AI, it functions as a true tool of accountability.[29]

As previously mentioned, this activity is only mandatory for certain deployers: for those of high-risk systems under Article 6, Paragraph 2 of the AI Act – that is, the AI systems listed in Annex III, except those intended for use in the sectors outlined in point 2 – for public organizations or private entities providing public services, and for deployers of high-risk AI systems listed in Annex III, point 5, letters b) and c) (i.e., those related to creditworthiness assessments and the determination of risks and prices in the life and health insurance sectors).

Preliminary elements to the Fundamental Rights Impact Assessment (FRIA) are undoubtedly those AI impact measurement models developed by supranational institutions. One example is the ALTAI (Assessment List for Trustworthy Artificial Intelligence), a tool that, based on the Ethical Guidelines for Trustworthy Artificial Intelligence presented by the High-Level Expert Group on AI (AI HLEG) to the European Commission in April 2019, supports the implementation of the key requirements outlined in those guidelines. From the reading of the Ethical Guidelines[30], it is evident that the necessary components to define a trustworthy AI system are three:

- Legality (the system must comply with all applicable laws and regulations);
- Ethicality (the system must adhere to shared ethical principles and values);
- Robustness (both technical and social, to minimize the risk of unintentional harm).

At the same time, seven requirements have been highlighted that AI systems must meet throughout their lifecycle to be considered trustworthy:

- Human intervention and oversight (including the respect of fundamental rights, human intervention, and human oversight);
- Technical robustness and security (including resilience to attacks, emergency planning, general security, accuracy, reliability, and reproducibility);
- Data privacy and governance (including data confidentiality, quality, integrity, and access);
- Transparency (including traceability, explainability, and communication);
- Diversity, non-discrimination, and fairness (including the prevention of unfair biases, accessibility, universal design, and stakeholder

[29] L Gatt and others, 'FRIA Implementation Model According to the AI Act' (2024) IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), 1224-1229. <https://ieeexplore.ieee.org/document/10796624>.

[30] AI HLEG, 'Ethics guidelines for trustworthy AI' (2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

participation);
- Social and environmental well-being (including sustainability, environmental respect, social impact, society, democracy, and collective interests);
- Accountability (including verifiability, minimization of negative effects, and reporting, trade-offs, and remedies).

Operationally, through the Assessment List for Trustworthy AI (ALTAI), the principles of AI are translated into an accessible and dynamic checklist that guides AI developers and distributors in the practical implementation of these principles, creating a self-assessment tool.[31] In the same vein and with the same self-assessment approach, the European Commission for the Efficiency of Justice of the Council of Europe (CEPEJ) developed in 2023 an evaluation tool for the operationalization of the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment. This self-assessment tool,[32] which refers to the principles outlined in the Ethical Charter adopted in December 2018,[33] proposes a questionnaire based on 29 questions to provide clear, practical guidance on how to concretely apply the Charter. Furthermore, this tool seems to complement the Human Rights, Democracy, and Rule of Law Impact Assessment (HUDERIA) presented by the Committee on Artificial Intelligence (CAI) of the Council of Europe during its 12th plenary session in November 2024.[34]

At the same time, significant contributions to the development of these tools have also come from other actors: such as Mantelero's studies,[35] ALIGNER Fundamental Rights Impact Assessment (AFRIA),[36] the FRAIA tool developed by the Dutch Government,[37] research presented by the Intesa Sanpaolo team,[38] or proposed solutions for improving AI fairness.[39]

Despite being based on excellent premises, these tools are not yet fully

---

[31] AI HLEG, 'Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment' (2020) <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>.

[32] CEPEJ, 'Assessment Tool for the Operationalisation of the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment' (2023) <https://rm.coe.int/cepej-2023-16final-operationalisation-ai-ethical-charter-en/1680adcc9c>.

[33] CEPEJ, 'European ethical Charter on the use of Artificial Intelligence in judicial system and their environment' (2018) <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

[34] CAI, 'Methodology for the risk and impact assessment of artificial intelligence systems from the point of view of human rights, democracy and the rule of law (HUDERIA methodology)' (2024) <https://rm.coe.int/cai-2024-16rev2-methodology-for-the-risk-and-impact-assessment-of-arti/1680b2a09f>.

