

## REFLECTIONS ON ARTIFICIAL INTELLIGENCE ETHICS PRINCIPLES

**Andra Nicoleta Puran<sup>1</sup>**

Politehnica University of Bucharest

**Ramona Duminică<sup>2</sup>**

Politehnica University of Bucharest

**Diana Maria Ilie<sup>3</sup>**

Politehnica University of Bucharest

DOI: <https://doi.org/10.62140/APRDDI92024>

**Abstract:** In today's society, characterized by rapid technological and informational changes and progress, frequently subject to major ethical and moral challenges, it becomes essential to base the development and use of artificial intelligence (AI) on the observance of ethical values. Both history and the experience of our days prove to us that most of the problems that humanity encounters cannot be solved without appealing to ethics.

We are crossing the so-called "fourth industrial revolution" or the "(r)evolution of AI", which is shaping its supremacy through interdisciplinary and cutting-edge technological advances. But, beyond the multitude of advantages it brings to humanity, it generates multiple moral and ethical problems. In this context, this paper proposes a 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.

The general opinion is that of promoting a human-centred AI that serves the common good, that respects fundamental rights and general ethical values, such as: fairness, non-discrimination, confidentiality and autonomy. This is just a starting point, the road to the implementation of these principles in real life is a long and rather arduous one, considering the generality that characterizes any principle and the dissensions that may appear in their regulation, dissensions mainly caused by fierce global competition for AI supremacy.

**Keywords:** ethics, principles, artificial intelligence, fundamental rights and freedoms, Union perspective, international perspective.

## I. Introduction

---

<sup>1</sup>Lecturer PhD., The National University of Science and Technology POLITEHNICA Bucharest, Pitești University Centre, Faculty of Economic Sciences and Law, Romania, e-mail [andradascalu@yahoo.com](mailto:andradascalu@yahoo.com), ORCID:0000-0002-8773-1548.

<sup>2</sup>Lecturer PhD., The National University of Science and Technology POLITEHNICA Bucharest, Pitești University Centre, Faculty of Economic Sciences and Law, Romania, e-mail [duminica.ramona@gmail.com](mailto:duminica.ramona@gmail.com), ORCID:0000-0002-3416-2329.

<sup>3</sup>PhD., Legal advisor, The National University of Science and Technology POLITEHNICA Bucharest, Pitești University Centre, Legal Department, Romania, e-mail [dianamaria.ilie@yahoo.com](mailto:dianamaria.ilie@yahoo.com), ORCID: 0009-0009-2487-1648

In a period in which humanity is marked by ambiguity, chaos, fragmentation, crises, in an extremely complex political and economic context, artificial intelligence (AI) has developed at a very alert pace that seems unstoppable.

Today, we live among digital assistants (Siri, Alexa, Cortana), autonomous cars, smart cameras with facial recognition or systems capable of making predictions about future behavior, fifth generation wireless technologies (5G), medical devices guided remotely through technology, lifesaving robots etc.

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 to have a stable and transparent legal framework (Duminică, Ilie, 2023, p. 21). The lack of a clear legal framework at the Union and international level “generates ambiguity and, proportionally, risks to social organization (...)” (Drăgușin, 2020).

## **II. The need to return to ethics**

As happened in other stages of human development, for example at the time of the discovery of nuclear energy, the question naturally arises whether or not the advantages brought by artificial intelligence outweigh its disadvantages for man and his living environment. Could AI even become a threat to the survival of the human species? How do we integrate a field like AI into the sometimes sensitive realities of life? How do we ensure that technology serves society and the individual? How do we protect human rights in the age of AI?

It is obvious that we are crossing the so-called “fourth industrial revolution” or the “Evolution of AI”, in which we have stepped, with no way back, into a new realm of practical applications among the most innovative, in which the challenges and questions of order moral and legal are proportional to the successes achieved. In the field of AI, the greatest risks can arise from the inadequate approach to ethical values, therefore the solution of new problems can only be done by returning to ethics. It is not for nothing that ethics is considered one of the oldest and most relevant disciplines in world history, at a time when, as the American theologian Carl F.H. Henry, “ethics has become an incisive and universal requirement for the survival of humanity as a species.” (Henry, 2006, p. 16)

In this context, the foundations of an applied ethics, known as Cyber-ethics and the ethics of artificial intelligence, also called *robo-ethics*, are already being laid.

Cyber-ethics is considered to be a set of rules that “aims to prevent the abuse or violation of human rights as a result of various IT operations, data transfer, the protection of human dignity, private life and intellectual property through the extensive use of computer networks” (Crăciun, Jianu, 2019, p. 49).

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

Even if cyber-ethics and the ethics of artificial intelligence are sometimes analysed separately, they are nothing but aspects of the same type of problems, caused by the interaction between man, his personality, his consciousness and new technologies (Crăciun, Jianu, 2019, p. 49).

