

PROTECTION OF FUNDAMENTAL RIGHTS IN THE CONTEXT OF THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE

Roxana Silvia TRUȚA*

ABSTRACT: *Currently, artificial intelligence (AI) has strengthened as a tool with a profound impact on multiple sectors, including the legal field. The accelerated evolution of these technologies arouses a special interest from specialists, becoming a focal point of contemporary doctrinal debates. This paper proposes a dual approach: on the one hand, analyzing how AI can support and strengthen the guarantee of fundamental human rights and freedoms, and on the other hand, critical assessment of the vulnerabilities and threats that its implementation can generate. In this respect, particular emphasis will be placed on the imperative of respect for fundamental rights, addressing challenges such as: ensuring decisional transparency, preventing algorithmic discrimination and protecting the right to privacy in the context of massive data processing.*

Keywords: artificial intelligence, fundamental rights, protection, discrimination, guarantee
JEL CODE: K24

DOI: 10.62838/cjjc-2024-0078

1. SOME INTRODUCTORY ASPECTS ON THE CONCEPT OF ARTIFICIAL INTELLIGENCE

The inherent polyvalence of the concept of artificial intelligence (AI) has generated, in recent doctrine and practice, a multitude of interpretative valences. This diversity of approaches is a direct consequence of the ability of AI to structurally penetrate and transform fundamental sectors, generating new challenges in areas such as legal, economic or purely technological (Vartires, 2023, pp. 592-593). An in-depth analysis of this paradigm by the European Union Agency for Fundamental Rights (Pop, 2024) (FRA)¹, reveals the absence of a terminological consensus at the unitary level. Despite this doctrinal effervescence, a functional approach has been crystallized at the level of the European Union, artificial intelligence being defined by reference to systems capable of intelligent behavior (Duțu, 2024).

* Assistant Professor, Faculty of Economics and Law, George Emil Palade University of Medicine, Pharmacy, Science and Technology from Tîrgu-Mureș; Lawyer, Mureș Bar, ROMANIA.

¹ European Union Agency for Fundamental Rights

This capacity translates into the active assessment of the environment and the adoption of – measures characterised by a variable degree of autonomy – with a view to achieving pre-established objectives.

Structurally, the architecture of these systems based on artificial intelligence knows a double manifestation. On the one hand, they can be eminently immaterial, based exclusively on software solutions with virtual operability (algorithmic architectures such as voice assistants, advanced imaging analysis programs, search engines or facial and voice recognition systems).

On the other hand, AI can materialize through direct incorporation into hardware infrastructures, a hypothesis that includes advanced robotics, fully autonomous vehicles, drones or integrated applications in the field of *Internet of Things* (IoT) ².

Although the genesis of theoretical debates and the early stages of research on artificial intelligence may be placed temporally as early as the last century, particularly in the academic space of the United States (Moroianu, 2020), we are currently witnessing a profound paradigm shift. The disruptive impact of emerging technologies has surpassed the purely theoretical stage, generating visible global reverberations that reshape not only social dynamics, but the very substance of contemporary legal relations and fundamental rights.

2. THE INTERSECTION OF FUNDAMENTAL RIGHTS AND ARTIFICIAL INTELLIGENCE

The intersection between the architecture of fundamental rights and the exponential expansion of artificial intelligence represents, in the current paradigm of contemporary society, a critical research field, marked by a continuous doctrinal effervescence. Fundamental rights, conceived as inherent and inalienable prerogatives of the human being, enshrined and protected by constitutional rules and international legal instruments, cover a broad spectrum of essential freedoms, from respect for privacy and freedom of expression to the principle of non-discrimination and the right to a fair trial within a reasonable time. However, the increasing intrusion of algorithmic decision-making systems into traditional mechanisms to guarantee and protect these rights has revealed a number of major structural vulnerabilities that require an urgent re-evaluation of existing legal instruments (Manu, 2023).

