ETHICAL AND LEGAL ASPECTS OF THE EUROPEAN UNION REGARDING THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE IN THE FIELD OF NATIONAL SECURITY – A SOCIAL PERSEPCTIVE

Cpt. Ilie IFTIME

PhD candidate at "Carol I" National Defence University, Bucharest, Romania

https://doi.org/10.66793/tituecir19proceeding16

Abstract:

In the context of an increasingly intense global competition (both among private corporations and various international actors) with regard to the research and development of artificial intelligence (AI), alongside the continuous stream of recent innovations and the growing demand among end-users for such emerging technologies, a fundamental question arises: **Does the current legislative framework proposed by the European Union adequately ensure security in the research, development, and operational deployment of AI systems?**

This paper seeks to explore several issues highlighted by public discourse, including but not limited to: What are the key legislative instruments that safeguard social protection in the event of adverse effects caused by AI technologies? Why is the national security sector exempted from the obligations set out under the current EU regulatory framework, and to what extent does this exemption impact the social sphere? Are the newly introduced legislative requirements sufficiently robust, clearly defined, and readily implementable? Do these regulations hinder innovation and the competitive capacity of major companies? What risks and challenges may arise from the current and future regulatory developments in the field of AI? Within this analytical context, the primary focus of scientific inquiry will be the social domain, given that it remains the principal beneficiary of products and services built upon artificial intelligence.

The current regulation of the field of artificial intelligence at the level of the European Union

The rapid and exponential rise of technology within contemporary society is an undeniable reality. This evolution has culminated in the attainment of a superior level of capability, namely the domain of artificial intelligence. Its utility is now recognized across all sectors of society (research and development, industry, medicine, defense, education, etc.), with its primary observable effects including: reduced time in service delivery; the removal of repetitive tasks from the workload of human resources; the capacity to analyze significantly larger volumes of data in a shorter period of time while simultaneously delivering higher-quality outputs; enhanced decision-making through the testing of scenarios and hypotheses, as well as the provision of predictions and recommendations; and increased social and cybersecurity protection, among others.

Nevertheless, due to the continuous development of artificial intelligence and the growing willingness to delegate an increasing number of tasks to such systems, so that they may become capable of independent reasoning and autonomous decision-making in certain circumstances, a set of risks to which society may be exposed is increasingly emerging. Consequently, states and international organizations have begun to more rigorously regulate this field.

At European Union level, the main legal instrument is the EU Artificial Intelligence Act, which entered into force in 2024 and was subsequently updated in 2025. The Act classifies and defines AI systems into four categories based on their associated risk: *"unacceptable risk, high risk, limited risk, and minimal risk"* **[1]**, with the first category being strictly prohibited from use. However, the national security domain is exempted from this rule, an aspect outlined in Article 2, paragraph 3: *"This Regulation shall not apply to areas outside the scope of Union law and shall not affect, in any way, the competence of Member States concerning national security, irrespective of the type of entity entrusted by Member States with tasks related to such competences"* **[1]**. In the same vein, it does not apply to AI systems already placed, or not yet placed, on the market and used exclusively for military, defense, or national security purposes, regardless of the entity performing the activity.

In this respect, it may be observed that the current legislative framework provides substantial flexibility to the national security sector with regard to the use of artificial intelligence systems. Consequently, the potential gap in the development and operationalization of AI compared with states governed by authoritarian regimes (a concern identified by several experts in the field) is mitigated on this strategic front.

Furthermore, one of the most significant questions that remains is what happens if certain authorities operating in the field of national security unjustly interfere with the private life of a citizen by using AI-based technologies, particularly those classified as presenting unacceptable risks, such as:

• *"systems for evaluating and classifying individuals based on social behavior, known, inferred, or predicted personal characteristics, and assigning a negative social score that will subsequently affect certain aspects of their private life;*

• *systems for assessing the risk that a person may commit specific criminal offenses;*

• *the creation and expansion of databases by extracting data from facial recognition systems;*

• *the classification of individuals according to biometric data in order to create target groups defined by features such as race, political opinions, religious or philosophical beliefs, sexual orientation, etc."***[11]**

Therefore, limiting abuses in the use of AI under the pretext of national security concerns is also achieved through other legislation closely linked to the EU AI Act, but having a greater impact on the use of artificial intelligence in the field of national security, namely the legislation concerning the protection of personal data (GDPR). Within this framework, specific ethical and legal constraints are identified, which affect the need to employ AI in contexts other than armed conflict and criminal prosecution.