In the face of these challenges, fortunately, the reactions did not take long to appear and are underway at the international level, at the European level and of the EU institutions.

As for Romania, it declared itself open to the visions promoted at the international and EU level from the beginning, it welcomed and took note of UNESCO's recommendations for ethics in AI and it follows the directions laid out step by step by the EU institutions. Thus, on May 4, 2023, the Romanian Committee for Artificial Intelligence was established in our country<sup>4</sup>. The main objective of the Committee stipulated in art. 3 paragraph 2 of O.M. no. 20.484/2023 is the creation of an artificial intelligence ecosystem based on excellence, trust and compliance with ethical principles, coherent, performing 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 Ethical Council of Artificial Intelligence, which will provide the Romanian Government with scientific advice on the responsible and ethical use of artificial intelligence.

### **III. The international view on the principles of AI ethics**

On November 23, 2021, during the General Conference of the United Nations Educational, Scientific and Cultural Organization (UNESCO), gathered in Paris in the 41<sup>st</sup> session, the Recommendation on the ethics of artificial intelligence was adopted. On this

---

<sup>4</sup>Order of the Ministry of Research, Innovation and Digitization no. 20.484/2023 regarding the establishment, organization and operation of the Romanian Committee for Artificial Intelligence, published in the Official Monitor of Romania, Part I no. 382 of May 4, 2023.

occasion, the 193 member states recognized the positive but also negative impact of artificial intelligence (AI) on societies, on the environment, ecosystems and human lives, including on the human mind.

This document thus becomes a first global agreement in AI ethics, a first “set” of common, universal rules, born from the need to regulate a booming field worldwide. The stated purpose was to define policies and principles that would constitute a basis for the operation of AI systems for the good of humanity, individuals, societies, the environment and to prevent damage. At the same time, it was aimed for it to be a globally accepted normative instrument, focusing not only on the identification of values and principles, but also on their practical implementation.

The core of the Recommendation lies in the idea that AI should be human-centered and serve the interest of humanity. This Recommendation was based on the instruments of the international human rights framework, namely: the Universal Declaration of Human Rights (1948), the Convention on 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.

From the content of the Recommendation, four values<sup>5</sup> emerge that can be considered the foundations of the ethics of artificial intelligence:

- a. respecting, protecting and promoting human rights and fundamental freedoms and human dignity
- b. sustainable development of the environment and ecosystems
- c. ensuring diversity and inclusion
- d. ensuring living in peaceful, just and interconnected societies.

At the same time, in addition to these values, the Recommendation outlines ten principles<sup>6</sup> of AI ethics, respectively:

- a. proportionality and the prohibition of doing harm
- b. safety and security
- c. fairness and non-discrimination
- d. durability
- e. the right to privacy and data protection

---

<sup>5</sup>UNESCO, *Recommendation on the ethics of artificial intelligence adopted on 23 November 2021*, [https://www.cnr-unesco.ro/uploads/media/f1077\\_recomandari-unesco-ai-site.pdf](https://www.cnr-unesco.ro/uploads/media/f1077_recomandari-unesco-ai-site.pdf), accessed on 27.11.2023.

<sup>6</sup>UNESCO, *Recommendation on the ethics of artificial intelligence adopted on 23 November 2021*, [https://www.cnr-unesco.ro/uploads/media/f1077\\_recomandari-unesco-ai-site.pdf](https://www.cnr-unesco.ro/uploads/media/f1077_recomandari-unesco-ai-site.pdf), accessed on 27.11.2023.

- f. human oversight and determination
- g. transparency and explainability
- h. responsibility and accountability
- i. AI awareness and literacy
- j. adaptive and multi-stakeholder governance and collaboration.

Summarizing, through this Recommendation a first step was achieved in creating a universal framework of values, principles and actions to guide states in drafting legislation, policies or other instruments regarding AI in compliance with international law, being protected 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 in all stages of the life cycle of the AI system, encouraging multidisciplinary dialogue and building consensus on ethical aspects related to AI systems (Duminică, Ilie, 2023, pp. 37-38).

However, UNESCO does not stop here, so among its latest actions in this field we point out the initiative of March 30, 2023 to ask governments again to implement much stronger ethical rules in the field of AI.

The Director-General of the UN Educational, Scientific and Cultural Organization, Audrey Azoulay, said in Paris that “this is the challenge of our times”, and the self-regulation of this industry is not enough to avoid the ethical damage caused by the continuous development of AI systems, and it is imperative to put in application of the Recommendation adopted in November 2021. The reaction came as a result of the call of more than 1,000 experts who requested a moratorium on new AI systems, the main issues raised being discrimination and stereotypes, as well as disinformation, personal data protection, human rights and environmental protection<sup>7</sup>.

