ARTIFICIAL INTELLIGENCE FROM THE PERSPECTIVE OF LAW AND ETHICS: CONTEMPORARY ISSUES, PERSPECTIVES AND CHALLENGES

Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of International Law and European Law

ARTIFICIAL INTELLIGENCE FROM THE PERSPECTIVE OF LAW AND ETHICS: CONTEMPORARY ISSUES, PERSPECTIVES AND CHALLENGES

Collection of proceedings of the international scientific conference
"Artificial Intelligence from the Perspective of Law and Ethics:
Contemporary Issues, Perspectives and Challenges"

November 11, 2021
Košice

Ján Klučka, Lucia Bakošová, Ľuboslav Sisák (eds.)

Leges

Citation:
KLUČKA, J., BAKOŠOVÁ, L., SISÁK, Ľ. (eds.): Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges. Praha: Leges, 2021, 257 p., ISBN 978-80-7502-590-6

Reviewers:
Doc. JUDr. Kristián Csach, LL.M., PhD.
Department of Civil and Commercial Law, Faculty of Law, Trnava University

Doc. JUDr. Marcel Dolobáč, PhD.
Department of Labor Law, Faculty of Law, Pavol Jozef Šafárik University in Košice

Editors:
prof. JUDr. Ján Klučka, CSc.
Mgr. Lucia Bakošová, PhD.
JUDr. Ľuboslav Sisák
Institute of International Law and European Law, Faculty of Law, Pavol Jozef Šafárik University in Košice

The publication was created within the project supported by the Commission for Economic and Legal Sciences VEGA of the Ministry of Education of the Slovak Republic under no. 1/0643/20 under the name "Artificial Intelligence from the Perspective of Law and Ethics."

Published by Nakladatelství Leges, s. r. o., Lublaňská 4/61, Prague, Czech Republic, in 2021.
First edition
Print run: 50 copies
Cover design Michaela Vydrová
Layout Gradis
Print Tiskárna Protisk, s. r. o., České Budějovice

ISBN (print) 978-80-7502-590-6
ISBN (e-book) 978-80-7502-591-3

KATALOGIZACE V KNIZE – NÁRODNÍ KNIHOVNA ČR

Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges (konference) (2021 : Košice, Slovensko)
Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenges : collection of proceedings of the international scientific conference "Artificial Intelligence from the Perspective of Law and Ethics: Contemporary Issues, Perspectives and Challenge" : November 11, 2021, Košice / Ján Klučka, Lucia Bakošová, Ľuboslav Sisák (eds.). – First edition. – Prague : Leges, 2021. – 1 online zdroj
Nad názvem: Pavol Jozef Šafárik University in Košice, Faculty of Law, Institute of International Law and European Law. – Obsahuje bibliografie a bibliografické odkazy

ISBN 978-80-7502-591-3 (online ; pdf)

* 004.8 * 34 * 17 * 341 * 342.7 * 341.231.14 * 343 * 347.951 * 347.77/.78 * 347.447 * 327.7 * 341.3 * (4) * (062.534)
– umělá inteligence – právní aspekty
– umělá inteligence – etické a morální aspekty
– mezinárodní právo
– lidská práva
– trestní právo – země Evropské unie
– soudní rozhodnutí a stanoviska
– ochrana duševního vlastnictví
– odpovědnost za škodu
– mezinárodní instituce a organizace
– mezinárodní právo humanitární
– sborníky konferencí

34 – Právo [16]

CONTENT

1 GENERAL OVERVIEW OF THE ARTIFICIAL INTELLIGENCE AND INTERNATIONAL LAW

Ján Klučka

My contribution provides a brief overview of the contemporary possibilities and problems of international law relating to its potential regulation of AI based Technologies (first part) and identification of some areas where artificial intelligence is more or less applicable in the international law today (second part).[1]

1.1 Problems with the new rules of international law relating to artificial intelligence: Lack of the definition of artificial intelligence in current international law

To date there is no internationally accepted definition of the basic concept of AI. The lack of a definition hampers further discussions on possible international cooperation in the AI field, and in practice it is difficult to adopt international legislation, the concept and subject matter of which is not clearly definable. Another problem concerns the specificities of the rule-making process of the new rules of international law. Besides content uncertainties of possible legal regulation of the artificial intelligence, the rule-making process of the new rules of international law is time-consuming, and therefore it cannot respond adequately to the fast development of the new artificial intelligence technologies. As a result, the states are rightfully reluctant to be bound by the regulation of this new technology as long as its definition is not clear. In addition, this notion should not be used in legislation of a regulatory nature, as the absence of its legal definition obscures the very subject matter and purpose of the legislation. The definition of AI should not be too "narrow", as it will not be able to absorb the development of new technologies and legislation will get "lagging behind" over time and may become obsolete. On the other hand, the definition should not be too broad either, as this could lead to several interpretation problems in the process of its national implementation. As regards the requirements that such a definition should meet, it is inclusiveness, accuracy, complexity, feasibility and consistency which are

[1] This paper was prepared within the framework of the research project VEGA no. 1/0643/20 "Legal and Ethical Perspective of Artificial Intelligence".

regularly emphasized. While the definition of AI is still lacking at the global level, such a definition can already be registered at the regional level. This applies specifically to the European Union, which points to the need to adopt a common European framework with harmonized definitions and common ethical principles of AI. The accepted AI definition contains appeal to intelligence, autonomy as well as to the achievement of specific objectives. The basic problem of artificial intelligence is therefore the creation of a machine that behaves in such a way that we would call it intelligent as we do in the case of human beings. According to the EU definitions of AI it shall be based either on a software base or embedded in a hardware device, which copies the intelligence of human through collecting and processing data, analysing and interpreting the environment and acting with a degree of autonomy in order to achieve specific objectives.

Given the scope and complexity of the tasks which the mechanisms equipped with AI technologies are capable of solving, legal writings speak of the narrow (weak AI) and general (strong AI), which differ in the degree and extent of autonomy in the performance of their functions. The first of them has intelligence limited to a specific area, within which the system is capable of repeated solution of programmed tasks, or only one task under constant human control. In contrast to narrow form strong AI intelligence will be able to replicate human levels of intelligence similar to natural persons.

Internal or international regulation of artificial intelligence?

First look confirms that artificial intelligence is currently subject mainly to national regulations, while international regulations are becoming more and more delayed. Priority of national legislations confirms that they are applied in almost 50 countries around the world. Its content and number depend on the importance of development of AI technologies in a particular country and its economy. In general, such national legal and other acts regulate mainly partial methods and conditions of commercial use of AI technologies of a civil nature. This situation is not satisfactory, because if such technologies cross national borders over time, they may be confronted with different national rules in the territory of another state. This fact can lead to conflicts as well as growing problems, the result of which can be a slowdown in the development of international transport, trade, but also increasing tensions in the political field, etc. The solution to this situation could therefore be an international treaty on uniform legislation, which would oblige the contracting

parties to uniformly regulate the criminal, civil and administrative aspects of AI technologies. For the sake of completeness, it should be noted that more than 160 organizational, national and international sets of rules concerning the ethical and guiding principles of AI have been developed to date, but unfortunately there is no common platform for converging or harmonizing them because there is no body of their international coordination and cooperation of states.

