ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS – LEGAL CHALLENGE FOR THE EUROPEAN UNION

Asst. Prof. Jovan Andonovski
"Goce Delcev" University, North Macedonia

Abstract

Artificial Intelligence is increasingly present in our lives, reflecting a growing tendency to turn for advice, or turn over decisions altogether. On the other hand, inviolability of human life is the central idea behind human rights. Artificial Intelligence generates some challenges for human rights. The European Union acts as a passive observer in the new debate that is of great importance. Between the formal regulations and 'ethics guidelines' in the field of Artificial Intelligence, the EU has to make a decision and position itself. Argument against regulation of emerging technologies is the stifling of innovation. On the other side of this argument is the need to provide a framework within which citizens can be protected from threats to privacy, autonomy, well-being and other aspects of human rights that may be affected as technologies like artificial intelligence are increasingly incorporated into everything. Nevertheless, the EU has to adopt a legal frame that can be developed in the future, bearing in mind the benefits of the AI.

Key words

Artificial Intelligence, human rights, EU, law

Artificial intelligence – the new nornal

Artificial Intelligence (AI) is one of the new terms which we can meet frequently in our daily life. The term AI is difficult to define and has been a question of debate in the recent times. In general, AI refers to the "ability of a computer or a computer-enabled robotic system to process information and produce outcomes in a manner similar to the thought process of humans in learning, decision making and solving problems" (PWC, 2017: 2).

The purpose of the AI is a complex one. AI has the potential to drive productivity, innovations and competitiveness and to significantly improve human life. It is of great importance for the economy to boost the growth of GDP and productivity by enabling companies to automate complex tasks and improve efficiency, to develop new business models, new products and services. It is also a very useful tool for the education, healthcare, energy and security and other sectors, like prevention of natural disasters, modernizing education, improving health services. However, bearing in mind the large number of ways in which AI is used; it is obvious that key actors around the globe can pursue different objectives with the AI. There are governments and/or legal entities that use the AI not only for the above-mentioned purposes, but also to monitor the population. China is using 176 million high-performance intelligent surveillance cameras with embedded facial recognition. It is walking on the edge of democracy and human rights (National Artificial Intelligence Research and Development Strategic Plan, US Government, 2016).

Another example is the United States of America, which is one of the leaders in the AI research and development. The USA was the first country to implement a comprehensive AI research and to develop a strategic plan in the year 2016. It has six key priorities including: investment, ethical and legal issues, developing data sets and testing infrastructure, and skills. One year later, in July 2017, China published its Next Generation AI Development Plan, in order to become a world leader in AI by 2030. The document is based on the following targets: tackling the key problems in research and development; pursuing a range of AI products and applications and cultivating an AI industry. The leading duo is followed by Japan, the UK, Germany and the UAE. All these countries adopted AI strategies in the course of 2017 targeting the same or very similar goals.

Towards an EU strategic plan for AI

In the EU, policymakers want to use AI to improve the decision-making and foster unbiased decisions. The European Commission is engaged in a process to ensure that the EU has the necessary tools to build responsible AI which will be market-driven but most important – human-centric. The European Council and the European Parliament invited the Commission to propose a **European approach to AI** by 2018. In April 2018, the European Commission delivered a Communication from the Commission

to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions named "Artificial Intelligence for Europe" (European Commission, 2018). This strategic document offers a great introduction in the topic, highlighting the importance of a coordinated approach by all key stakeholders at the EU level in order to make the most of the AI. The strategic document is based on the following pillars: launch of European initiative on AI; the EU position in a competitive landscape; combined financial frame between the Member States and the private sector, with a main focus on research and development; vision beyond 2020; ethical and legal framework and monitoring of the development.

In December 2018, the European Commission delivered new Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions named "Coordinated Plan on Artificial Intelligence" (Coordinated Plan on Artificial Intelligence, European Commission, 2018). The document is produced in order to promote the development of Artificial Intelligence (AI) in Europe.

The **Coordinated EU Plan** was needed because:

- Only when all European countries work together we can make the most of the opportunities offered by AI and become a world leader in this crucial technology for the future of our societies.
- Europe wants to lead the way in AI based on ethics and shared European values so that citizens and businesses can fully trust the technologies they use.
- Cooperation between Member States and the Commission is essential to address new challenges brought by AI.

Needles to say that, the EU established a High-Level Expert Group (AI HLEG) which is responsible for drafting **AI ethics guidelines** as well as to propose policy recommendations on investment and regulatory framework. On 18 December 2018, AI HLEG delivered Draft Ethics Guidelines for Trustworthy AI as a working document (Draft Ethics Guidelines for Trustworthy AI, European Commission, 2018). The final version is expected at the end of the first quarter of 2019. The leading principle of this document is the maxim: "maxims the benefits of AI while minimizing its risks". The core orientation is the "human-centric approach in the development of the AI". The final aim of the document is to produce "trustworthy AI as a north star", since human beings will only be able to confidently and fully reap the

benefits of AI if they can trust the technology. The Trustworthy AI has two components: (1) it should respect the fundamental rights, applicable regulation and core principles and values, ensuring an "ethical purpose" and (2) it should be technically robust and reliable since, even with good intentions, a lack of technological mastery can cause unintentional harm.