#### **IV. Union and European view on AI ethics**

An essential part of the Digital Single Market strategy, artificial intelligence (AI) and its regulation are currently also on the EU’s priority list.

Thus, the EU leaders did not miss the start, and in this sense, we recall the Resolution of the European Parliament of February 16, 2017, which contained recommendations

---

<sup>7</sup>UNESCO calls for stronger ethical rules on artificial intelligence, <https://www.agerpres.ro/zigzag/2023/03/31/unesco-cere-reguli-de-etica-mai-solide-privind-inteligenta-artificiala--1085125>.

addressed to the Commission regarding civil law rules on robotics<sup>8</sup>, this being a first alarm signal regarding the fulminant developments of AI that must be subordinated to ethical and legal principles.

Later, in 2018, the European Commission adopted the “Communication on artificial intelligence”<sup>9</sup>, which aimed at numerous measures to develop technological capabilities at the EU level, recognizing that the EU was lagging behind in the field of AI investments compared to Asia or the US.

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

From his perspective, reliable AI features three components that ideally must work in harmony and overlap, and should be met throughout the system’s lifecycle 🙄 a) AI should be lawful, complying with all applicable laws and regulations; (b) it should be ethical, ensuring adherence to ethical principles and values (c) it should be solid, both technically and socially, because even if well-intentioned, AI systems can cause unintended harm. The group of experts also identified four ethical principles that must underpin a reliable AI: respect for human autonomy; damage prevention; equity and explainability. These rules are complemented by seven requirements considered key requirements for the realization of trustworthy AI: human involvement and supervision; technical robustness and safety; data privacy and governance; transparency; diversity, non-discrimination and fairness; the well-being of the environment and society; responsibility<sup>10</sup>.

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 issues associated with artificial intelligence, robotics and related technologies*<sup>11</sup>, in which the following rules of an ethical nature are outlined:

<sup>8</sup>Resolution of the European Parliament of 16 February 2017 containing recommendations addressed to the Commission regarding civil law rules on robotics, [https://www.europarl.europa.eu/doceo/document/TA-8-2017-02-16\\_RO.html#sdocta9](https://www.europarl.europa.eu/doceo/document/TA-8-2017-02-16_RO.html#sdocta9) accessed on 27.11.2023.

<sup>9</sup>European Commission, Communication 237/2018 of 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 28.11.2023.

<sup>10</sup>High Level Expert Group on Artificial Intelligence, Ethical Guidelines for Trusted Artificial Intelligence, p. 9, <https://op.europa.eu/ro/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1> accessed on 28.11.2023.

<sup>11</sup>European Parliament resolution of 20 October 2020 containing recommendations to the Commission on the framework for ethical issues 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 28.11.2023.

- artificial intelligence, robotics and related technologies to be human-centered, created and controlled by humans;
- mandatory compliance assessment of artificial intelligence, robotics and related high-risk technologies;
- safety, transparency and accountability;
- safeguards and remedies against bias and discrimination;
- the right to reparation of the damage;
- social responsibility and gender equality in artificial intelligence, robotics and related technologies;
- artificial intelligence, robotics and related green and sustainable technologies;
- respecting privacy and limiting the use of biometric recognition;
- good governance in the field of artificial intelligence, robotics and related technologies, including the data used or produced by such technologies.

Subsequently, on 9 March 2021, the European Commission adopted the Communication entitled *Compass for the Digital Dimension 2030: The European Model for the Digital Decade*<sup>12</sup>, which aimed to provide a vision of what a successful digital transformation means by 2030. This context created the premises for the initiation on April 21, 2021 of the *Proposal for a Regulation of the European Parliament and of the Council establishing harmonized rules on artificial intelligence (the Regulation on artificial intelligence)* and amending certain legislative acts of the Union, with the stated aim to turn the EU into a global hub for trusted AI. The regulation proposal<sup>13</sup> aims to support the development and adoption of trustworthy, human-centred 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 the ethical principles of the use of artificial intelligence in judicial systems is of particular importance in establishing the principles of AI ethics. It is the “*European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their*

---

<sup>12</sup>European Commission, Communication 118/2021 of the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Compass for the Digital Dimension 2030: The European Model for the Digital Decade, <https://eur-lex.europa.eu/legal-content/ro/ALL/?uri=CELEX:52021DC0118>, accessed on 28.11.2023.

<sup>13</sup>Proposal for a Regulation of the European Parliament and of the Council establishing harmonized rules on artificial intelligence (Artificial Intelligence Regulation) and amending certain Union legislation, <https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX:52021PC0206>, accessed on 28.11.2023.