With regards to the appropriate "timing" of future AI international legislation, there is a concern that early regulation may be inaccurate or incomplete, and may be an obstacle to further development and potential benefits of incoming AI technologies. On the other hand (and paradoxically) the absence of reliable legal environment may also be a problem for the future AI technological innovations because it can open the door to arbitrary action by states.

The solution could be flexible international law capable of responding more quickly to future technological changes. Finally, several developed countries are currently not interested in the international regulation of AI, or its oversight by an international body and/or organization, and their relations regarding the development of new technologies are marked by competition (USA v. China) order to achieve world leadership in strategic, military, commercial or other benefits that would result from such leadership. The contemporary reality is that, at best, it will only be possible to register the sporadic and gradually emerging traditional arrangement of AI through international law over the next few years, despite its growing proliferation and application. At least temporarily, such a regulatory gap will be filled by soft law rules. However, the existence of international legal regulation cannot be ruled out for the regulation of partial questions, and/or to modify AI-related procedural rules. One of the specific issues that could be the subject of international regulation may be AI technology used in war conflicts (so-called killer robots. LAWS), which is generally believed to be prohibited by international law. Nevertheless and beyond this topic there are general values and principles that should guide AI and the adoption of which is of interest to the international community, regardless of the speed and development of its AI technologies. Due to the fact that AI technologies are getting closer and closer to people, there is a need to adjust their interrelationships so that AI technologies respect people's lives and health, as well as their fundamental rights and freedoms. This is a general need of the international community to guarantee the safe development of AI and its harmonious "cohabitation"

with the human community. With respect to the future regulation of artificial intelligence it is reasonable to assume that this will be of a multilevel nature consisting of the rules of the soft law, international standards and international law whereas their mutual relations and their importance may vary in the future.

Non-binding rules relating to the artificial intelligence

The above-mentioned international law gap relating to AI is currently filled by the non-binding rules of soft law and international standards, which are developed much faster and able to react more appropriately to the new incoming challenges of the artificial intelligence. These rules can be also inspirational for the rule-making process of the new legal rules. However, these rules also have their advantages and disadvantages. Their advantages consist of the fact that soft law rules are of the "non-state" nature, and accepted by various international NGOs, professional and scientific communities, public-private partnerships and even major business entities (Microsoft, Google, IBM, Facebook), which undertake to keep their products compliable with such rules. Experience confirms that these rules can take various forms, starting with voluntary programs, standards, codes of conduct, certification files, model laws, guidelines, declarations of principles and guidelines, and do not require any form of formalized consent. As they are not territorially limited, they can also be used on an international scale by states, non-governmental organizations, or private companies. It is possible to mention their freedom of content, as they are able to address a wide range of issues, starting with the health risks of AI and ending with ethical problems of AI. They can be used simultaneously by different entities and, if they prove successful in the practice, they can be gradually "transformed" into legally binding rules. In general soft law rules can therefore initiate and/or facilitate the development of new rules of international law relating to AI.

On the other hand soft law rules have also their disadvantages. As the acceptance of these rules is based on a voluntary approach, there is a risk of their selective use in the practice based on the individual interest of a particular user instead of their general acceptance. Another problem is the possibility to choose between existing soft law sets that have been processed by various non-governmental entities. The legal doctrine therefore rightly points to the need for a coordinating body which should try to address these shortcomings. These rules are usually proposed by small groups of experts

on the basis of their own group selection. Their next disadvantage is also considered to be the generality and vagueness of their formulations which allow subjective interpretation, and reduces the stability of their application practice. And finally their compliance is not enforceable, which is generally considered to be the Achilles' heel of soft law rules.

General principles of AI

Despite the persistent lack of binding international law regulation, the international community has been able to adopt the general principles of artificial intelligence. They were however adopted by the "non-specialized" OECD in 2019, when 36 of its member states as well as several non-member (partner) states signed a document containing the OECD Principles of Artificial Intelligence. Although these principles are not legally binding, their uniqueness lies in the fact that, for the first time, they have been agreed by the states that intend to apply them in their national AI policies and respect them in future AI research and development processes. The document embedded five basic value principles related to the responsible deployment of trusted AI technologies, as well as five recommendations for international cooperation in this field. These general principles should serve as a global reference point for ensuring the credibility of AI. The principles emphasize *inter alia* that AI should serve the benefit of humanity and the planet, as well as the sustainable development and well-being (Principle 1), AI systems should be designed to respect the rule of law, democratic values and include adequate safeguards in where human intervention is necessary to ensure social justice (Principle 2). In relation to AI, transparency and disclosure of information should be ensured to guarantee that people are aware that they are in contact with AI (Principle 3). During their "life cycle", AI systems must operate in a robust and safe manner, and potential risks should be continuously evaluated and managed (Principle 4). Organizations and individuals developing and operating AI systems should be responsible for their proper functioning and respect (Principle 5). These principles prefer human-centered approaches in relation to the future development of AI, to reduce disparities between countries and to ensure a minimum level of guarantees for all people. These principles became the inspiration for the G20 group which adopted a similar document in June 2019 fully based on OECD principles.

Interpretation and application of international law in relation to AI

It would be an incorrect statement to say that the future regulation of AI technologies should be limited only to new rules of international law and/or soft law and international standards. However, in relation to artificial intelligence, legal rules may also be provided by current international law, if the interpretation of the text of valid treaty allow their application. If this is not the case, the creation of new international treaty is possible.

Lethal autonomous weapons systems

The growing technological capacity in the development of autonomous systems in the military field has caused a number of problems of a legal, moral, technological and security nature. The special attention of the international community was attracted by the so-called autonomous killing systems known as killer robots – LAWS. The most important aspect of their autonomy (from a humanitarian, legal and ethical perspective) is that they are able to activate and launch an attack without the influence and control by the human operator. In this respect, autonomous lethal systems clearly differ from other existing weapon systems in which the determination of time, place of attack as well as target identification is determined by the weapon operator on the basis of his own analysis.

As this is a new type of conventional weapon (albeit equipped with AI technology), the application of the Convention to the Prohibition or Restrictions on the Use of Certain Conventional Weapons, which may be considered excessively injurious or indiscriminate from 1980 the application of just this Convention came into consideration first. It should be however noted that current international law (including humanitarian law) is insufficient to apply to the autonomous lethal weapons, as its basic rules are to be applied only by human operators and not by machines without significant human control. The failure to comply with this requirement is contrary to international humanitarian law and the Martens clause and customary rules. According to most opinions. the international regulation of the prohibition of killer robots should have the form of a *new international treaty* and/or *additional protocol* to the Convention from 1980. The special group of experts started the preparatory discussion on the various aspects of such treaty since 2016. In short, such regulation should provide significant human control over the use of force and ban weapons that should operate without such

control. Such a ban would not allow any exceptions and should apply in all circumstances.