[35] A Mantelero, 'The Fundamental Rights Impact Assessment (FRIA) in the AI Act: roots, legal obligations and key elements for a model template' (2024) 54(106020) CLSR <https://www.sciencedirect.com/science/article/pii/S0267364924000864>.

[36] Aligner, 'ALIGNER Fundamental Rights Impact Assessment (AFRIA)' (2020) <https://aligner-h2020.eu/>.

[37] Government of Netherlands, 'Impact Assessment Fundamental Rights and Algorithms' (2022) <https://www.government.nl/documents/reports/2022/03/31/impact-assessment-fundamental-rights-and-algorithms>.

[38] S Bertaina and others, 'Fundamental Rights and Artificial Intelligence Impact Assessment: A New Quantitative Methodology in the Upcoming Era of Ai Act' (2025) 56(106101) CLSR. <https://www.sciencedirect.com/science/article/pii/S0267364924001663>.

[39] A Calvi, D Kotzinos, 'Enhancing AI fairness through impact assessment in the European Union: a legal and computer science perspective' in FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, 2023, 1229–1245. <https://dl.acm.org/doi/abs/10.1145/3593013.3594076>.

comprehensive for ensuring the anticipated preventive protection: these systems do not provide an overarching view of fundamental rights but rather focus on specific aspects, and a complete analysis is lacking, which is necessary to ensure the effective protection of individuals, especially in specific application contexts. In this context, due to the absence of an effective tool suited to evaluating the impact of AI on fundamental rights in a quantifiable, automatic, and objective manner, assessment errors could occur, potentially leading to excessive or unexpected costs for individuals and businesses.[40]

The difficulties that legislators face in developing tools to prevent and measure the negative impacts of new technologies on fundamental rights stem primarily from the fact that the so-called fourth digital revolution[41] and the social implementation of AI in all fields have placed contemporary society in a condition, described by Anders as characteristic of the technical man, of a Promethean gap, [42] such that we are no longer able to control our own technological actions, neither morally nor cognitively. This condition translates into a *de iure condendo* framework where the relationship between new technologies and law resembles the relationship between the hare and the tortoise, forcing regulation to set fixed points while also continuously updating them to the evolving AI systems that are the subject of legislative action. This creates the risk of legislative hypertrophy, leading to ever-decreasing legal certainty. The most vulnerable groups, both as individuals and as communities (the elderly, people with disabilities, minors, digital illiterates in general), are particularly affected by this situation, as they are most exposed to the risk of digital alienation of their rights and the erosion of social protection mediated by the law.

In this context, a critical issue that the legislator faces, within the framework of an inescapable need for an ethical approach to AI, is the feasibility of developing machine-readable ethical principles to be embedded within AI systems – an "ethical language" that machines could manage independently, thereby potentially relieving AI producers and operators from legal responsibility or redistributing such responsibilities in a more advantageous manner.

In this regard, although central to a digital ethics, modern technology makes the traditional concept of ethics insufficient. [43] In this view, Benanti and Maffettone [44] recognizing the criticality of a simplistic association between traditional ethics and the moral dilemmas posed by new technologies, argue that it would be wrong to think that it is possible to implement moral theories in AI systems if ethics is understood as the human competence to consciously discern what is right or wrong in a defined context of choice regarding the pursued ends or the consequences of one's actions. The algorithm cannot possess ethical capability, but it can – and must – be ethically designed.

[40] L Gatt and others, 'FRIA Implementation Model According to the AI Act' (2024) IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), 1224-1229. <https://ieeexplore.ieee.org/document/10796624>.
[41] L Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo* (Raffaello Cortina editore 2017).
[42] G Anders, *L'uomo è antiquato. I. Considerazioni sull'anima nell'epoca della seconda rivoluzione industriale*, trad. L. Dallapiccola (Bollati Boringhieri 2007).
[43] Portinaro (n 7) XXII.
[44] P Benanti, S Maffettone, *Noi e la macchina* (Luiss University Press 2024).

