# DIMENSIONS OF ARTIFICIAL INTELLIGENCE ETHICS FROM AN INTERNATIONAL AND EU PERSPECTIVE

**A. N. PURAN**

**Andra Nicoleta Puran**

National University of Science and Technology Politehnica Bucharest, Pitești University Centre, Faculty of Economic Sciences and Law, Romania
https://orcid.org/0000-0002-8773-1548, E-mail andradascalu@yahoo.com

*Abstract: Today, we live in a time when humanity is marked by an extremely complex political and economic context, by chaos, crises, but also by a fulminant development of science and, above all, of artificial intelligence (AI). In the period of the "fourth industrial revolution" (or "(r)evolution AI"), we must be aware that, in addition to the multitude of advantages it brings to humanity, it also generates multiple ethical problems, given the lack of a clear legal framework at Union and international level. In the face of these challenges, fortunately, the reactions did not take long to appear at the international level, at the European level and of the EU institutions. This study proposes a brief reflection on the issue of ethical principles that have been taking shape in recent years at the European and international level in the field of AI, based on the fundamental idea of promoting a human-centered AI that serves the common good, that respects fundamental rights and general ethical values.*

*Keywords: artificial intelligence, principles of ethics, fundamental rights, international perspective, union perspective.*

## 1. Introduction

Because today we live among digital assistants (Siri, Alexa, Cortana), autonomous cars, smart cameras with facial recognition or systems capable of making predictions regarding future behaviors, fifth generation wireless technologies (5G), medical devices remotely guided through technology, robots that save lives, etc., this paper aims to bring to attention several aspects of the AI Ethics Dimension from an international and EU perspective.

It is undeniable that Artificial Intelligence is a reality, a way of life installed subtly and imperceptibly in our behavior, a reality that generates both undeniable positive effects for humanity, but also negative ones, especially in conditions where the development and use of AI takes place without a stable and transparent legal framework. (Duminică & Ilie, 2023).

Artificial intelligence systems promise widespread benefits for society, while also presenting substantial risks in almost all sectors. Over the past few years, a number of different groups and initiatives have tried to outline and agree on principles to guide the application of AI in society, but the lack of a clear legal framework at EU and international level "generates ambiguity and, proportionately, risks to social organisation [...]" (Drăgușin, 2020).

Thus, both at the international and EU level, the foundations of an applied ethics are already being laid, known as Cyber-ethics and the ethics of artificial intelligence, also called robo-ethics.

Cyber-ethics is considered to be a set of rules that "aim to prevent abuse or violation of human rights as a result of various operations on computer media, data transfer, the protection of human dignity, privacy and intellectual property through the extensive use of computer networks" (Crăciun & Jianu, 2019).

Artificial intelligence ethics, also known as robo-ethics, is a multidisciplinary discipline made up of the totality of moral values and principles that guide the creation, development and use of artificial intelligence with the aim of optimizing the beneficial impact of AI and, simultaneously, reducing its risks and negative consequences.

Even though cyber-ethics and the ethics of artificial intelligence are sometimes analyzed distinctly, they are nothing more than aspects of the same type of problems, caused by the interaction between humans, their personality, their consciousness, and new technologies. (Crăciun & Jianu, 2019).

As a result of the international and EU-level efforts to outline the legal framework for the use of AI, on May 4, 2023, in Romania, the Romanian Committee for Artificial Intelligence was established. The main objective of the Committee stipulated in art. 3 paragraph 2 of Ministerial Order no. 20,484/2023 (published in the Official Gazette, Part I no. 382 of May 4, 2023) is to create an artificial intelligence ecosystem based on excellence, trust and respect for ethical principles, coherent, efficient and sustainable, which generates added value in social and economic terms through the use of artificial intelligence technologies. Several structures, councils and working groups will operate within this Committee, including the Scientific and Ethics Council for Artificial Intelligence, which will provide the Government of Romania with scientific advice on the responsible and ethical use of artificial intelligence.

## 2. Research
### 2.1. AI ethics at the international level

The 193 member states of the U.N., within the General Conference of the United Nations Educational, Scientific and Cultural Organization (UNESCO), meeting in Paris in its 41st session, recognizing the positive but also negative impact of artificial intelligence (AI) on societies, the environment, ecosystems and human lives, including the human mind, adopted, on November 23, 2021, the Recommendation on the Ethics of Artificial Intelligence.