A first layer of analysis that arouses controversy concerns the right to privacy and the guarantee of equal treatment (Bîrsan, 2006), principles under constant pressure in the digital age. Artificial intelligence systems base their functional architecture on the collection, storage and processing of massive volumes of personal data, generating substantial and imminent risks to individual privacy by resorting to intrusive technologies of generalized surveillance, data extraction and advanced algorithmic profiling techniques. Closely related to this vulnerability, the principle of non-discrimination is seriously exposed, as artificial intelligence models are likely to take over, perpetuate and even exacerbate historical biases or hidden errors in the datasets used to train them. This

² For more information on this, see the official FRA page:

https://fra.europa.eu/sites/default/files/fra_uploads/fra-2021-artificial-intelligence-summary_ro.pdf, accessed on 29 May 2026.

assimilation of inequities invariably leads to discriminatory results and flawed decisions, altering the very substance of the fundamental right to equal treatment and imposing the need for rigorous auditing mechanisms of systems.

On another plane of interest, freedom of expression and the right to a fair trial (Bodea, 2024) also feel the shock waves of technological innovation. Recommendation and filtering algorithms today have the ability to directly model information flow and, by extension, public discourse, generating harmful phenomena such as algorithmic censorship, filter bubbles and systemic propagation of disinformation, aspects that highlight the imperative need to achieve a right balance between the implementation of intelligent systems and the protection of freedom of information. Correlatively, in the context in which algorithmic entities participate in the adoption of decisions with a direct impact on the legitimate rights and interests of individuals, the guarantee of respect for the right to a fair trial becomes an ineluctable challenge. It is therefore necessary to enshrine legal transparency of algorithms, to guarantee the right of the litigant to challenge purely automated judgments and, in particular, to establish clear mechanisms of legal responsibility for developers and implementers of these architectures.

Last but not least, individual autonomy and freedom of thought are under siege by technologies capable of subversively influencing human cognitive and decision-making processes. Artificial intelligence is profoundly reshaping the dynamics of decision-making in vital sectors of society, with direct impact on healthcare, criminal justice and, quite particularly, on employment relations, where new forms of algorithmic management have the potential to considerably restrict the autonomy and dignity of the individual (Coman, 2024). The protection of self-determination imperatively requires the maintenance of significant and real human control over any decision-making system that produces major legal effects on the person's life. Thus, mitigating these multiple vulnerabilities requires an integrated response from the legislator, which brings together solid and imperative legal frameworks – illustrated eloquently by the requirements of the General Data Protection Regulation (GDPR) at European level –, with rigorous ethical standards and technical security guarantees, thus ensuring the pre-eminence of human rights in the face of technological advance.

In light of technological developments in recent years, artificial intelligence has ceased to be a mere tool of computation or scientific curiosity, becoming a structural force that profoundly reconfigures human interactions, economic systems, democratic institutions and the very essence of human rights. This societal metamorphosis brings with it extraordinary promises of medical efficiency, administrative optimization and scientific progress, but it also generates major systemic risks to individual dignity, equity and fundamental civil liberties. In this context marked by a fluid and often unpredictable dynamic, the defence of fundamental rights is no longer only a secondary ethical option or a theoretical concern, but an urgent existential necessity of the rule of law, which calls for the implementation of coherent strategies and rigorous mechanisms capable of ensuring that algorithmic systems are designed, developed and used in full compliance with universal humanist values. An integrated and successful approach to managing these major challenges requires an indissoluble synergy between three central pillars of contemporary society: the establishment of imperative legal frameworks and regulations by governmental authorities and international bodies, the assumption of responsible self-regulation and stringent ethical guidelines by the research community and industry, and as

well as cultivating an active civic consciousness through public education and open dialogue among the entire population.

The first and most authoritarian pillar in this complex protection architecture is the firm legislative intervention of national governments and supranational bodies, the only entities invested with the public power to impose coercive legal limits and clear sanctions against technological abuses. The legal regulations specifically designed for the artificial intelligence era must go beyond the declaration of good intentions and encode precise obligations relating to the protection of personal data, absolute confidentiality, the prohibition of automated discrimination, ensuring algorithmic transparency and establishing clear lines of legal responsibility for damages caused by automated decisions.

An example of a historical reference in this direction is the General Data Protection Regulation (GDPR) adopted by the European Union, a pioneering legal instrument that has redesigned global standards for privacy protection and established extremely strict guidelines for the processing of personal information used in the training and running of applications based on artificial intelligence. The GDPR enshrined fundamental principles such as the right of the person not to be subject to a decision based solely on automated processing, including profiling, if it produces legal effects or affects it similarly.