Algorithmetica, if we use this term to refer to the ethical reflection applied to the engineering of the algorithm, must be nothing more than this precautionary and responsibility-based principle in its design. There cannot be an ethics of the algorithm in the sense of a language, a moral competence that it can handle on its own, as one of the possibilities of a generative AI. The ethicality of the algorithm can only translate into an ethics-by-design approach by human operators who program the machines while being aware of the risks involved in implementing AI across various operational environments with social impact.

In a vision of trustworthy and ethical AI,[45] the choice of an applicable ethical theory, whatever it may be, must take into account the fact that relevant rights may conflict with each other (e.g., security and privacy regarding biometric data processing) and that not all human beings share the same values.

However, there are two concepts that, according to Benanti and Maffettone, remain constant, regardless of the ethical theory of reference: justification and value. First, any moral judgment must be justified. To do this, the agent who expresses the judgment indicates a norm that may not be shared by the reference community. From there, the agent refers to a moral principle – which in turn is based on a value – and ultimately to the chosen ethical theory. In this sense, justification and value are interconnected.

The challenge then lies in defining what values are considered relevant in the face of the risks of the digital revolution and why.

There are two manifestations of value that we experience daily: self-realization (leading a morally worthy life achieves the well-being of each individual) and respect (every human being recognizes the limits of their freedom in that of others). The ethical theory that best combines both forms, according to the two authors, is one that views value as an organic unity, where a surplus of value is realized when the value of the whole exceeds that of the sum of its individual parts.[46] This interpretation of value implies recognizing an objective – not just subjective – value for certain behaviors. Thus, we can say that human beings choose how to behave after resolving the tension between their moral actions as individuals and as members of a community, whatever that community may be. What must not be lost sight of, therefore, is the centrality of the human being, both generally and specifically in the governance of new machines that transition from being automatic to increasingly autonomous.

In conclusion, if it is true that ethics cannot be applied like engineering, regulation must address all use cases of AI systems that could negatively impact fundamental rights and therefore pose moral dilemmas (e.g., autonomous machines and aircraft, drones used for military purposes, etc.). This is the socio-technological character of AI: it loses its apparent neutrality in its practical applications.

If one were to apply the general theory that sees value as an organic unity to specific problematic cases, the following steps should be taken:[47]

---

[45] AI HLEG, 'Ethics guidelines for trustworthy AI' (2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
[46] Benanti, Maffettone (n 44) 88.
[47] Benanti, Maffettone (n 44) 91.

1. Carefully describe the case in question;
2. Explicitly identify how the facts affect shared values;
3. Know who is involved and who has a moral right in the case;
4. Highlight the foreseeable consequences of a particular decision;
5. Understand the relationships between the people involved and the foreseeable consequences of the choice made on them;
6. Opt for the choice linked to the maximum value in terms of organic unity.

What the two authors actually propose is the recourse to an ethical theory of value as an organic unity, in whose application the moral, social, and legal situation is crucial.

Addressing the issue of ethics for the digital realm is imperative: technological artifacts in general, and particularly computer algorithms, have caused an anthropological shock such that they delineate new hierarchies of values (for example, the conception of the common good changes), create new needs, and even radically transform the perception that humans have of themselves and the world in which they live. This condition inevitably has consequences for the moral actions of human beings. Hence the need to identify the ethical principles that the design and use of AI systems should be inspired by, taking into account the possible impacts these systems may have on politics, the economy, and law.

Any discourse on algorethicality begins with the need to rethink subjectivity as no longer individual, as a privileged prerogative of single human actors:[48] the human being who cooperates in their moral actions with non-human subjectivities is defined as *homo algorithmus.*[49] Within this reciprocal influence, it is necessary to identify ethical principles that can be shared by both components of the newly formed subjectivity. The fundamental ethical norm is the protection of human dignity, from which a series of principles descend: intuition, intelligibility, and adaptability.

First, the algorithm must be capable of perceiving human intentions and adapting accordingly. This is made possible by machine learning and, therefore, by human experience. The principle of intelligibility implies that humans must be able to intuit the action the algorithm will perform. Thus, it must be transparent and traceable, meaning that humans should be able to understand the underlying logic that led the AI system to produce a certain output based on input data, and this information must always be accessible to avoid the opacity (black box) typical of more complex AI systems. Finally, the adaptability of the algorithm to human intentions stems from the ability of the technological artifact to relate to its human counterpart, whether programmer or user.