This document thus becomes the first global agreement on AI ethics, a first "set" of common, universal rules, stemming from the need to regulate a rapidly expanding field worldwide. The stated goal was to define policies and principles that would form the basis for the functioning of AI systems for the good of humanity, individuals, societies, the environment and to prevent harm. At the same time, it was intended to be a globally accepted normative instrument, focusing not only on identifying values and principles, but also on their practical implementation.

The foundation of this Recommendation lies in the instruments of the international human rights framework, namely: the Universal Declaration of Human Rights (1948), the Convention relating to the Status of Refugees (1951), the Convention on Discrimination (1958), the International Convention on the Elimination of All Forms of Racial Discrimination (1965), the International Covenant on Civil and Political Rights (1966), the International Covenant on Economic, Social and Cultural Rights (1966), etc.

Given that artificial intelligence technologies can be beneficial to humanity, but also raise fundamental ethical concerns, which can lead to discrimination, inequality, digital divides, exclusion and a threat to cultural, social and biological diversity and social or economic divisions, the purpose of the Recommendation is found in the idea that AI must be human-centered and serve the interest of humanity.

The objectives set by the Recommendation (UNESCO, 2021) consist of:

(a) to provide a universal framework of values, principles and actions to guide States in formulating legislation, policies or other instruments on AI, in accordance with international law;

(b) guide the actions of individuals, groups, communities, institutions and private sector companies to ensure the incorporation of ethics into all stages of the AI system life cycle;

(c) protect, promote and respect human rights and fundamental freedoms, human dignity and equality, including gender equality, in order to safeguard the interests of present and future generations, to conserve the environment, biodiversity and ecosystems; and to respect cultural diversity at all stages of the AI system's life cycle;

(d) encourage multidisciplinary and pluralistic dialogue and consensus building on ethical aspects related to AI systems;

(e) promote equitable access to developments and knowledge in the field of AI and the sharing of benefits, paying particular attention to the needs and contributions of low- and middle-income countries, including least developed countries, landlocked developing countries and small island developing States.

The Recommendation (UNESCO, 2021) enshrines four values that can be considered the foundations of artificial intelligence ethics:

a) respect, protection and promotion of human rights and fundamental freedoms and human dignity.

The enshrinement of human rights and freedoms in international legal instruments and in modern democratic constitutions is the result of a long historical process that includes philosophical, legal and political conceptions, theories, doctrines, but also different forms of recognition and normative consecration of them. Fundamental human rights and freedoms and human dignity are a constitutional reality, with profound implications for the existence of each human being, also representing an existential reality of each person, of society as a whole and a dimension of democracy. Thus, the respect, protection and promotion of human dignity and rights, as established by international law, are essential throughout the life cycle of AI systems.

According to the Recommendation, "new technologies should provide new means to uphold, defend and exercise human rights, not to violate them."

b) sustainable development of the environment and ecosystems.

Sustainable development, that development that meets the needs of the present without compromising the ability of future generations to meet their own needs, must be promoted throughout the life cycle of AI systems.

c) ensuring diversity and inclusion.

The Recommendation states that this objective can be achieved by promoting the active participation of all individuals or groups, without discrimination on any grounds.

d) ensuring living in peaceful, just and interconnected societies. "This value requires that peace, inclusion and justice, equity and interconnectedness are promoted throughout the life cycle of

AI systems, to the extent that the processes of the life cycle of AI systems should not lead to segregation, objectify or undermine the freedom and autonomous decision-making, as well as the safety of human beings and communities, divide and turn individuals and groups against each other, or threaten the coexistence of humans, other living beings and the natural environment." (UNESCO, 2021)

At the same time, in addition to these values, the Recommendation outlines ten principles (UNESCO, 2021) of AI ethics, namely:

a. proportionality and non-maleficence

b. safety and security

c. fairness and non-discrimination

d. sustainability

e. right to privacy and data protection

f. oversight and human determination

g. transparency and explainability

h. responsibility and ownership

i. AI awareness and literacy

j. adaptive and multi-stakeholder governance and collaboration

In summary, this Recommendation represents a first step in creating a universal framework of values, principles and actions to guide States in drafting legislation, policies or other instruments on AI, in compliance with international law, while protecting human rights and fundamental freedoms, human dignity and equality, including gender equality. The objective of the Recommendation is to protect the interests of present and future generations, to conserve the environment, biodiversity and ecosystems and to respect cultural diversity at all stages of the life cycle of the AI system, encouraging multidisciplinary dialogue and consensus-building on ethical aspects related to AI systems (Duminică & Ilie, 2023).