Interpretation of the term "driver" in Conventions on the road traffic relating to driverless cars

The second aspect of the relationship of AI technology to the current international law represents the problem of interpretation of the term "driver" in Article 8 of the 1949 Geneva Convention on Road Traffic and the Vienna Convention on Road Transport of 1968. The relevant provisions of these conventions [Article 4 (1) of the Geneva Convention and Article 1 (v) of the Vienna Convention] require the driver to be a natural person who fulfils the required physical and professional conditions. This problem arose in connection with the gradual introduction of driverless cars and the question of whether an autonomous car driving control system using AI could be considered as a "driver" within the meaning of the said conventions.

A closer analysis of the individual paragraphs of Article 8 confirms that such a system does not meet the conditions imposed on the driver, because it is a natural person who must have the necessary physical and mental fitness and must be in good physical and mental condition (paragraph 3), and must have also sufficient knowledge and skills to drive a vehicle (paragraph 4).

The Road Convention however offers another possibility of interpretation of the term driver. It´s Article 8 (5) that states that every driver must be able to control his vehicle at all times. The driverless cars therefore raised the question whether only a natural person complied with the required conditions but also another entity that is also able to exercise control over the vehicle and different from a natural person (the so-called functional principle) can be considered a driver. To address this issue, an approach has been taken whereby road transport conventions can be considered as "living instruments", which allows the interpretation of the term "driver" to be extended to include not only the natural person but also another other "non-human" entity capable of constantly performs control the vehicle. By amending the Vienna Convention of 1968, a paragraph 5bis was inserted in Article 8 of in March 2014, stating that a system affecting the vehicle's driving style will be considered to comply with para. 5 (This system is capable of performing control over vehicle).

With this amendment, the presence of the driver in the vehicle is still necessary, but the vehicle can be steered by an automatic system, which can be

taken over by the driver at any time. In this connection, reference should be made to the common opinion adopted by the Governments of Austria, Belgium, France, Germany and Italy, which points out that the vehicle system need not "allow" the driver to switch off completely in the event of an emergency braking when the maximum braking mechanism is initiated. The reason is that in such a case the driver may not be able to react adequately in a potentially extremely dangerous situation, which could be exacerbated if his vehicle's autonomous system is deactivated or switched off. As such a system helps the driver to maintain control of the vehicle in the event of dangerous traffic situations, it can be considered to comply with Article 8 § 5 of the Convention. The legal writing characterizes such a situation as shared driving and adopted change create a legal framework for semi-autonomous vehicles with the presence of the driver. In such a sense Vienna Convention has been amended to allow the use of autonomous control systems, provided that they comply with international regulations and can be switched off by the driver with exception of emergency and dangerous situation on the road.

1.2 Artificial Intelligence in international law

Following areas in international law can be identified as conceivable for practical application of AI technology mainly learning machines: the stage of preparation of international treaty, its negotiation, translation of adopted treaty, monitoring of compliance of treaty by their contracting parties, computational analysis of treaty and automation of international law.

The basic consideration underlying the possibility of *automating international law* is that the activities of lawyers in national law are increasingly automated by various AI technologies, especially in the field of legal research and information services. These can also assist in routine legal acts, such as the preparation of draft contracts, routine appeals against court decisions, legal classification of the facts of the case, etc. Due to its development and expansion, the role of the human factor in the field of domestic legal services and expertise will probably decrease in the near future, which will give lawyers more time to solve more complex legal problems. Against this background the question arises as to whether international law is also appropriate for such automation and, if so, to what extent. It can be said that international law is different in that it lacks the necessary quantitative and qualitative characteristics because it is relatively small and diverse. It contains only a limited number of international treaties of various contents, which often contain

broad and vague wording that can be interpreted differently under the political and cultural influence of concrete contracting parties. The data is too limited for learning international law related to case law of international judicial bodies providing relatively small amount of data sufficient for example for prediction of trial in international law. Such a characteristic also fully applies to the rules of customary international law. The only exceptions are international trade law under the auspices of the World Trade Organization (WTO) and international criminal law, which are more homogeneous and "denser". Based on these facts it is reasonable to assume that at least for the near future neither international law nor the agenda of international lawyers can be automated.

AI technologies can be used to identify priorities in those areas which the future party (parties) focused on in preparing the same or similar treaty in the past, identification of frequently used legal terms and their common interpretation, identification of foreign legislation, the use of which is possible, the process of preparing the negotiators of the future treaty. Next step represents the *negotiation process of the draft treaty when* AI is most often used to improve and speed up the quality of translation of oral discussion, working materials including international precedent materials, international conference materials and other *travaux preparatoires* documents. The correct translation is fully required with respect of norms contained in the original text of the adopted treaty because its mistranslation has the potential to cause a great deal of damage, trigger the procedure of responsibility and sanction regime of contracting parties.

Another area for the use of AI technologies is *in the monitoring of the fulfilment* of international obligations of states issuing from international treaties due to their ability to collect and analyse large amounts of data by using machine learning. Using this capability increases the effectiveness of monitoring facilities, which makes it possible to determine whether states are properly fulfilling their international obligations, for example in the field of disarmament, environmental protection, etc. For the sake of completeness, it can be mentioned that the internet also contributes to the areas related to the application of international law, *opinio iuris* in the process of creating new customary rules, etc.

As far as *customary international law* is concerned, the ability of AI allows to gather information on the previously unknown practices of states and *opinio iuris* in large number of states and in UN archives and thus speed up the process of creating new customary rules. The main technical problem

is that texts of the documents regulating the practice of the States and their *opinio iuris*, is not "legible" because it does not have a form in which relevant AI technologies could be used. The goal of so-called digitization is therefore the transformation of written text into the form of data that AI is able to read and analyse.

Finally, to examine the text of existing international treaties, the so-called *computational analysis,* could be used with respect of the texts which are human-readable but not readable by artificial intelligence technologies, although the number of treaties made in a format readable by AI (pdf, xml, html) has recently increased. In the case of computational analysis a legal text is transformed into digital format in combination with natural language processing. The use of such tool helps to categorize, process, analyse and extract the required information faster, and also from a large amount of data, and to make it accessible and accelerate the process of their analysis by humans.

Suggested list of literature and resources

1. ILSCHNER, W.: The Computational Analysis of International Law. In: *Ottawa Faculty of Law Working Paper,* 2019, No. 33. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3428762 (quoted 3 November 2021).
2. BOUTIN, B.: Technologies for International Law & International Law for Technologies. In: *GroJIL-blog,* 2018. Online: https://grojil.org/2018/10/22/technologies-for-international-law-international-law-for-technologies/ (quoted 3 November 2021).
3. BURRI, T.: International law and Artificial Intelligence. In: *German Yearbook of International Law*. 2017, vol. 60, p. 95. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3060191 (quoted 3 November 2021).
4. CIHON, P., MAAS, M. M., KEMP, L.: Should Artificial Intelligence Governance be Centralised? Design Lessons from History. In: *Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)*. Online: https://arxiv.org/pdf/2001.03573.pdf (quoted 3 November 2021).
5. DEEKS, A.: Introduction to the Symposium: How Will Artificial Intelligence Affect International Law? In: AJIL Unbound, 2020, vol. 114, p. 138. Online: https://www.cambridge.org/core/journals/american-journal-of-international-law/article/introduction-to-the-symposium-how-will-artificial-intelligence-affect-international-law/CD26AD55818677B9B28FB59EAD96D4BB (quoted 3 November 2021).
6. DEEKS, A.: High-Tech International Law. In: *Virginia Public Law and Legal Theory Research Paper, 2020*, No.10, p. 11. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3531976 (quoted 3 November 2021).