A moral choice generally involves three elements: awareness, freedom, and responsibility. [50] The intelligence attributed to artificial entities does not correspond to human intelligence: the former is capable of finding solutions to specific problems, while the latter is of a general type. Thus, the algorithm may

[48] C Accotto, *Il mondo dato: cinque brevi lezioni di filosofia digitale* (Egea 2017) 93.
[49] A Picchiarelli, 'L'algoretica e lo sviluppo etico dell'Intelligenza Artificiale' in R. Razzante (ed.), *Manuale sull'Intelligenza Artificiale* (Giappichelli Editore 2024).
[50] Picchiarelli (n 49).

only have practical awareness related to the specific problem it addresses and for which it was created.

The freedom that leads humans to perform moral actions is influenced by the new filter between the agent and the world, represented by the algorithm: new technologies have exponentially increased human capabilities, but with unpredictable outcomes. This leads to a concept of uncertain collaboration (uncertainty) between humans and algorithms.[51]

Regarding the concept of responsibility, since moral action is the result of cooperation between multiple agents – human and non-human – some philosophers have used the concept of distributed responsibility. All moral agents are accountable for their actions, but they are not always considered responsible:[52] in the event of damage caused by an algorithmic system used by a human, it is not appropriate to speak of full responsibility on the part of the artificial agent but rather of accountability (attribution) for the effect of the influence the algorithm has on the human agent.

To avoid technocratic developments, where technologies – or rather, those who govern them – determine the relevant values rather than communities, and to achieve authentically anthropocentric technological development, an interdisciplinary approach is necessary. This approach should develop governance frameworks that ensure collaboration between humans and algorithms according to the human-in-the-loop paradigm.

Although from different perspectives, both Benanti and Maffettone, and Picchiarelli, question the possibility of applying moral categories to artificial agents and the need to develop a new ethics that is not a replacement but an integrative extension of traditional ethics.

A measured and measurable law seems clearly possible, but it is essential to follow the digital innovative path that scholars are committed to design, as scales, values, and norms serve as closure elements of the law and guide its related outcomes.[53]

At the same time, just as there is an increasing need to valorize a decision-making method that is humanly verifiable in terms of logic, controllability, and justification,[54] there is an ever-stronger demand to monitor the impact of artificial intelligence models on equity.

Social equity proves to be a concept of significant complexity: it is necessary to consider the various applications of AI, capture the different types of damage it could potentially cause, and reflect the perspectives of different existing communities. On this point, although simple rankings were relied upon for a long time to assess the potential damage to various communities, more recently, the need for a more comprehensive approach has emerged: shifting from AI equity rankings, which attempt to reduce equity to a single number, to a comprehensive suite of benchmarks that perform a variety of measurements and can provide tools to understand both the broad range of potential harms and trade-offs between different aspects of equity.[55]

---

[51] Picchiarelli (n 49).
[52] Picchiarelli (n 49) 8.
[53] G Corasaniti, *Informatica giuridica e progettazione innovativa digitale* (Cedam 2024) 185.
[54] Corasaniti (n 53) 187.
[55] A Wang, A Hertzmann, O Russakovsky, 'Benchmark suites instead of leaderboards for evaluating AI fairness' (2024) 5 (11) Patterns. <https://www.cell.com/patterns/fulltext/S2666-3899(24)00239-3>.

5.1 Possible new developments.

It is indisputable that the FRIA (Fundamental Rights Impact Assessment) in its current state presents limitations: the AI Act alone is insufficient to specify how to effectively implement a FRIA, as it does not provide the parameters for creating appropriate measurement pathways.[56] Instead, it only includes a generic reference to Article 27, paragraph 5, regarding the preparation by the AI Office of a 'template for a questionnaire, including through an automated tool, to facilitate deployers in complying with their obligations'. At the same time, there is an increasing need for an assessment conducted by public and impartial bodies acting as true certifying entities for AI that respects fundamental rights.