However, the galloping evolution of generative models and autonomous systems requires a rapid transition to more specialized legal frameworks, such as the recent European Artificial Intelligence Regulation (EU AI Act), a revolutionary regulatory act that ranks algorithmic systems according to the level of risk they pose to citizens' rights and safety, completely banning unacceptable practices such as social score mass biometric surveillance or cognitive-behavioral manipulation and imposing draconian requirements of compliance, transparency and prior auditing for high-risk classified applications in the fields of medical, justice, human resources or critical infrastructure management.

Beyond the formal and often rigid constraints of written law, a crucial importance in the preservation of human rights lies in the internal conduct and moral responsibility of the technological community, made up of industry companies, researchers, data engineers and software developers, who have the duty to conceive and promote sound ethical guidelines throughout the life cycle of any algorithmic product.

This ethical dimension should not be seen as a mere superficial exercise in corporate image or as a form of regulatory evasion, but as a rigorous technical and philosophical discipline transposed directly into the source code and systems architecture through modern concepts such as „Ethics by Design” and „Privacy by Design”.

The responsible development of artificial intelligence requires an unwavering adherence to fundamental principles such as distributive fairness, decisional transparency, fiduciary responsibility, privacy protection and unconditional respect for human dignity, which are absolutely essential to mitigate the major risks related to the emergence of algorithmic biases and prejudices. Practically, transposing these principles requires abandoning opaque „black box” models and investing heavily in the development of explainable artificial intelligence systems, which allow end-users, independent experts and supervisors to clearly understand the logical reasoning and variables considered behind an automated decision, whether it is the refusal of a bank loan, the statement said selecting an employment profile or diagnosing a medical condition.

Industry organizations must also establish strict internal protocols for independent auditing of training data, ensuring that the datasets used are representative, properly

sampled and free from historical distortions that could perpetuate or amplify systemic discrimination related to race, gender, political orientation, ethnicity or socio-economic status.

The third pillar, indispensable and perfectly complemented by the first two dimensions, focuses on the social and community plan, directly targeting public awareness, digital literacy and active civic engagement of the population in the mechanisms of technological governance. Educating the general public about the complex implications of artificial intelligence on fundamental rights is a vital democratic shield in the face of attempts at mass behavioral manipulation and abuses of power by either the great monopolistic actors in technology or authoritarian state entities.

When citizens have a sound conceptual understanding of how social platform recommendation algorithms work, the major risks associated with systematic disinformation through deepfake technologies, or the invisible way in which their digital traces are collected and monetised, they become able to exercise their rights firmly, protect information self-determination, and hold political and economic decision-makers accountable.

Promoting an inclusive public dialogue and stimulating civil society participation in public policy-making on artificial intelligence ensures a genuine democratisation of technological progress, transforming the citizen from a mere passive and vulnerable consumer of technology into a sovereign co-creator of his own digital future.

This collective mobilization and involvement is fundamental to prevent deep socio-technological faults from occurring and to ensure that the colossal economic and social benefits of artificial intelligence will not be confiscated by a small group of interests, but will be distributed fairly, serving the common public welfare, strengthening social cohesion and unabatedly defending the inalienable freedoms of each individual.

3. NORMATIVE DELIMITATIONS - PROHIBITED PRACTICES IN THE USE OF ARTIFICIAL INTELLIGENCE ACCORDING TO ART.5 OF AI ACT³

The regulatory architecture of the European regulation on artificial intelligence shall, by means of the provisions of Article 5, establish (Duțu, 2025), a genuine absolute prohibitive regime with regard to technological systems likely to give rise to disproportionate interference in the exercise of fundamental rights and freedoms. From a teleological perspective, the text crystallizes an increased protection of the autonomy of will and the inner forum of the subject of law, sanctioning the placing on the market or use of those systems that vitiate the cognitive decision-making process (Maior, 2026). In this respect, the provisions of Article 5 (1) (a) prohibit *de jure* the use of imperceptible subliminal techniques or diligent maneuvers designed to circumvent the conscious filter of the person, thereby altering free will and generating material or moral material damage.