7. DIALLO, B. S.: The Binding Force of International Legal Standards in the Face of the Recurrent Practice of Soft Law. In: *Adam Mickiewicz University Law Review*, 2017, vol. 7, p. 88. Online: https://pressto.amu.edu.pl/index.php/ppuam/article/view/12190/12031 (quoted 3 November 2021).
8. DOCHERTY, B.: The Need for and Elements of a New Treaty on Fully Autonomous Weapons. *Human Rights Watch,* June 1, 2020. Online: https://www.hrw.org/news/2020/06/01/need-and-elements-new-treaty-fully-autonomous-weapons (quoted 3 November 2021).
9. European Parliament Resolution of 20 January 2001 on artificial intelligence: questions of interpretation and application of international law in so far as the EU is affected in the area of civil and military uses and of state authority outside the scope of criminal justice (2020/2013(INI)). Online: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0009_EN.html (quoted 3 November 2021).
10. ERDÉLYI, O. J., GOLDSMITH, J.: Regulating Artificial Intelligence: Proposal for a Global Solution. In: *Association for the Advancement of Artificial Intelligence,* 2018. Online: https://par.nsf.gov/servlets/purl/10066933 (quoted 3 November 2021).
11. Framework for Artificial Intelligence Systems Using Machine Learning (SC 42 WD 23053). Online: https://www.iso.org/standard/74438.html (quoted 3 November 2021).
12. G 20 Principles on Artificial Intelligence. Online: https://www.meti.go.jp/press/2019/06/20190610010/20190610010-1.pdf (quoted 3 November 2021).
13. HILL, S., MARSAN, N.: Artificial Intelligence and Accountability: A multinational Legal Perspective. Online: https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-160/MP-IST-160-PP-4.pdf (quoted 3 November 2021).
14. MAAS, M. M.: International Law Does Not Compute: Artificial Intelligence and the Development, Displacement or Destruction of the Global Legal Order. In: *Melbourne Journal of International Law*. 2019, vol. 20. Online: http://classic.austlii.edu.au/au/journals/MelbJIL/2019/3.html (quoted 3 November 2021).
15. MARCHANT, G.: "Soft Law" Governance of Artificial Intelligence, January 25, 2019. Online: https://aipulse.org/soft-law-governance-of-artificial-intelligence/ (quoted 3 November 2021).
16. MARCHANT, G. E.: The Growing Gap Between Emerging Technologies and the Law. In: MARCHANT, G. E., ALLENBY, B. R., HERKERT, J. R. (Eds.): *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem*. Springer: Dordrecht, 2011, p. 19–33. ISBN 978-94-007-1355-0.
17. MIAILHE, N.: AI&Global Governance: Why We Need an Intergovernmental Panel for Artificial Intelligence. *United Nations, Centre for Policy Research,* December 12, 2018. Online: https://cpr.unu.edu/publications/articles/ai-global-governance-why-we-need-an-intergovernmental-panel-for-artificial-intelligence.html (quoted 3 November 2021).

18. NASH, L.: Advancing Intelligence and Global Society: International Law´s Role in Governing the Advance of Artificial Intelligence. In: *Kentucky Law Journal Online*, 2019-2020, vol. 108, p. 107. Online: https://www.kentuckylawjournal.org/online-originals/advancing-intelligence-and-global-society-international-laws-role-in-governing-thenbsp-advance-of-artificial-intelligence (quoted 3 November 2021).
19. New Weapons, Proven Precedent: Elements of and Models for a Treaty on Killer Robots. October 20, 2020. Online: https://www.hrw.org/report/2020/10/20/new-weapons-proven-precedent/elements-and-models-treaty-killer-robots (quoted 3 November 2021).
20. OECD Principles on Artificial Intelligence on May 22, 2019. Online: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (quoted 3 November 2021).
21. PERRITT, H. H., Jr.: The Internet is Changing the Public International Legal System. In: *Kentucky Law Journal 1999 – 2000,* vol. 88. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=446102 (quoted 3 November 2021).
22. SCHUETT, J. et al.: *A legal Definition of AI*. August 26, 2019. Online: https://deepai.org/publication/a-legal-definition-of-ai (quoted 3 November 2021).
23. SCHERER, M. U.: Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. In: *Harward Journal of Law & Technology*, 2016, vol. 29, no. 2, p. 359. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2609777 (quoted 3 November 2021).
24. VIHUL, L.: International Legal Regulation of Autonomous Technologies. November 16, 2020. Online: https://www.cigionline.org/articles/international-legal-regulation-autonomous-technologies (quoted 3 November 2021).
25. RHIM, Y.-Y., Park, K.-B.: The applicability of Artificial Intelligence in International Law In: *Journal of East Asia & International Law*, 2019, Vol.12, No. 1. Online: http://journal.yiil.org/home/pdf/publications/2019_12_1_pdf/jeail_v12n1_01.pdf (quoted 3 November 2021).

Author details

prof. JUDr. Ján KLUČKA, CSc.
Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of the International Law and European Law
j.klucka@klucka.eu

2 TRUSTING THE USE OF AI IN THE LAW FIELD: FOREMOST, AN ETHICAL CHALLENGE

Ruxandra Andreea Lăpădat

Abstract

Acknowledging the deep change that humanity faces today is the premise of any discussion about the artificial intelligence role in any field. There is growing evidence that an irreversible transformation is already on the way. However, the law field represents a controversial one to integrate AI, given its ancient philosophy as a milestone and its delicate place in the engine of the society. The ethics in this matter is the keyword that cleaves into engaging issues that thrive to be discussed. Going deeper into the matter, the immediate challenge translates into the impact of this issue over one of the core human rights affected: the right to a fair trial. The analysis about the huge impact of the digital era over the act of justice is pulsating over an ethical denouement: is it normal for robots to decide over people? Besides, in order to offer a practical perspective to this paper, we also encourage the reader to accompany us on a short journey to the future of professions in the law field, as a correlative matter to the act of justice in the digital era. Bearing in mind that digitization is only the first step of the digital transformation we're speeding to, the tendency is to believe that the law experts will be replaced by advanced systems, or by less costly workers supported by technology or standard processes, or by lay people armed with online self-help tools. Can this be totally true? We'll try offering an argumentative point of view about how the future of law professions might look.

Introduction

The use of AI in the law field started *in bona fide,* aiming to improve the access to justice, facilitating the right to a fair trial and minimizing the time spent in order to receive justice. The process was slow, hard to implement at first and to accept in the beginning, but today, AI consolidated its benefits and role as a game-changer in any industry that it is used in, leading to a digitalization of justice. What do we understand through the concept of digitalization in the law field? Is the use of AI intelligence reduced to programming and automatization? Is it keeping in line the healthy reasons we started using it for? We would argue not. In this regard, it is necessary to provide a perspective over both definitions of the process and the AI itself.