The guarantee of ethical and legal standards may also be achieved through an internal mechanism for certifying AI systems by the state concerned. Thus, algorithms are subject to verification by the provider, and consequently, not every AI technology will be permitted to enter the market. Additionally, as in any sensitive issue identified at the national level, training programs for different professional categories and awareness initiatives for the general public have a considerable impact on the responsible use of AI.

Another legal issue, which may be interpreted on a case-by-case basis, concerns the field of electronic communications, where the confidentiality of data imposed on service providers may be compromised. According to current legislation, the direct collection and processing of personal data by national security authorities does not fall under the scope of the GDPR; however, when such data are obtained with the assistance of third parties, the situation differs, and the European Court of Human Rights must examine each case individually.

Alongside the EU AI Act, which has internal applicability only within the Union, there is also the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, which holds the status of an international treaty and may be used by other states outside the EU as an international benchmark for addressing AI governance. It likewise becomes a standard for states aspiring to obtain EU membership. This convention draws attention to the risk that AI may *"undermine human dignity and individual autonomy, human rights, democracy, and the rule of law"* **[2]** and emphasizes several issues such as the growing social inequality, discrimination, unlawful and arbitrary surveillance, blackmail, manipulation, disinformation, etc. In this regard, the document provides a set of limitations, principles that must be respected, rights, safeguards, as well as the possibility of pursuing legal remedies by individuals who consider that AI systems have been abusively applied in relation to their person.

In the field of national security, the convention largely maintains the same guidelines as the EU AI Act, but introduces an important condition: compliance with international human rights law as a fundamental prerequisite. Specifically: *"AI activities related to national security are excluded from the scope of the convention, provided that such activities are in all cases conducted in a manner consistent with applicable international human rights law and with respect for democratic institutions and processes"* **[2]**. Although the European Convention on Human Rights also states that certain civil rights may be restricted in cases of criminal prosecution or national security concerns, it nonetheless provides a set of criteria for assessing the intervention of authorities in private life (whether a protected right was interfered with, whether the interference was lawful with a legitimate aim, and whether it was strictly necessary). Unfortunately, a general truth that applies to a greater or lesser extent depending on each state suggests that certain issues, such as for example, "the manipulation of nuclear weapons", may be considered matters of high politics, while violations of human rights may be categorized as matters of low politics.

National security, as discussed at the level of the European Union, is closely linked to that of NATO. Consequently, it can be observed that, under the same overarching guidelines, NATO's vision does not significantly differ from that of the EU when referring to a "responsible development of AI" (as set out in the NATO Artificial Intelligence Strategy launched in 2021 and updated in 2024) **[4]**. Under this imperative, new terms have emerged in

the field, such as TAI — Trustworthy AI, and XAI — Explainable AI. The European Commission places strong emphasis on these concepts and has imposed a set of criteria in this regard, often known as the seven principles of a reliable AI system:

• "the protection of fundamental human rights and interaction limited solely to the provision of assistance;
• *robustness and safety through the ability to self-correct certain errors;*
• *preservation of privacy and data security against unauthorized users;*
• *ensuring transparency of system functioning and explainability of processes and outputs;*
• *respect for diversity and fairness by understanding differences among users and avoiding discrimination;*
• *development grounded in environmental protection, sustainability, and ensuring social well-being;*
• *accountability of the AI system through the reporting of any internally identified issues and responsibility for the outcomes generated"* **[5]**.

These principles aim to ensure that AI becomes safer and more accessible, while providing well-reasoned information, presented in an intelligible manner, enabling the user to understand aspects such as: why specific results were provided, why alternatives were not offered, when the system is unable to provide the desired output, under what circumstances it can or cannot be trusted, and when its responses are ambiguous or unreliable.

Risks and challenges in the process of regulating artificial intelligence

Given that the research and development of artificial intelligence is conducted today at an unprecedented level—amounting to a veritable "algorithmic arms race," in which most actors invest significant resources—we are witnessing a regulatory framework for AI that attempts to keep pace with ongoing innovations in the field. Consequently, a series of risks and challenges emerge in the process of regulating artificial intelligence.

One such challenge concerns the terminology used in this domain. The main legal document itself, the EU Artificial Intelligence Act, has been criticized for being overly complex and lacking clarity **[6]**. Over time, various ambiguities have arisen in the definition of certain terms, concepts, components, and specific processes. Although many of these issues have been clarified, there remain several concepts that are still vague, insufficiently delimited, and not clearly regulated, such as *human vulnerability* and *artificial intelligence system*. In the case of *human vulnerability*, for instance, although the expression is widely used, it is still not precisely defined so as to encompass all aspects reflecting the full range of possible human error. Therefore, the lack of clarity surrounding some of the concepts being used may, on the one hand, create uncertainty for authorities, operators, and AI users, and, on the other hand, open small loopholes enabling the avoidance of responsibility or the use of technology for unintended purposes. This limitation, however, may be understood primarily through the lens of the pioneering stage of legislation specific to artificial intelligence.