This paradigm of the protection of invalid consent is supplemented organically by point (b) of the same paragraph, which criminalizes at administrative level the abusive exploitation of the vulnerability state of certain categories of persons –, whether it stems from reasons of age of limited physical or mental capacity or of precarious socio-economic

³ https://eur-lex.europa.eu/legal-content/RO/TXT/PDF/?uri=OJ:L_202401689

conditions – whenever such an AI system tends to hijack the subject's conduct and cause an unfair injury.

Beyond individual guardianship, the Community legislator has established peremptory legal barriers to prevent abuses derived from algorithmic analysis at the macro-social level and from criminal determinism. Article 5 (1) (c) thus suppresses the possibility of implementing social scoring systems, prohibiting the classification of citizens based on their behavioural history or on the extrapolation of personality traits, a practice whose illegality derives from the arbitrary application of unfavourable legal or social treatments, by circumventing the principle of proportionality and contextuality of data. At the same time, in matters of criminal justice and the work of judicial bodies, point (d) proscribes the use of artificial intelligence systems for the purpose of assessing or predicting individual criminogenic risk solely on the basis of automated profiling. The prohibition of this instrumentation of predictive justice bases the imperative of observing the presumption of innocence and the individualization of the act, stating that the algorithm cannot substitute the concrete assessments of the criminal investigation body or of the court, but can, at most, support a human analysis based on objective factual elements, verifiable and circumscribed to the objective side of a criminal illicit.

The guarantees established in the sphere of privacy, protected including in the light of Article 8 of the European Convention on Human Rights, acquire new valences through strict regulation of biometric data processing and emotional intrusions.

Article 5 (1) (e) enshrines the protection of the right to image and personal data by expressly prohibiting the non-selective collection (scraping practice) of facial images from open sources or closed-circuit television networks (CCTVs) for the purpose of aggregating databases operation that blatantly violates the principle of data minimization provided by the General Data Protection Regulation. On the other hand, the emotional forum is protected against intrusions by the provisions of point (f), which qualify as unlawful the automatic deduction of affective states in employment reports and educational premises, except for strict derogations motivated by higher medical interests or public safety. In addition to this framework, point (g) prohibits biometric categorisation operations intended to infer personal data with a special protection regime (in the sense of Article 9 GDPR), thereby blocking any attempt to algorithmically infer racial origin, political beliefs, trade union membership, religious orientation or sexual life, elements related to the hard core of human identity.

The legal regime of real-time remote biometric identification, circumscribed to publicly accessible spaces and actions of law enforcement bodies, is governed by the rule of the ban on plano, the derogations being of strict interpretation and application (exception est strictissimae interpretationis). According to point (h) of Article 5 (1), these systems may be operationalized exclusively for reasons of last ration: the detection of victims of crimes of deprivation of liberty or trafficking in human beings, the removal of a state of imminent danger to life or the prevention of a terrorist attack, as well as the location of suspects targeted by warrants for offences whose penalty limit is at least four years. Even in the presence of these situations of urgent need, paragraph (2) establishes a rigorous proportionality test, in conjunction with imperative procedural obligations. Operationalization is strictly conditional on teleological, spatial and temporal limitation, requiring a prior assessment of the impact on fundamental rights under Article 27 and registration of the system in the EU database in accordance with Article 49. Exceptions to

the obligation to carry out the prior procedure are allowed exclusively as interim measures, in duly justified cases of urgency and subject to immediate administrative regularisation, thus ensuring the rule of law in the face of technological intrusions.

4. CONCLUSIONS

The accelerated evolution of artificial intelligence has demonstrated that this technology has surpassed the purely theoretical or computing tool stage, becoming a structural force that profoundly reconfigures human interactions, economic systems, democratic institutions and the very essence of human rights. While it promises tremendous progress and efficiency in multiple sectors, the implementation of algorithmic decision-making systems also generates major systemic risks to individual dignity, equity and fundamental civil liberties, requiring an urgent reassessment of existing legal instruments.