Firstly, AI is a concept hard to exhaustively define, but in our opinion, the most relevant definition refers to AI as the digital intelligence of a machine that could successfully perform any intellectual task that a human being can.[1] Being characterized by the digital environment, it gives the law field a new fictional side that appears as unknown and expanding. The processes implying AI are divergent today, even if they started from the same premises. Understanding these processes that take place today means understanding more clearly the steps to the deep transformation that law principles face today.

In this regard, *specialists draw attention to the use of interchangeable terms of digitalisation, digitisation and digital transformation, albeit describe distinct processes. Thus, digitization is a simpler operation, namely the transition of an element from analogous form to digital form. In contrast, the term digitialisation is wrongly equated in common knowledge to that of digital transformation, although it is more limiting than the latter. Specifically, digitalisation involves the use of digital technologies and data to change the way we work and the interaction between stakeholders, to attract new revenue, to replace or transform certain processes and to create an environment centered on the use of digital information. Digital transformation involves the integration of digital technology in all aspects of an area of activity, which fundamentally changes the way it is operating, and it provides distinct added value for stakeholders. Unlike digitalisation, which involves adapting to new technologies, digital transformation involves profound changes. It requires a new cultural orientation, with an emphasis on challenging the status quo and on continually testing the best solutions.*[2]

Thus, being presented, we invite you to discover three current problems being discussed today on an international level, emerging efforts in the same time in order to find sustainable solutions.

[1] STĂNILĂ, L.: Artificial intelligence. Criminal law and the criminal justice system. Memories of the future, 2020, p. 35.

[2] DURACH, F., MĂRCUȚ, M., PUCHIU, R., ȘTEFAN, V.: *De la digitalizare la transformare digitală în România*, Institutul European România, 2021, p. 6. Online: http://ier.gov.ro/wp-content/uploads/2021/07/Policy-brief-nr.-9_iulie-2021_Final.pdf (quoted 1.11.2021).

2.1 The issue of the legal existence and intention of artificial intelligence – who takes responsibility?

Using any dictionary of synonyms, we can observe that the meaning for the word *intelligence* goes slightly from brain, reason, knowledge to statistics, judgement, surveillance, inclining to emphasize an increasing subjectivism when we analyze deeper. Why? A simple guess would get us quickly to the motive behind. The root of the intelligence is undoubtedly reflected by intent.

Consequently, the most stringent problem of using AI in the law field arises in regard to the core of any mental aspect behind a human action, replicated now by a machine: its intent. Even if the AI is created through an exclusively objective and mathematical process, the paradox is that even a sum of tech processes results in a cognitive-like outcome. This being said, even if using a strong control over the AI involved in the law field, deviant behavior may take place. The essential question remains: who takes responsibility for these risks?

The quick answer may be intuitively directing us to the human person that created the robot or the AI it is using. This may apply only in a world where people can totally control the AI and use it only for automatic processes, where the AI was not created with learning skills. But today, the automatic robots with no potential to grow represent no longer the vast reality we're living in, but only a minority of it, because they are simply not enough to meet the new needs existing in every field today. Hence, the answer may be more nuanced than we would expect.

In this regard, we find two relevant practical examples that show the scary stage where we might find ourselves in all this AI development speeding around.

The first example refers to TESLA, one of the biggest electric vehicle creating company worldwide, which put on the market fully autonomous cars, which use AI intelligence for self-driving. One of these cars killed two men in Spring 2021, in Texas, after a Tesla car which was driving itself when it ran off road, hit a tree and burst into flames,[3] the company is being sued for the accident.

[3] The New York Times: 2 Killed in Driverless Tesla Car Crash. Online: https://www.nytimes.com/2021/04/18/business/tesla-fatal-crash-texas.html (quoted 1.11.2021).

Secondly, we would like to draw the attention to the prejudices that can emanate form the data biased system, resulting in big breaches of the ethical principles, such as inequity and discrimination created by the use of AI which is not strictly controlled. In this regard, we cite a research study from Harvard and a study case in California Law which are emblematic, as follows:

In a paper published earlier this year, she found that Internet searches for "black-sounding" names — such as Darnell or Ebony — were 25 percent more likely to result in the delivery of an ad suggesting that the person had an arrest record, even when no one with the name had an arrest record in the company's database.[4]

A real-world example from a different context as to how biased data labeling can skew results. St. George's Hospital, in the United Kingdom, developed a computer program to help sort medical school applicants based on its previous admissions decisions. Those admissions decisions, it turns out, had systematically disfavored racial minorities and women with credentials otherwise equal to other applicants. In drawing rules from biased prior decisions, St. George's Hospital unknowingly devised an automated process that possessed these very same prejudices.[5]

Understanding the spread of the phenomenon, international organizations started as a first step to consider offering legal personality to the AI robots. In this regard, we mention as a first the European Parliament's report with recommendations to the Commission on Civil Law Rules on Robotics, from 27.01.2017, where it is mentioned that *thanks to the impressive technological advances of the last decade, not only are today's robots able to perform activities which used to be typically and exclusively human, but the development of certain autonomous and cognitive features – e.g., the ability to learn from experience and take quasi-independent decisions – has made them more and more similar to agents that interact with their environment and are able to alter it significantly; whereas, in such a context, the legal responsibility arising through a robot's harmful action becomes a crucial issue.*[6]

[4] REUELL, P.: Seeking fairness in ads. The Harvard Gazette, 30.04.2013. Online: https://news.harvard.edu/gazette/story/2013/04/seeking-fairness-in-ads/ (quoted 1.11.2021).

[5] BAROCAS, S., SELBST, A. D.: Big Data's Disparate Impact. In: *California Law Review*, vol. 104, 2016, p. 682, ISSN 0008-1221. Online: https://ssrn.com/abstract=2477899 (quoted 1.12.2021).

[6] European Parliament: Committee on Legal Affairs. Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), 27.01.2017. Online:

On a country level, we mention the Federal Court of Australia as a pioneer, that ruled this year that an "inventor can be non-human", meaning that an AI system can own a patent. This landmark decision was made shortly after South Africa became the first country to approve AI as an inventor in July 2021. Unlike the UK and the US, which refused to recognise AI systems as patent owners, Australia seems to open the door for offering a full legal personality to the AI.[7]

The problem of the intent and its responsibility is far away from being solved by offering legal personality to the AI, tipping only the ethical iceberg underneath, but the solutions taken into consideration by the European Parliament are interesting to debate. There are two options proposed in the abovementioned report. First, according to the principle of strict liability, there should be a manufacturer who is liable, because he is best placed to limit the damage. Then he can turn against his suppliers. The other option is a risk assessment approach whereby tests must be performed beforehand and compensation must be shared between all parties involved.

Moreover, a compulsory insurance can be taken into consideration, at least for large robots. In terms of liability, customers need to be sure that they will be insured in the event of damage. Another big issue that rises is security and data protection. Robots cannot operate without a data exchange, so there is also the question "who will have access to this data?".

We can observe that the subject of the responsibility of the AI is no longer theoretical, but in an imminent need to be rethought and regulated from the legislative perspective.

Coming back to the ethical perspective, the solutions proposed above seem practical, but odd. Is it ethical for a manufacturer to be held accountable for something he cannot control, as well as for deeds he has not committed? If we put the responsibility on the shoulders of the manufacturer and every other party involved, without certain conditions being settled, do we not inhibit the whole technological development?