Another issue reflects the opposite extreme of the previous one, namely *overregulation*. This challenge has been flagged by various AI-driven enterprises that claim such an approach will slow the pace of innovation (due to increasingly complex standards, the requirement to modify or adapt infrastructure, and the additional cost and time involved **[9]**). Up to a certain point, companies are encouraged, having prioritized, simplified, and free access to so-called *sandbox environments* that allow the testing of technologies in real-world conditions; however, in subsequent stages the process becomes more burdensome (adapting to new requirements, market authorization, provision of guarantees, etc.). At the level of EU Member States, this phenomenon translates into slower implementation of AI technologies due to rising costs and more extensive compliance obligations, while also generating inequalities among states in terms of security and social protection resulting from differing speeds of implementation **[7]**. This aspect must be prioritised given that states are part of international alliances and agreements, and the EU's role as a global actor is contingent upon coordinated and unified implementation in this area. Europe's competitiveness with other major international actors (such as the United States and China) will also be negatively affected, as the strategic advantage conferred by innovation may be neutralized by the complex and cumbersome mechanism required for its deployment. Therefore, overregulation of AI may generate vulnerabilities both in the field of security and within the social sphere.

Another aspect that may be considered a challenge in clearly delimiting the boundaries of AI research and development relates to the field of national security, specifically those military and defense technologies that are exempted from the strict rules of the EU AI Act. This category, which may include AI products classified as presenting an unacceptable level of risk, has the potential to generate real threats to societal safety if not managed under strict oversight. Consequently, a new grey area emerges, where subjective interpretations may arise when labeling issues as matters of national security, thereby justifying the use of such technologies. Moreover, abuses may

be enabled through the coordination of targeted and discriminatory actions lacking transparency, involving infringements upon private life, with the aim of exercising control, manipulation, maintaining power, influencing decisions, etc. This concern arises from the fact that most legal obligations fall on developers and providers, with fewer responsibilities applying to end users **[8]**. A concrete example in this regard is the AI facial recognition systems installed on the streets of major Russian cities. These systems have enabled the identification and detention of numerous protesters criticizing the Kremlin's intervention in the war in Ukraine at various locations and times of day. Another illustrative example is the social credit system implemented in China several years ago. AI systems make automated decisions regarding individuals' access to certain services (healthcare, justice, education, etc.) solely based on a profile and social score generated from the continuously monitored behavior of the respective citizen (professional performance, civic conduct, political affiliation, etc.). Thus, a wide array of basic civil rights and liberties fundamental to a democratic society are more easily and severely infringed upon through the use of such AI technologies.

From a more technical perspective on the research and development process, it must be understood that AI security risks may arise at every stage of the life cycle (*"initiation, design and development, verification and validation, deployment, operation and monitoring, reassessment, retirement"* **[10]***)*, as proposed by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) under standard ISO/IEC 22989:2022:

• *"initiation: the intended objectives of AI applications may violate national laws/regulations and ethical principles;*
• *design and development: flawed AI infrastructures, technical vulnerabilities, errors in design and development;*
• *verification and validation: insufficient testing, failure to remedy previously identified risks;*
• *deployment: use of insecure or unverifiable hardware/software, unauthorized access and usage;*
• *operation and monitoring: direct attacks / backdoor attacks / model inversion / user simulation attacks, model theft, error generation through manipulated feedback, exploitation of code vulnerabilities, misuse and malicious use of the system;*
• *continuous validation: delays in updating data, inability to detect and correct errors in due time;*
• *reassessment: similar issues to those in the initial stage;*
• *retirement: incomplete destruction and disclosure of data and information"* **[10]**.

Therefore, at any stage, hidden purposes, faulty designs, security breaches, misuse, data leaks, system hijacking, and other incidents may arise, enabling technology to be used for ends other than those initially intended. These risks, as detailed by experts in the field, implicitly highlight a set of requirements such as: the need for highly qualified and responsible personnel; strengthening the legislative framework by addressing technical aspects in detail; constant and strict monitoring of all stages; and rigorous testing of AI products that will be used by the population in order to prevent data theft and avoid misleading users through erroneous services (the system must constantly communicate its limitations to the end-users).