It is precisely in this context that the project born from the collaboration between the Research Centre ReCEPL (Research Centre in European Private Law),[57] CNF (National Forensic Council)[58] and CNR-Isti (National Research Council - Institute of Science and Technology for Information "Alessandro Faedo"),[59] the FRIA Project, is positioned.

The partnership project, in which the CNR-Isti has been replaced by the HABES research laboratory (Hardware Assisted and Blockchain Empowered Security lab)[60] and the CoLAB-AI Observatory, in addition to the typical challenges of complex legislation, however, presents a further challenge, this one not only of a technical nature: it is indeed essential to establish a shared methodological approach that allows for effective collaboration both among the project's various partners and with the public authorities, both Italian and European, some of which are still in the process of consolidation.

The multidisciplinary team is working on a project to assess the impact that artificial intelligence systems can have on the fundamental rights of individuals, with the aim of developing a tool capable of semi-automatically identifying the risks associated with the use of these technologies. The ultimate goal is to provide a tool that ensures effective, rigorous, and timely analysis, fully compliant with the complex regulatory framework currently in place regarding artificial intelligence.

This is, therefore, an initiative that requires not only high-level legal and scientific expertise but also an interdisciplinary and collaborative approach, aimed at building a concrete and productive dialogue between research and various institutions, in order to jointly address the regulation of artificial intelligence and the protection of fundamental rights – issues that are as complex as they are important.

---

[56] L Gatt and others 'FRIA Implementation Model According to the AI Act' (2024) IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), 1224-1229. <https://ieeexplore.ieee.org/document/10796624>.
[57] RECEPL - Research Centre in European Private Law. <https://www.unisob.na.it/ateneo/c008.htm?vr=3>.
[58] CNF - Consiglio Nazionale Forense. <https://www.consiglionazionaleforense.it/>.
[59] CNR-Isti - Consiglio Nazionale delle Ricerche - Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo". <https://www.isti.cnr.it/it/>.
[60] HABES Lab - Hardware Assisted and Blockchain Empowered Security lab. <https://habes.cs.unisa.it/>.

6. Conclusions.

It was difficult to imagine that on May 30, 2023, a statement of this kind would come from the upper echelons of digital technocracy: 'AI could lead to the extinction of humanity'. This alarm was raised by AI gurus such as Sam Altman (ChatGPT OpenAI), Demis Hassabis (Google DeepMind), and Dario Amodei (Anthropic), following some hearings in the US Congress on AI, Elon Musk's call for a moratorium on AI research to assess its risks and social impact, and a contribution by Yuval Harari on The Economist, where he argued that by taking control of language in general in our increasingly mediated society, AI has essentially "hacked" the operating system of our civilization. These concerns thus reached the general public, beyond the specialized debate.

This makes even more urgent the question underlying this contribution on what we can expect from AI in the coming years, after the first European law on the matter. A question that challenges the common narrative about AI, according to which the US develops the technology, China copies it, and Europe limits itself to regulating it. This argument is often used to push the European Union and its Member States to invest more in the development of AI systems in order to position themselves credibly within the strategies of major private and public players on an international scale, but it risks undermining the systemic scope – and the global momentum – of the importance of regulating AI implementation in the maturing digital society. It is now a regulatory request even from the Big Tech companies themselves, and it is absolutely decisive in order to ensure not only the overall well-being of society, with an anthropocentric use of AI, but specifically the fundamental rights of individuals, not only as single individuals but as members of a social community.

In response to this regulatory need, since 2020 – with the European Commission's Communication A European Strategy for Data[61] – the European Union has expressed its ambition to become a global benchmark for the regulation of new technologies and data governance. Unlike the U.S. and Chinese models – which respectively entail a high concentration of data space management by the private sector, and a combination of state surveillance and control by Big Tech – the European Union proposes a third way for regulating the processing of data essential to the development of AI models. This approach seeks to reconcile the legitimate interest in collecting, processing, and using data with the protection of fundamental rights. This model translates into the creation of a single data market through the establishment of European data spaces, the enhancement of computing infrastructure, the promotion of data portability and sharing, and the development of digital skills. The Continent Action Plan[62] of 9 April represents a further step in the European political strategy aimed at pursuing a 'third way' compared to the U.S. and Chinese models mentioned above. In this communication, the Commission

[61] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "A European Strategy for data, Brussels, 19.2.2020, COM (2020) 66 final. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066>.