We think that more nuances need to be taken into consideration over this problem. In our opinion, it will only be fair for a manufacturer to obey the legal rules existing today for any offense, because they need no update: he

https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html#title2 (quoted 1.11.2021).

[7] LO, D.: Can AI replace a judge in the courtroom? UNSW Newsroom, 01.10.2021. Online: https://newsroom.unsw.edu.au/news/business-law/can-ai-replace-judge-courtroom (quoted 27.10.2021).

can be held accountable if there is a proof for any kind of intent when producing the robot. If there is an error or an intent that cannot be proven in the producing process and the learning skills of the robot acted on deviant behavior or deviant development, then we argue that the manufacturer is not trustworthy and the situation presented may remain uncovered by the legal framework.

What does need an update is how we legally translate the robot' s behavior. We cannot put an equal between the intent of the manufacturer and a possible intent of a machine. Even if the machine cannot have emotions, therefore it may not have intent as we understand the term in human behavior, it can conclude an intent as a result of a mathematical process developed in his software. Consequently, we need to define first in a legal framework for the behavior a robot may exhibit.

Furthermore, this fact opens the discussion for the sanctions that may be set into place. We believe that sanctioning the manufacturer will not cover all the situations that may arise. A *sui generis* sanction shall be created for the AI itself that can go as far as destroying a robot if it becomes dangerous.

The need to discuss the problem of using AI in the law field was raised by UN too, from a controlling perspective, as it follows:

Urgent action is needed as it can take time to assess and address the serious risks this technology poses to human rights, warned the High Commissioner: "The higher the risk for human rights, the stricter the legal requirements for the use of AI technology should be."

Ms. Bachelet also called for AI applications that cannot be used in compliance with international human rights law, to be banned. "Artificial intelligence can be a force for good, helping societies overcome some of the great challenges of our times. But AI technologies can have negative, even catastrophic effects if they are used without sufficient regard to how they affect human rights."[8]

This aspect opens a debate for another imminent problem related to the subject: protecting the human rights in the digital era.

[8] UN News: Urgent action needed over artificial intelligence risks to human rights. 15.9.2021. Online: https://news.un.org/en/story/2021/09/1099972 (quoted 1.12.2021).

2.2 Robots affecting human rights: study case regarding the right to a fair trial

From the judicial perspective, the human rights will be the most impacted by the use of AI. It may not be 100% of them, but for sure the vast majority will suffer alteration of content or exercise as the digital era develop. We will look at a brief analysis over the right to a fair trial and its expression nowadays.

The electronic file, the usual use of electronic signatures in trials, as well as the multitude of information provided in the virtual environment by public judicial institutions are some essential aspects that define the beginnings of a deeply restructuring process regarding the right to a fair trial. The COVID-19 pandemic offered an impulse of using the AI and its qualities in the justice field, giving the impression that AI will be only beneficial to the humanity. However, we would argue that it only opened the Pandora's box, letting loose a lot of ethical problems that we just started to get a glimpse of.

Prima facie, we emphasize that digitalisalition of justice can expand its scope of implementation, offering a potential to increase the level of defense insured at the constitutional level. In this sense, we exemplify by the possibility of the party to participate in the process, including in virtual format, by videoconference, in case of physical incapability of the person to appear. In addition, this format also facilitates the development of processes with a cross-border component, offering flexibility and speed.

However, the right of the party to appear in person is a component of the right of defense that cannot be denied and must always remain an option. It has been restricted in the context of the present pandemic, but it cannot be annihilated now or in the future. We cannot deny the importance of this fact, also revealed by the legislator's option to impose, in certain processes, for example those of family civil law, the presence in person of the parties in order to facilitate the communication. This applies both to the lawyers representing the parties and to the judge.[9]

Observing some of these aspects, on 21 April 2021, the European Commission presented a proposal for a regulation laying down harmonized rules

[9] BĂNICĂ, R. A.: Digitization of justice in the context of the COVID-19 pandemic and the implications of digitalization on constitutional rights. In: *Constitutional Law Review*, no. 2, 2020, p. 11, ISSN: 2457-8754. Online: http://www.revistadedreptconstitutional.ro/wp-content/uploads/1contents/2020_2/2020_2_Ruxandra_A_Banica_Digitization_of_justice_in_the_context_of_the_COVID_19_pandemic.pdf (quoted 29.10.2021).

on Artificial Intelligence (Artificial Intelligence Act[10]) and amending certain Union legislative acts. The proposal is supplemented by 9 annexes and it contains a broad regulatory framework for safely using the AI, including the justice field.

We emphasize that the proposal is here just in time, given the pandemic boost that digitalization received. The EU institutions sensed the need to intervene through a legislative initiative in this matter. We salute the proposal, but there is a lot more to be discussed and included in the regulation, despite the apparent density of the paper. In our opinion, the initiative is incipient, offering only very general definitions and broad regulations, with no specific chapters and specific measures to protect the vulnerable sectors, such as justice. However, a good basis to start is represented by the chapters of the regulation referring to the prohibited artificial intelligence practices and the code of conduct proposed. Also, a good perspective of the steps to take is offered through the calculation of the estimated financial impact.

The proposal already started to receive important comments and contributions from the practitioners in the law field, in order for it to be improved. We mention here the CCBE position paper on the proposal for a regulation laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act),[11] a paper that stressed in a realistic way many ethical aspects to be taken into consideration, views that we totally share and support:

a) **the exclusion of the use of AI tools** which may infringe a person's fundamental rights; for example: for the purposes of so-called "predictive policing" and for the purposes of determining risks of future offending as an aid to making of decisions as for example granting of a bail, imposing of a sentence, following conviction, making of decisions concerning probation and, generally, during prosecution and trial. Furthermore, the output of an AI system should not, on its own, be treated in judicial proceedings as having the status of evidence;

[10] European Commission: Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. COM (2021) 206 final. 21.4.2021. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 27.10.2021).

[11] Council of Bars and Law Societies of Europe: CCBE position paper on the proposal for a regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act). 8.10.2021. Online: https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Position_papers/EN_ITL_20211008_CCBE-position-paper-on-the-AIA.pdf (quoted 29.10.2021).

b) following **the transparency rules**: AI systems, if used in the justice system, do not obstruct the right to a fair trial and do not violate the rights of the defence. Moreover, given that the manner in which some AI systems produce their output may not be reasonably capable of explanation (the "black box" problem), and the fact that the transparency requirement might not always be met whether for that or some other reason, the proposal must provide for other safeguards, for instance the outcome provided by an AI tool must not be taken into account in case of doubt or when the requirements of transparency or explainability are not met;
c) **the inequality of arms** that may arise between the more advanced capabilities which prosecutors may have at their disposal and the more limited resources lawyers may have;
d) **non-allowing a judge to delegate all or part of his/her decision-making power** to an AI tool: there should be prohibited in the field of Justice not only automated decision making by AI systems but also the use of those AI systems which produce "decisions" of a nature which might tempt a human judge simply to adopt such decisions uncritically – effectively rubber-stamping what in effect would be automated decision-making.