Thus, systems based on massive volumes of data can perpetuate historical prejudices and inequities, altering the substance of the right to equal treatment, but also freedom of expression through algorithmic censorship or the propagation of disinformation. In this context, the defence of fundamental rights is an existential necessity of the rule of law, the management of these challenges assuming an indissoluble synergy between three central pillars: the coercive legislative intervention of governments, the ethical development assumed responsible by the technological industry through concepts such as „Ethics by Design” and the education of the general public for active civic engagement. In order to guarantee the right to a fair trial and to protect self-determination, it becomes imperative to abandon opaque decision models of type „black box” in favour of developing an explainable artificial intelligence, capable of logically justifying automated decisions, together with the obligation to maintain a significant human control over any system that produces major legal effects.

As regards legislative intervention, the provisions of Article 5 of the new European Artificial Intelligence Regulation function as a genuine absolute prohibitive regime, constituting a permanent legal shield against systems likely to generate disproportionate interference. Within the scope of legislative intervention, the provisions of Article 5 of the new European Artificial Intelligence Regulation function as a genuine absolute prohibitive regime, constituting a permanent legal shield against systems likely to generate disproportionate interference.

This law prohibits *de jure* the marketing or use of systems that vitiate consent and cognitive decision making by resorting to subliminal imperceptible techniques or to abusive exploitation of the vulnerability of certain categories of people.

Beyond individual guardianship, the Community legislature has established peremptory legal barriers to prevent macro-social abuses, definitively suppressing the possibility of implementing social scoring systems based on the extrapolation of personality traits and prohibiting the use of artificial intelligence to assess criminogenic risk based solely on automated profiling, thus blocking dangerous criminal determinism. The scope of privacy and privacy also receives increased protection, with regulations prohibiting non-selective collection (scraping) of open source facial images for aggregation of databases, as well as automated deduction of affective states in educational spaces and in the workplace. Biometric categorization operations intended to infer

sensitive data with a special protection regime, such as racial origin, political beliefs or sexual life, related to the hard core of human identity are also strongly prohibited.

As regards real-time remote biometric identification in publicly accessible spaces, it is governed by the rule of the plano prohibition, being tolerated exclusively for reasons of last ratio, such as the prevention of a terrorist attack or the detection of victims of serious crimes, its operationalisation being strictly conditioned by a rigorous proportionality test and imperative prior evaluation procedures

REFERENCES:

- Bodea, B., 2024. *Drepturile omului și libertățile fundamentale*. 35-45 ed. București: Hamangiu.
- Bîrsan, C., 2006. *Convenția europeană a drepturilor omului. Comentariu pe articole*. București: C.H.Beck.
- Coman, R. M., 2024. *Integrarea inteligenței artificiale în colectarea și analiza probelor digitale în procesul penal*. București, Universul Juridic, pp. 152-54.
- Duțu, M., 2024. [Online] Available at: <https://www.juridice.ro/essentials/7647/primul-tratat-international-privind-ia>. [Accessed 2026].
- Duțu, M., 2025. *Elemente de dreptul inteligenței artificiale*. București: Universul Juridic.
- Duțu, M., n.d. [Online] Available at: <https://www.juridice.ro/essentials/7647/primul-tratat-international-privind-ia>. [Accessed 2026].
- Maior, D. A., 2026. *De la psihoputere la practici interzise: protecția libertății cognitive în contextul art. 5 din Regulamentul AI și al jurisprudenței CEDO*. Craiova, Universitaria Craiova.
- Manu, G., 2023. *Impactul digitalizării asupra dreptului public*. București, Universul Juridic, p. 95.
- Moroianu, A., 2020. *Drepturile omului în era inteligenței artificiale*. *Drepturile omului*, Volume 2, p. 75.
- Pop, C. T., 2024. *Perspective ale utilizării inteligenței artificiale în domeniul justiției constituționale*. București, Universul Juridic, pp. 37-39.
- Vartires, D., 2023. *Impactul inteligenței artificiale asupra domeniului juridic*. Digitalizarea sistemului judiciar din România. In: *In honorem Brândușa Ștefănescu. Interpretarea și aplicarea dreptului. Supremație, prioritate, preeminență și subsidiaritate*. București: Hamangiu.
-
-
-