[62] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "AI Continent Action Plan", Brussels, 9.4.2025, COM (2025) 165 final. <https://eur-lex.europa.eu/legal-content/EN/ TXT/?uri=COM:2025:165:FIN>.

reiterates the Union's ambition to become a global leader in artificial intelligence, while ensuring its development remains human-centred and trustworthy. This objective is to be achieved through concrete actions along five main lines: enhancing computing infrastructure; ensuring greater access to high-quality data; developing AI algorithms and increasing their application in strategic sectors; strengthening digital literacy at all levels of the population; and optimising single market regulation to improve regulatory compliance.

Within the Union, computing capacity – which is essential for the development of AI models throughout their entire life cycle – is made accessible through the network of cutting-edge European supercomputers coordinated by the European High-Performance Computing Joint Undertaking (EuroHPC), such as LUMI, Leonardo, and MareNostrum. This partnership provides substantial resources to start-ups, industry, public administrations, and researchers, enabling broad participation in the development of state-of-the-art AI models.

The EuroHPC has recently been enhanced through the AI Innovation Package (2024), which provides for the creation of AI Factories – currently thirteen – designed to foster innovation, cross-sector collaboration, and the application of new models in strategic areas such as health, technology, the environment, education, industry, finance, agriculture, cybersecurity, space, and the public sector. These infrastructures are open to participation by EU Member States and associated third countries, which may establish AI Factory Antennas to provide remote access to the supercomputing resources of the corresponding AI Factory located in another Member State.

The EuroHPC will serve as a single point of access to the support services offered by the AI Factories for users across Europe, particularly start-ups, SMEs, and EU-funded research projects, ensuring streamlined access to computational resources while reducing administrative costs. The next step will be the establishment of AI Gigafactories – large-scale infrastructures dedicated to the development and training of next-generation AI models with trillions of parameters. To this end, these infrastructures will focus on power capacity, reliable supply chains, advanced networks, energy efficiency, and AI-driven automation. AI Gigafactories are essential to enhancing Europe's global competitiveness and safeguarding its autonomy in scientific progress and strategic industrial sectors. These infrastructures could foster the development of a European AI processor manufacturing system, thereby reducing the Union's dependence on third countries for the supply of critical technologies. Such a complex infrastructure network requires substantial investment. Accordingly, the European Union is exploring new hybrid investment models that engage both the public and private sectors. At the AI Action Summit held in Paris last February, European Commission President von der Leyen announced the InvestimentAI initiative, a public-private partnership intended to mobilise approximately €200 billion for the development of AI Gigafactories.

European enterprises must be placed in a position to close the gaps in general cloud capacity – typically provided by large-scale data centres – and in edge capacity, which enables similar services to be delivered with significantly lower response times. This is because European data centres have significantly lower

capacity than those in the United States and China, and therefore depend on infrastructures located in non-European countries. Such dependence may negatively impact economic security and public administrations. Through the proposed Cloud and AI Development Act, the Commission aims to encourage investment in cloud and edge capacity to meet European cloud demand by 2035. Given that data centres can adversely affect energy consumption and the environment, the European digitalisation roadmap includes measures designed to ensure that these facilities are energy sustainable.

In order to enhance both the quantity and quality of data for the development of AI models, a new European data strategy will be developed with the aim of creating a single market for cloud services at the European level, targeting both public and private sector users, and designed to be as balanced, secure, and competitive as possible. To provide this infrastructure with a solid regulatory framework, the Commission intends to systematise the fragmented rules applicable to cloud services by developing a Cloud Code that integrates existing codes of conduct and certification schemes on security, energy efficiency, quality, data protection, and data portability. An example is CISPE (Cloud Infrastructure Services Providers in Europe), a European code of conduct for data protection in cloud services. It is compliant with the rules established by the GDPR and was validated by the EDPB in its Opinion 17/2021. In particular, providers adhering to this code are prohibited from using their customers' data for profiling or direct marketing. Common European rules and conditions will also be established for the award of public contracts for data processing services. The Commission will also seek to simplify legislation to reduce administrative burdens and ensure the effectiveness and efficiency of data governance structures. Among the tools to be deployed are Data Labs, which will collect and aggregate data from various AI Factories operating in the same sectors. These infrastructures will not only guarantee access to European data spaces, such as the European Open Science Cloud, but also enhance the interoperability of different data processing and collection systems.