In addition, we will highlight this idea that physical justice should always be an option for the citizens, in order to fully respect their rights. The right to be judged by a man can transform in time in a real human right, due to the AI delicate ethical nature. A natural distrust will always hover over the use of AI, for that no person will fully agree to have his acts judged by a robot that lacks empathy and emotional intelligence, vital elements when judging a case and expecting to respect a given decision. The responsibility of making a decision also pictures a key problem that will always considered by a human being: *The entire decision-making process must remain a human-driven activity and human judges must be required to take full responsibility for all decisions. An AI system may be used to "assist" judicial authorities, the possibility of it doing so to, in effect, reach decisions or formulate the expression of such decisions is excluded.*

Moreover, another aspect that thrives to be discussed is represented by the notion of judicial authority. If we start to delegate judges' and prosecutors' attributions to robots, irrespective of their lengths, does this not affect the judicial authority itself? It will need to be redefined with such care for the ethical perspective. In our opinion, the definition should be limited only to the hu-

man acts and not include the delegated attribution to the robots, because no scenario should include delegating authority to the robots.

The discussion may go deeper when we look to the judicial authority definition form a constitutional perspective, even if it may seem forced at first sight. Romanian Constitution provides in the first article that *the state is organized according to the principle of separation and balance of powers – legislative, executive and judicial authority – within the constitutional democracy*. Consequently, not defining clearly the meaning of the judicial authority and that it only resides in the human power may result in catastrophic meanings, giving the AI constitutional powers, interfering with the state affairs.

So, where should the limit be when AI interferes with the law professionals? Where is the future heading? This points out the last problem we would like to open for discussion: the future of law professions in the digital era.

2.3 Losing face: replacing experts in the field with artificial intelligence – is it ethical?

Reasons which supported the start of using AI in the law field resided in:

- lack of human resources, given the growing number of the case law;
- low costs implied by using a programme or a robot, supported also by the digitalization of justice, which includes online proceedings that are also cost effective;
- celerity of the process, leading to a better respect for the right to a fair trial.

Bearing in mind that digitization is only the first step of the digital transformation we're speeding to, what does the future of the law professions look like in the next century?

Regarding the issue of the disappearance of certain manual jobs through digitization, we consider the risks in the field of justice to be low. Most professions have a strong creative and intellectual component, as well as a low repetitive character, aspects that cannot be replaced by artificial intelligence at this time. However, as a small preview, we opine that the digital transformation in the law field will imply at least 50% of the sector to be automatised, with the same effects on the judicial professions. The benefits listed above are directly proportionate to the reality today, offering no argument to stop the process. However, with a careful approach to the ethical part of replacing judicial experts with robots, there will be mostly advantages in this regard.

Nevertheless, the profession of a judge remains a delicate problem. There are numerous initiative to substitute judges with robots, naming here the following:

a) the judicial decisions of the European Court of Human Rights (ECtHR) have been predicted to 79% accuracy using an artificial intelligence (AI) method developed by researchers at UCL, the University of Sheffield and the University of Pennsylvania;[12]
b) In Estonia, the simplest civil lawsuits, with a stake of less than 6,400 euros, are decided by an artificial intelligence system. Annually, 30,000 such cases are judged by a robot with a prior human interface. The online process can be initiated (on www.e-toimik.ee), because since 2001 every citizen has a verified electronic identity. You can even file for divorce online. Regarding the contractual disputes that involve the return of sums of money up to 6,400 euros, there is a semi-automated payment order system.[13]

Despite these initiatives, we keep expressing the point of view of not letting a trial be conducted by a robot. It may answer to a need for the judge to be objective and incorruptible, but the human component is vital in order for the trial to keep being seen as making justice.

The subject of law professions is also treated bluntly in the book written by Daniel and Richard Susskind *The Future of the Professions: How Technology Will Transform the Work of Human Experts*, who changed the perspective over professionals and our limitative beliefs:

Professionals play such a central role in our lives that we can barely imagine different ways of tackling the problems that they sort out for us. But the professions are not immutable. They are an artefact that we have built to meet a particular set of needs in a print-based industrial society. As we progress into a technology-based Internet society, however, we claim that the professions in their current form will no longer be the best answer to those needs.

Moreover, overcoming the cost effective part of digital justice when using the AI, the authors offer another interesting point of view: the preventive part of justice will develop by the use of AI. In this regard, there is a strong

[12] UCL News: AI predicts outcomes of human rights trials. 24.10.2016. Online: https://www.ucl.ac.uk/news/2016/oct/ai-predicts-outcomes-human-rights-trials (quoted 02.11.2021).

[13] SAVESCU & THE ASSOCIATES: Civil lawsuits with a stake of less than 6,400 euros are solved by a robot. JURIDICE.ro, 16.06.2019. Online: https://www.juridice.ro/643353/estonia-procesele-civile-cu-miza-de-sub-6-400-euro-sunt-rezolvate-de-robot.html (quoted 02.11.2021).

initiative in the matter in the last decade to develop the pre-trial and online dispute resolution mechanisms in order to decongest the courts, as following:

In another direction, interest is developing in embedding legal requirements into our social and working lives, so that, for example, automatic compliance with health-and-safety regulations can be integrated into the design of buildings that can identify and respond when temperature levels are above some statutory level. In this way, human beings do not need to know the law and make a conscious decision to comply, and consequently, lawyers' direct involvement is not needed.

There are moves also towards a new discipline—legal risk management—where the spirit is dispute avoidance rather than dispute resolution, and towards multi-disciplinary practice, where lawyers work alongside accountants, consultants, and tax specialists in providing an integrated professional service.[14]

Conclusion

The analysis over the implication of the AI in the law field needs to be continuous and deeper than the format of this paper may permit, but the present paper will offer some stepping stones regarding the stage of development existing today.

Resuming the ideas presented and their conclusions, we propose to retain the following aspects:

a) there is an urge to control the development of AI and to gain clarity over its liability and a possible intent. This may imply as a step *deep legislative framework changing*, *bigger number of human resources* or *heavier attributions for the existing ones*, such as double checking the coding and programing of the AI, the direct results that robots present, the effects of using the robots that exist, such as affecting the human rights;
b) an imminent need to update the sphere of the human rights, their content, exercise and even their existence (new human rights may be born, such as the right to internet, if the internet or digital means will be mandatory in different requests);

[14] SUSKIND R., SUSSKIND, D.: The Future of the Professions: How Technology Will Transform the Work of Human Experts, OUP Oxford, 2015, p. 92, ISBN 978-0198799078.

c) reviewing the legal professions, assigning time to analyze their position, anticipating their future, updating the professional training for them (training that should include digital skills or using AI in their activity).