Ultimately, although the legislation is complex and restrictive, it must be understood that both the authorities and institutions responsible for AI regulation and the companies involved in AI research and development must find a balance between current and future legal requirements and AI-based products, ensuring that these technologies remain useful and safe for society, compliant with the legislation and norms in force, while also supporting the necessity of maintaining high levels of innovation and market competitiveness — particularly in the context of unsafe alternatives proposed by other international actors whose values and interests differ from those of the European Union.

Conclusions:

The regulation of AI by the European Union, and subsidiarily by the Member States, may represent a challenging task in the context of rapid developments in this sector, as well as due to the legislative pioneering in a relatively new field. Nevertheless, the responsible authorities and institutions must anticipate and proactively establish the legal framework for the research, development, and use of technologies based on artificial intelligence. A subsequent adaptation to such evolutions, and the time gap between the emergence of an innovation and the creation of the appropriate legal framework, may provide malicious actors with genuine vulnerabilities that could be exploited.

Furthermore, beyond the legislative framework and the deadlines set, the EU bears responsibility for providing mechanisms that facilitate the Member States' adaptation to the new requirements and their uniform implementation, so as to avoid various internal blockages and to prevent the loss of valuable time in this global strategic competition

in the field of AI. Even if we refer to a complex legislative framework, filled with new standards, requirements, and obligations, such mechanisms can reduce rigidity, latency, and even related compliance costs.

Although exceptions exist in the field of national security and defence, these must be strictly delimited through legislative instruments, physical protection measures, and continuous monitoring, as their extraction and utilisation for purposes other than those defining the present exemption may generate major negative effects on all security sectors (social, military, political, economic, and environmental).

At the same time, both the legislative framework and the companies involved in AI research and development must prioritise the provision of safe and useful products for the end-users, ensuring that no direct or indirect harm is inflicted upon democratic values and interests.

Bibliography:

1. *** *Legea EU privind AI*, 2024, URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689, accessed in 22.10.2025.
2. *** *Convenția cadru a Consiliului Europei privind inteligența artificială, drepturile omului, democrația și statul de drept*, 2024, URL: https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX:52024PC0264, accessed in 16.10.2025.
3. *** *An Artificial Intelligence Security Framework,* published in *Jorunal of Physics: Conference Series vol. 1948*, 2021, URL: https://www.proquest.com/docview/2546087129/C5CF9145AB714A98PQ/24?accountid=88069&sourcetype=Scholarly%20Journals, accessed in 12.10.2025.
4. *** *NATO releseas revised AI strategy*, 2024, URL*:* https://www.nato.int/cps/en/natohq/news_227234.htm, accessed in 30.10.2025.
5. *** *A review of Trustworthy and Explainable Artificial Intelligence,* published in *IEEE Access*, vol. 11, 2023, URL: https://0k113688w-y-https-ieeexplore-ieee-org.z.e-nformation.ro/document/10188681, accessed in 03.10.2025.
6. *** *Critics rise concerns about the EU AI Act*, URL: https://symbio6.nl/en/blog/criticism-of-eu-ai-act?, accesat în 29.10.2025.
7. Ahern Deirdre, *Operationalising AI Regulatory Sandboxes under the EU AI Act: The Triple Challenge of Capacity, Coordination and Attractiveness to Providers*, URL: https://arxiv.org/pdf/2509.05985, accesed in 22.10.2025.
8. Caspar Catherine, *The EU AI Act: A new legal framework for ethical, safe and innovative use of artificial intelligence*, 2025, URL: https://www.novagraaf.com/en/insights/eu-ai-act-new-legal-framework-ethical-safe-and-innovative-use-artificial-intelligence?, accesed in 28.10.2025.
9. Ditsche Jochen, Mikhaylenko Maria, *European AI Act: Opportunities and challenges,* 2024, URL: https://www.rolandberger.com/en/Insights/Publications/European-AI-Act-Opportunities-and-challenges.html?, accesed in 25.10.2025.
10. Javil Abdul, Welch Amber, *Artificial Intelligence lifecycle risk management*, 2025, URL: https://aws.amazon.com/blogs/security/ai-lifecycle-risk-management-iso-iec-420012023-for-ai-governance, accessed in 05.10.2025.
11. Juric Marko, *Legal regulation on the use of artificial intelligence for national security purposes in Europe,* 2024, published in *European Integration Studies, vol.20, nr. 2*, 2024, Zagreb, URL: https://ojs.uni-miskolc.hu/index.php/eis/article/view/3803, accessed in 14.10.2025.