Currently, the application of AI models involves less than 20% of European companies, particularly SMEs. To address this shortcoming, the Commission aims to implement the so-called Apply AI Strategy, promoting innovation and the development of European AI solutions both cross-sectorally and within sectors. This strategy will foster the adoption of AI by the private sector (including strategic industrial sectors such as research, defence, pharmaceuticals, and robotics) and the public sector (services such as health, justice, education, and public administration), with the aim of balancing the protection of fundamental rights with the Union's economic interests.

One of the instruments to support the digitalisation of these sectors will be the European Digital Innovation Hubs, which, in combination with the AI Factories, will enable companies to access computational resources and experimentation spaces through Testing and Experimentation Facilities, particularly in the fields of health, industry and energy, agriculture, and mobility.

Significant investments will also be allocated to Research and Innovation (R&I) through a series of initiatives such as GenAI4EU, under the Horizon 2026–2027 programme, for the development of advanced European AI models, pilot

projects for the digitalisation of public administrations, and the establishment of the European Centre for AI Research, known as RAISE (Resource for AI Science in Europe).

Last March, the Commission launched the Union of Skills strategy[63] with the aim of strengthening the Union's competitiveness through the development and mobility of skills. Its main objectives are: fostering lifelong skills through a solid educational foundation; upskilling and reskilling; promoting skills mobility to unlock the potential of the single market; and attracting and retaining talent from third countries to address skills shortages and nurture top talent within Europe.

In this context, a series of initiatives will be launched to promote digital literacy at all levels of education. These include the 2030 Roadmap for Digital Education, aimed at implementing digital skills among school students and teachers; the establishment of the AI Skill Academy, which will develop curricula focused on new technologies – particularly GenAI – as well as projects to support the reintegration into the workforce of women who have interrupted their careers. To promote skills mobility, measures will be introduced to optimise European legislation that allows highly qualified workers and researchers from third countries to reside in Member States, along with the Marie Skłodowska-Curie Actions MSCA Choose Europe, a European funding programme that supports the training of researchers at both doctoral and postdoctoral level.

The European Union's ambition to become a global benchmark for the development and application of new technologies must be supported by a solid and clear regulatory framework. The proper functioning of the European single market for AI will only be possible if appropriate measures are adopted to implement the AI Act. To this end, the European institutions will promote a series of actions, including the establishment of an AI Act Service Desk within the Office for AI – in collaboration with the network of European Digital Innovation Hubs – to provide practical guidance to AI model providers and deployers (particularly small-scale ones) in achieving regulatory compliance, also offering self-assessment tools; and the creation of the AI Pact, managed by the Office for AI, which encourages and supports European companies in anticipating the adoption of certain provisions of the AI Act. Moreover, public consultations are planned to identify the main shortcomings of the new regulation, along with the establishment of partnerships with third countries in the field of research and innovation.

From what has been examined so far, it is clear that the goal of making the European Union the first 'AI Continent' cannot be achieved without the shared commitment of European and national institutions, researchers, and entrepreneurs.

The significance of the regulatory and normative issues surrounding AI, in relation to the ambivalence of its social implementation as a tool for social inclusion and the anthropocentric promotion of economic development, or for governance and control in what Zuboff calls "surveillance capitalism", pertains

[63] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "The Union of Skills", Brussels, 5.3.2025, COM (2025) 90 final. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri =CELEX:52025DC0090>.

to a framework of digitization that defines the world and society as an infosphere, augmented by AI. An infosphere that – thanks to Metaverse technologies – makes it possible, within the mass social experiences, to transition from the real (analog) experience to the virtual (digital) one, which now frames life as "Onlife".[64] This mass digital immersion raises concerns about disruptive effects of colonization on social language, where the individual is no longer an empirical individual valued in their singularity (person) at the center of the legal experience[65] but is reduced to a "profile" available for economic and political marketing, governed by algorithms.[66] A condition that could be defined as algocracy, which is fueled by unequal exchanges between "free" services offered to users and their mass datafication for economic and political marketing purposes, and ultimately systemic social control, on the other hand, by an AI imaginary that promises individuals an indefinite expansion of their "creative" experience in the virtual environment provided by the Metaverse(s).