*environment*<sup>14</sup> which lists the following key principles to help policy makers, legal professionals and private sector companies ensure that the use of artificial intelligence respects human rights, privacy and data protection:

1. *Principle of respect for fundamental rights*: ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights.
2. *Principle of non-discrimination*: specifically prevent the development or intensification of any discrimination between individuals or groups of individuals.
3. *Principle of quality and security*: with regard to the processing of judicial decisions and data, use certified sources and intangible data with models elaborated in a multi-disciplinary manner, in a secure technological environment.
4. *Principle of transparency, impartiality and fairness*: make data processing methods accessible and understandable, authorise external audits.
5. *Principle “under user control”*: preclude a prescriptive approach and ensure that users are informed actors and in control of the choices made.

## V. Conclusions

Although we welcome the efforts that are being made at the international and European level to lay the foundations of AI ethics, as we have shown in this paper, we still appreciate that it is not enough to outline some sets of principles that can sometimes be too general, can be interpreted differently, or at other times they may come into conflict when applying them to concrete situations. They only form the basis of AI ethics, providing simple guidelines and helping to compress complex ethical issues into a few central elements accessible to all, but they are not sufficient, but must be supplemented with a detail of how to apply in practical situations (Whittlestone et al., 2019, pp. 195-200).

Without disputing the advantages of artificial intelligence, we appreciate that caution must be exercised and we must exclude or at least limit the acceptance of possible violations of fundamental rights or abuses by accelerating the process of adopting uniform norms in the field. Ethics become all the more necessary as AI evolves, and its negative outcomes can become more and more serious and frequent. In the end, AI is a “tool” that, if not used ethically, could lead to discriminatory decisions in processes such as hiring or firing, in legal proceedings, could seriously damage the right to privacy, data protection, dignity, through

---

<sup>14</sup> 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 28.11.2023.



the use of facial recognition equipment, for online tracking and profiling of individuals, through the creation of highly realistic fake video-audio material, and the examples go on.

Like any other science, ethics is also 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 mankind by AI, without jeopardizing security, human rights and the good of societies.

## REFERENCES:

Crăciun, Daniel; Jianu, Florin- 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 electrical distribution networks), EMERG vol. 5, Issue 9 (2019), DOI: 10.37410/EMERG.2019.09.03.

Drăgușin, Claudiu Marian- Stadiul actual al eticii în domeniul inteligenței artificiale (The current state of ethics in artificial intelligence). Dreptul no. 3 (2020).

Duminică, Ramona; Ilie, Diana Maria- 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 no. 19 (2023).

Henry, Carl F.H. apud. Vasile Boari, in Peter Singer- Tratat de etică (Ethics), translation coordinated by Boari, Vasile and Mărincean, Raluca, Iași, Romania: Polirom, 2006.

Whittlestone, Jess [et al.] - The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions, Cambridge, UK: Apollo – University of Cambridge Repository, 2019, pp. 195-200, <https://doi.org/10.17863/CAM.37097>.

UNESCO, Recommendation on the ethics of artificial intelligence adopted on 23 November 2021, [https://www.cnr-unesco.ro/uploads/media/f1077\\_recomandari-unesco-ai-site.pdf](https://www.cnr-unesco.ro/uploads/media/f1077_recomandari-unesco-ai-site.pdf), accessed on 27.11.2023.

Resolution of the European Parliament of 16 February 2017 containing recommendations addressed to the Commission regarding civil law rules on robotics, [https://www.europarl.europa.eu/doceo/document/TA-8-2017-02-16\\_RO.html#sdocta9](https://www.europarl.europa.eu/doceo/document/TA-8-2017-02-16_RO.html#sdocta9) accessed on 27.11.2023.

European Commission, Communication 237/2018 of 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 28.11.2023.

High Level Expert Group on Artificial Intelligence, Ethical Guidelines for Trusted Artificial Intelligence, p. 9, <https://op.europa.eu/ro/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1> accessed on 28.11.2023

European Parliament Resolution of 20 October 2020 containing recommendations to the Commission on the framework for ethical issues 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 28.11.2023

European Commission, Communication 118/2021 of the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Compass for the Digital Dimension 2030: The European Model for the Digital Decade, <https://eur-lex.europa.eu/legal-content/ro/ALL/?uri=CELEX:52021DC0118>, accessed on 28.11.2023.

Proposal for a Regulation of the European Parliament and of the Council establishing harmonized rules on artificial intelligence (Artificial Intelligence Regulation) and amending certain Union legislation, <https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX:52021PC0206>, accessed on 28.11.2023,

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 28.11.2023.

Order of the Ministry of Research, Innovation and Digitization, no. 20.484/2023 regarding the establishment, organization and operation of the Romanian Committee for Artificial Intelligence, published in the Official Monitor of Romania, Part I no. 382 of May 4, 2023.