UNESCO does not stop there, however, so among its latest actions in this field we point out the initiative of March 30, 2023 to once again ask governments to implement much more solid ethical rules in the field of AI. The reaction came as a result of the call of over 1,000 experts who called for a moratorium on new AI systems, the main issues raised being discrimination and stereotypes, as well as disinformation, personal data protection, human rights and those regarding environmental protection (www.agerpres.ro, Unesco calls for more solid ethical rules regarding artificial intelligence, 2023).

### 2.2. The Union and European perspective on AI ethics

At the European Union level, given that it is an essential part of the digital single market strategy, artificial intelligence (AI) and its regulation are currently on the list of priorities, and in this regard we recall the European Parliament Resolution of February 16, 2017 containing recommendations addressed to the Commission regarding civil law rules on robotics (European Parliament Resolution, 2017), this being a first alarm signal regarding the fulminating developments of AI that must be subsumed under ethical and legal principles.

In 2018, the European Commission adopted the "Communication on Artificial Intelligence" (European Commission, Communication 237/2018), which aimed at numerous

measures to develop technological capabilities at EU level, recognizing that the EU was lagging behind in AI investment compared to Asia or the US.

Also in 2018, the European Commission established the High-Level Expert Group on Artificial Intelligence (Al HLEG), an independent group of experts to support the implementation of the European Strategy on Artificial Intelligence.

From the group's perspective, trustworthy AI has three components that ideally should work in harmony and overlap and should be fulfilled throughout the system's lifecycle: (a) AI should be legal, complying with all applicable laws and regulations; (b) it should be ethical, ensuring compliance with ethical principles and values; (c) it should be sound, both from a technical and social perspective, because even with good intentions, AI systems can cause unintended harm. The expert group also identified four ethical principles that should underpin trustworthy AI: respect for human autonomy; prevention of harm; fairness and explainability. These rules are complemented by seven requirements considered key requirements for achieving trustworthy AI: human involvement and oversight; technical robustness and security; data confidentiality and governance; transparency; diversity, non-discrimination and fairness; environmental and societal well-being; responsibility (High-Level Expert Group on Artificial Intelligence).

Also in the field of AI ethics, it is worth mentioning that on October 20, 2020, the European Parliament adopted the Resolution containing recommendations addressed to the Commission on the framework of ethical aspects associated with artificial intelligence, robotics and related technologies (European Parliament Resolution, 2020), which outlines the following ethical rules:
- artificial intelligence, robotics and related technologies should be human-centered, created and controlled by humans;
- mandatory conformity assessment of high-risk artificial intelligence, robotics and related technologies;
- security, transparency and accountability;
- safeguards and remedies against bias and discrimination;
- the right to compensation for damage;
- social responsibility and gender equality in artificial intelligence, robotics and related technologies;
- artificial intelligence, robotics and related green and sustainable technologies;
- respect for privacy and limitation of the use of biometric recognition;
- good governance in the field of artificial intelligence, robotics and related technologies, including data used or produced by such technologies.

Subsequently, on 9 March 2021, the European Commission adopted the communication entitled Digital Compass 2030: a European blueprint for the digital decade (European Commission, Communication 118/2021), which aimed to provide a vision of what a successful digital transformation means by 2030. This context created the premises for the initiation on 21 April 2021 of the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Regulation) and amending certain Union legislative acts, with the declared aim of transforming the EU into a global hub for trustworthy AI. The proposal for a regulation (Proposal for a Regulation of the

European Parliament and of the Council on harmonised rules on artificial intelligence) is intended to support the development and adoption of trustworthy, human-centric AI, while ensuring the health, safety and fundamental rights of people in line with Union values and principles.