The Framework for Trustworthy AI is composed of the following chapters:

- Chapter I deals with ensuring the AI ethical purpose, by setting out the fundamental rights, principles and values that it should comply with;
- Chapter II derives guidance on the realization of Trustworthy AI, tackling both ethical purpose and technical robustness. This is done by listing the requirements for Trustworthy AI and offering an overview of technical and non-technical methods that can be used for its implementation; and
- Chapter III subsequently operationalizes the requirements by providing a concrete but non-exhaustive assessment list for Trustworthy AI.

The aim of the Guidelines is to offer guidance and it is not intended as a substitute to any form of policymaking or regulation.

Humans have an ethical code, an understanding what is right and wrong, and a system of regulations (laws) to support it. A similar approach is needed in AI, a set of rules that will allow the AI to perform in a best way not harming the people. On the other side, the frame should not be narrow and limiting the development and the use of AI. The AI ethical code should be finalized in the following way: the AI should not be used to harm humans in any way; it should not restrict the human rights and freedom; it should be used fairly and not to discriminate or stigmatize; it should operate transparently and it should be developed for the well-being of the people and society as a whole.

Create a trusted regulatory framework

The impact of AI on society and human beings depends on the way we will use the AI. There is a possible future where artificial intelligence drives inequality, inadvertently divides communities and is even actively used to deny human rights.

Human rights are universal and binding. They are on the pedestal in global terms. All governments, human beings and legal entities are obliged to respect the law and human rights. On international level, the United Nations developed special set of rules and guidelines named UN Principles on

Business and Human Rights (Guiding Principles on Business and Human Rights, (2011).

One might ask how AI can impact human rights. One of the most common cases of violation of human rights by AI is discrimination (Andersen, 2018). But the list is long and not limited only on discrimination.

Articles 6, 9 and 14 of the International Covenant on Civil and Political Rights guarantees the *rights to life, liberty and security, equality before the courts and a fair trial*. (International Covenant on Civil and Political Rights, UNHRq 2019). The growing use of AI in the criminal justice system risks interfering with rights to be free from interferences with personal liberty. The list continues with the so called "criminal risk assessment software", used in some legal systems to assist judges in their sentencing decisions. For example, in the USA it led to more black defendants falsely labeled as high risk and given higher bail conditions, kept in pre-trial detention, and sentenced to longer prison terms (Andersen, 2018). Practically, this AI system may attribute a level of future guilt, which may interfere with the presumption of innocence required in a fait trail. Also, Facial Recognition Software within law enforcement raises the risk of unlawful arrest due to error and overreach. History is rife with examples of people wrongly arresting people who happen to look similar to wanted criminals. Given the error rates of current facial recognition technology, these inaccuracies could lead to increased wrongful arrests due to misidentification, exacerbated by the lower accuracy rates for non-white faces (Goode, 2018).

With the increased level of shared data and information about our lives, including the social media, the *rights to privacy and data protection* is at risk. Privacy is one of the fundamental rights as well as the personal data protection is. Guaranteed with the ICCPR (Article 17) and the EU Charter of Fundamental Rights (Article 8). The AI systems are often composed to access and collect data and information. Namely, the analysis of data using AI may reveal private information about individuals or information that qualifies as protected. A very serious concern is the use of government social media monitoring programs. There is justification on the use of the AI to detect alleged threats. However, the misuse of AI monitoring programs is very possible to happen by collecting massive unwarranted data or information.

The further use and development of AI for surveillance might interfere the *right to freedom of movement*. AI can provide detailed depiction of individuals' movement as well as to predict future location. The justification about public safety reasons is in place. But weak governments and lack of democracy can

contribute to misuse of the AI in order to restrict the freedom of movement by locating people in real time and predict the movement of people.

Next risk on human rights regarding the use of AI is the *rights to freedom of expression, thought, religion, assembly and association.* A law passed in 2018 in Germany requires social media sites to remove a wide range of content within 24 hours after the content is "flagged". Flagging means that the authorities have the right to mark content as a risk – terrorist content, hate speech and fake news. However, based on the practice and implementation as well the fact that AI is imperfect, it is reported that much of the content is removed in error (Nolasco and Micek, 2018). There is a risk of censorship, self-censorship, persecution based on different grounds, including political and religious.

Usually, AI is designed to sort and filter data and information. This service can cause discrimination which might collide with *human rights: equality and non-discrimination*. Article 21 of the EU Charter of Fundamental Rights defines that any discrimination based on any ground such as sex, race, color, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited. Within the scope of application of the Treaty establishing the European Community and of the Treaty on European Union, and without prejudice to the special provisions of those Treaties, any discrimination on grounds of nationality shall be prohibited (EU Charter of Fundamental Rights, article 21). However, AI can discriminate in case when it treats different groups of people differently.