In addition to the clear and concrete economic and social problems (labor market, redefinition of the value chain, access to new digital citizenship), there is the risk of disabling cognitive, operational, and decision-making skills in individuals in favor of greater algorithmic efficiency. This presents a framework where human rights as a whole are at stake. Hence the need for governance of AI technologies and a detailed assessment of the effectiveness of the European regulation on AI in terms of expected outcomes for the protection of fundamental rights related to human dignity, equality, democracy, and the rule of law (Art. 2 TEU).

On the topic of participation in democratic life, De Luca[67] explores the current debate on the relationship between AI and politics, especially after the rise of Generative AI, providing an overview of the challenges and opportunities AI presents in the contemporary political context. The author highlights how AI is altering power dynamics within nation-states and the way these states relate to each other, pointing out an increasing dependence of governments on Big Tech for executive functions such as security and policing. GenAI has also facilitated the emergence of virtual politicians and new forms of political interaction, such as the so-called Augmented Democracy, where digital twins can be used to enhance participation and decision-making processes, thus impacting how political campaigns are conducted.

GenAI has shown a superior persuasive capacity compared to human arguments, making it a powerful tool for influencing public opinion. This raises concerns about public manipulation and the spread of misinformation, due to the non-neutrality of algorithms and techniques like deepfakes. Hence, the need for continuous vigilance from civil society, as well as preventive ad hoc measures. [68] In the AI ecosystem, Big Data, computational power, and automation play a significant role, particularly in how they can impact

[64] Floridi (n 41).
[65] P Grossi, 'Uno storico del diritto in colloquio con Capograssi' (2006) 83(1) Rivista Internazionale di Filosofia del Diritto 13-40.
[66] A Simoncini, 'Il diritto nell'era del dominio tecnico: la lente dell'esperienza giuridica' in G Cazzetta, G Morbidelli and O Roselli (eds.), *Alla ricerca della «dimensione giuridica». Omaggio a Paolo Grossi* (il Mulino 2024) 347-396.
[67] S De Luca, 'Intelligenza Artificiale e politica: l'ultima frontiera' (2024) 3 Storia del pensiero politico.
[68] See also Corasaniti (n 20) 41 ff.

fundamental rights and informational pluralism. The increasing use of generative AI by the media to produce content puts traditional journalistic organizations under pressure, risking market saturation in the information industry, with increased competition from automated systems generating low-cost content.

A key concept addressed is digital constitutionalism, which involves not only the emergence of new rights but also the necessary development of new tools to limit the concentration of power that has shifted from governments to large private digital operators. It is, therefore, necessary to rediscover the dual nature of the concept of constitution, not only as a bill of rights but also as a frame of government. Given that self-regulation by Big Tech is largely illusory,[69] a balance between judicial and legislative power is needed, so that the political-representative circuit can reclaim its role as lawmaker.[70] Additionally, there is the evidence that the impact of AI and digital technologies is not uniform but has asymmetric effects on democracies and authoritarian systems.

In conclusion, in a world increasingly shaped by AI and ICTs, defined in informational terms, a machine may help humans "save the world", but it must never be left alone to decide or choose.[71] It will be crucial, on one hand, to keep in mind the need for constant updates to the regulatory framework in relation to a continually evolving technological reality, and on the other hand, to monitor the effectiveness of the regulatory solutions proposed so far by the AI Act. This implies that it should not be a mere activity of receiving what has been "handed down from above", but rather an ongoing research activity aimed at fully understanding the legal provisions and preparing the necessary technical tools to keep pace with this Revolution.

[69] S De Luca, 'Intelligenza Artificiale e politica: l'ultima frontiera' (2024) 3 Storia del pensiero politico.
[70] O Pollicino, P Dunn, *Intelligenza artificiale e democrazia: opportunità e rischi di disinformazione e discriminazione* (Egea 2024).
[71] Corasaniti (n 20) 43.