Last but not least, at the European level, the adoption in 2018 by the European Commission for the Efficiency of Justice of the Council of Europe of the first European document establishing ethical principles for the use of artificial intelligence in judicial systems is of particular importance in establishing the principles of AI ethics. This is the "European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment" (2018), which lists the following key principles aimed at helping policymakers, legal professionals and private sector companies to ensure that the use of artificial intelligence respects human rights, privacy and data protection:

1. *Principle of respect for fundamental rights:* the design and implementation of artificial intelligence tools and services are compatible with fundamental rights.
2. *The principle of non-discrimination:* specifically prevents the development or intensification of any discrimination between individuals or groups of individuals.
3. *The principle of quality and security:* with regard to the processing of court decisions and data, certified sources and intangible data with models developed in a multidisciplinary manner will be used, in a secure technological environment.
4. *Principle of transparency, impartiality and fairness:* data processing methods shall be accessible and easy to understand, external audits shall be authorized.
5. *The "user-controlled" principle:* excludes a prescriptive approach and ensures that users are informed actors and control the choices made.

## 3. Conclusions

Like any other science, ethics is constantly changing and its principles must be adapted to the times and the field so that we can benefit from all the advantages offered to humanity by AI, without endangering security, human rights, and the well-being of societies.

Without contesting the advantages of artificial intelligence, although we welcome the efforts being made at international and European level to lay the foundations of AI ethics, we nevertheless consider that it is not sufficient to outline sets of principles that can sometimes be too general, can be interpreted differently or can sometimes come into conflict when applied to concrete situations.

## REFERENCES

1. Crăciun, D. & Jianu, F. (2019). Aspecte de etică în dezvoltarea și utilizarea inteligenței artificiale în rețelele electrice de distribuție (Ethical aspects in the development and use of artificial intelligence in electricity distribution networks). *EMERG* 5(9), DOI: 10.37410/EMERG.2019.09.03
2. Drăgușin, C.M. (2020). Stadiul actual al eticii în domeniul inteligenței artificiale (The current state of ethics in the field of artificial intelligence). *Dreptul* 3
3. Duminică, R. & Ilie, D.M. (2023). Ethical and legal aspects of the development and use of robotics and artificial intelligence. Protection of human rights in the era of globalization and digitisation. *Journal of Law and Administrative Sciences* 19

4. UNESCO. (2021). *Recommendation on the ethics of artificial intelligence,* https://www.cnr-unesco.ro/uploads/media/f1077_recomandari-unesco-ai-site.pdf accessed on 11.04.2024.

5. *UNESCO calls for stronger ethical rules regarding artificial intelligence*, https://www.agerpres.ro/zig-zag/2023/03/31/unesco-cere-reguli-de-etica-mai-solide-privind-inteligenta-artificiala--1085125 accessed on 11.04.2024.

6. European Parliament Resolution.(2017) containing recommendations to the Commission on civil law rules on robotics, https://www.europarl.europa.eu/doceo/document/TA-8-2017-02-16_RO.html#sdocta9 accessed on 11.04.2024.

7. European Commission, *Communication 237/2018 from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Artificial Intelligence for Europe*, https://eur-lex.europa.eu/legal-content/RO/ALL/?uri=CELEX%3A52018DC0237, accessed on 11.04.2024.

8. High-Level Expert Group on Artificial Intelligence, *Ethical guidelines for trustworthy artificial intelligence*, https://op.europa.eu/ro/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1 accessed on 11.04.2024.

9. European Parliament Resolution (2020) with recommendations to the Commission on the framework of ethical aspects associated with artificial intelligence, robotics and related technologies, *OJ C 404, 6.10.2021, p. 63–106*, https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A52020IP0275 accessed on 11.04.2024.

10. European Commission, *Communication 118/2021 from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Digital Compass 2030: The European blueprint for the digital decade, https://eur-lex.europa.eu/legal-content/ro/ALL/?uri=CELEX:52021DC0118*, accessed on 11.04.2024.

11. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Regulation) and amending certain Union legislative acts, https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX:52021PC0206, accessed on 11.04.2024.

12. European Commission for the Efficiency of Justice (CEPEJ), *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*, 2018, p. 7, https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c, accessed on 11.04.2024.