Other rights such as *rights to political participation and self-determination, rights to work and adequate standard of living, right to health, right to education,* etc. For example, AI can be used in social media algorithms and bots which lead the campaign to increase the reach of false information and potentially influence the voters. Citizens are exposed to different disinformation which is directing their behavior (voting pro or con) or even losing the trust in the legitimacy of elections. AI is a very useful tool for improving the quality and the quantity of the productivity. AI and its role in the automation of jobs, creates a real threat to the right to work. Additionally, AI plays a very important role in the healthcare sector in assisting doctors more accurately diagnose diseases or to provide more individualized patient treatment, etc. But there are also ways in which AI can endanger the right to health by discriminating the patient based on cost reduction (constructed over the income of the patient). AI can recommend different treatment depending on the socio-economic status.

Last but not least, soon it is very possible to expect a debate in the EU about the question of legal personality of the AI. In 2017, the European Parliament urged the European Commission to propose what it called "an electronic personality" as a legal status. There are a few pre-questions which need to be tackled. First, whether the EU has the power to decide about this issue or it is a matter of each Member State to decide? Second, beside the legal aspect, we should also underline the ethical aspect.

Conclusion

AI is an ocean without an end. Its further development will improve the quality of life of the human beings. The use of AI will continuously increase in the upcoming period. The EU is not in a lead in AI compared to some other countries worldwide. We are absolutely aware that if there is no frame, ethical or legal, there is a great potential for violation of human rights. Swift actions are needed to deal with this risk. Also, it is very important for the EU to take careful steps avoiding over-regulation. Over-regulation might influence further development of the AI. Having in mind the fact that different stakeholders are involved in research and development, as well as in implementation of the AI, it is very unlikely that only ethical limits will safeguard and protect the human rights from violation. Therefore, a legal approach is needed. The EU has to develop a certain legal frame for the AI.

The Union is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights, including the rights of individuals belonging to minorities. These values are common to the Member States in a society in which pluralism, non-discrimination, tolerance, justice, solidarity and equality between women and men prevail (Lisbon Treaty, 2009, art.2). These values constitute the foundation of the rights enjoyed by those living in the EU. Additionally, the EU Charter of Fundamental Rights brings together all the personal, civil, political, economic and social rights enjoyed by people within the EU. Also, the General Data Protection Regulation ensures a high standard of professional data protection, including the principles of data protection by design and by default. It guarantees the free flow of data within the EU. It contains provisions on decision-making based solely on automated processing, including profiling. In such cases, data subjects have the right to be provided with meaningful information about the logic involved in the decision (GDPR – Regulation (EU) 2016/679 of the Eu-

ropean Parliament and of the Council). The European Commission closely follows the Regulation's application in the context of AI.

To gain trust, which is necessary for societies to accept and use AI, the technology should be predictable, responsible, and verifiable, respect fundamental rights and follow ethical rules. In that sense, urgent actions are needed in:

a) Creation of trusted regulatory framework;
b) Training for the EU workforce on AI;
c) Development of comprehensive data protection legislation;
d) Government use of AI should be governed by high professional standards;
e) Private sector should develop accountability, ethical policies and transparency; and
f) Human rights impact assessment of use of AI.

The list of recommendations should be continued and further developed in an inclusive process of consultations with all key stakeholders. The final results should not be over-regulation which will negatively impact the innovations. Instead, the governments should monitor the development of the AI and create a track record of possible and actual violations of human rights. By establishing a list of best practices, the EU should develop soft regulation which, on one hand, will guarantee the human rights, and, on the other hand, will support the development of the AI. Member States and governments have the natural role to protect, promote, respect and fulfill human rights under the national and the international law and should avoid situations of engagement in or supporting practices that violate rights.

The AI systems are changing the way things function in governments and in companies. It brings the potential for serious violation of human rights. Current legislation protects human rights in an "old-fashioned" way and it doesn't fully fit in the new reality. The existing human rights framework must be applied for all the concerns until a new one is developed.

Bibliography

Andersen, L. (2018). *Human Rights in the Age of Artificial Intelligence*. New York, Access Now,

Communication Artificial Intelligence for Europe, (2018) European Commission

Coordinated Plan on Artificial Intelligence, (2018) European Commission

Draft Ethics Guidelines for Trustworthy AI, (2018) European Commission

EU Charter of Fundamental Rights, (2012) European Commission
"Face Value" (2018) https://irlpodcast.org/season2/episode3/
GDPR – Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)
OJ L 119. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679
Goode, L. (2018) *"Facial recognition software is biased towards white men, researcher finds"* the Verge, https://www.theverge. com/2018/2/11/17001218/facial-recognition-software-accuracy-technology-mit-white-men-black-women-error
Guiding Principles on Business and Human Rights, (2011) UN Human Rights
International Covenant on Civil and Political Rights, (1966) UNHR
Lisbon Treaty, Article 2, OJ C326/1 https://eur-lex.europa.eu/eli/treaty/teu_2012/art_2/oj
National Artificial Intelligence Research and Development Strategic Plan, (2016) US Government
Nolasco, D. and Micek, P. (2018) Access Now responds to Special Rappoteur Kaye on *"Content Regulation in the Digital Age"* Access Now.
PWC, (2017), Artificial Intelligence and Robotics – 2017 Leveraging artificial intelligence and robotics for sustainable growth