So, after countless analyses of pros and cons for using AI in the law field, is it worth it? Not the easiness of the trial, not the comfort it offers to the the justice seekers, nothing should weight the most in the justice balance, but always keeping in mind that all these must be sacrificed if justice is not obtained. *Justice, and only justice, shall always be our motto*! (Woodrow Wilson)[15]

Bibliography

1. The New York Times: 2 Killed in Driverless Tesla Car Crash. Online: https://www.nytimes.com/2021/04/18/business/tesla-fatal-crash-texas.html (quoted 1.11.2021).
2. UCL News: AI predicts outcomes of human rights trials. 24.10.2016. Online: https://www.ucl.ac.uk/news/2016/oct/ai-predicts-outcomes-human-rights-trials (quoted 02.11.2021).
3. BAROCAS, S., SELBST, A. D.: Big Data's Disparate Impact. In: *California Law Review*, vol. 104, 2016, p. 671, ISSN 0008-1221. Online: https://ssrn.com/abstract=2477899 (quoted 1.12.2021).
4. BĂNICĂ, R. A.: Digitization of justice in the context of the COVID-19 pandemic and the implications of digitalization on constitutional rights. In: *Constitutional Law Review,* no. 2, 2020, p. 11, ISSN: 2457-8754. Online: http://www.revistadedreptconstitutional.ro/wp-content/uploads/1contents/2020_2/2020_2_Ruxandra_A_Banica_Digitization_of_justice_in_the_context_of_the_COVID_19_pandemic.pdf (quoted 29.10.2021).
5. Council of Bars and Law Societies of Europe. CCBE position paper on the proposal for a regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act). 8.10.2021. Online: https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Position_papers/EN_ITL_20211008_CCBE-position-paper-on-the-AIA.pdf (quoted 29.10.2021).
6. DURACH, F., MĂRCUȚ, M., PUCHIU, R., ȘTEFAN, V.: *De la digitalizare la transformare digitală în România*, Institutul European România, 2021. Online: http://ier.gov.ro/wp-content/uploads/2021/07/Policy-brief-nr.-9_iulie-2021_Final.pdf (quoted 1.11.2021).
7. European Commission: Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Ar-

[15] LYONS, T.: The Little Black Book of Lawyer's Wisdom, Skyhorse Publishing, 2016, p. 40, ISBN 978-1510704152.

tificial Intelligence Act) and amending certain union legislative acts. COM (2021) 206 final. 21.4.2021. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 27.10.2021).

8. European Parliament: Committee on Legal Affairs. Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), 27.01.2017. Online: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html#title2 (quoted 1.11.2021).
9. LO, D.: Can AI replace a judge in the courtroom?. UNSW Newsroom, 01.10.2021. Online: https://newsroom.unsw.edu.au/news/business-law/can-ai-replace-judge-courtroom (quoted 27.10.2021).
10. LYONS, T.: *The Little Black Book of Lawyer's Wisdom*. Skyhorse Publishing, 2016, ISBN 978-1510704152.
11. REUELL, P.: Seeking fairness in ads. The Harvard Gazette, 30.04.2013. Online: https://news.harvard.edu/gazette/story/2013/04/seeking-fairness-in-ads/ (quoted 1.11.2021).
12. SAVESCU & THE ASSOCIATES: Civil lawsuits with a stake of less than 6,400 euros are solved by a robot. JURIDICE.ro, 16.06.2019. Online: https://www.juridice.ro/643353/estonia-procesele-civile-cu-miza-de-sub-6-400-euro-sunt-rezolvate-de-robot.html (quoted 02.11.2021).
13. STĂNILĂ, L.: Artificial intelligence. Criminal law and the criminal justice system. Memories of the future, 2020.
14. SUSKIND R., SUSSKIND, D.: *The Future of the Professions: How Technology Will Transform the Work of Human Experts*. OUP Oxford, 2015, ISBN 978-0198799078.
15. UN News: Urgent action needed over artificial intelligence risks to human rights. 15.9.2021. Online: https://news.un.org/en/story/2021/09/1099972 (quoted 1.12.2021).

Author details

Ruxandra Andreea LĂPĂDAT
PhD student
Bucharest University of Economic Studies
Faculty of Law
ruxandrabanica@hotmail.com

3 THE IMPACT OF ARTIFICIAL INTELLIGENCE ON FUNDAMENTAL RIGHTS AND FREEDOMS AND *VICE-VERSA*

Radoslav Benko

Abstract

*The paper discusses the impact of the use of artificial intelligence on fundamental rights and freedoms and the potential impact of fundamental rights and freedoms norms on the development of the artificial intelligence systems. It outlines the cases of the breach of certain fundamental rights and freedoms by the use of artificial intelligence systems, the interplay between the fundamental rights and freedoms norms and the ethics and vulnerabilities of artificial intelligence systems. Finally, it outlines the role which fundamental rights and freedoms norms may play in constructing/development of artificial intelligence systems.**

Introduction

Artificial intelligence (hereinafter "AI") is everywhere and its development, deployment and use is moving forward rapidly. It brings many benefits, but at the same time it raises many concerns, e.g., impacts on human autonomy, privacy and fundamental rights and freedoms.[1]

Marvin Minsky, one of the founding AI scholars, defines AI as *"the science of making machines do things that would require intelligence if done by men"*.[2] Another founding scholar of AI, John McCarthy, defines AI as *"the science and engineering of making intelligent machines"*.[3] According to Stanford University report AI is *"the science and a set of computational technologies that are inspired by, but typically operate quite differently from, the ways people use their nervous systems and bodies to sense, learn, reason, and take*

* This paper was prepared within the framework of the research project VEGA no. 1/0643/20 "Legal and Ethical Perspective of Artificial Intelligence".

[1] Artificial intelligence in society. Paris: OECD Publishing, 2019. Online: https://doi.org/10.1787/eedfee77-en (quoted 1.11.2021).

[2] Report of the World Commission on the Ethics of Scientific Knowledge and Technology (COMEST) on Robotics Ethics from the 14th of September 2017, p. 17. Online: https://unesdoc.unesco.org/ark:/48223/pf0000253952 (quoted 1.11.2021).

[3] McCARTHY, J.: What is AI? / Basic Questions. Online: http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html (quoted 1.11.2021).

action".[4] One of a key term of AI and one of its sub-field is machine learning. Harry Surden defines it as *"computer algorithm that have the ability to learn or improve in performance over time on some task*".[5] It is a machine that learns from date over time. This learning is through *"a statistical process that starts with a body of data and tries to derive a rule or procedure that explains the data or can predict future data*".[6] Machine learning is currently used at a variety of tasks, from driverless driving to diagnosing diseases. A machine learning technique that uses structures called "neural networks" that are inspired by the human brain is designated as deep learning. There are three main specific machine learning approaches: (i) machine vision that allows computers to recognize and evaluate images (e.g., facebook tags people in photos), (ii) natural language processing that helps computer understand, interpret, and manipulate human language (e.g., Google Translate, chatbots), and (iii) speech recognition that allows computers to translate spoken language into text (e.g., Siri, Alexa). Artificial intelligence works through algorithm (e.g., neural networks are type of algorithm), even though not all algorithms involve AI. Algorithm is defined a set of guidelines that describe how to perform a task. Computer science defines algorithm as a sequence of instructions that tell a computer what to do.[7]

The legal writings relating to the deployment and use of AI cover many different legal issues. Some writings are of broad scope and cover variety of risks and challenges (e.g., unfairness, bias, discrimination or transparency of AI), some cover very specific issue (e.g., legal personality of AI, accountability for harms, liability for damage, insurance) and some of them are domain-specific (focusing on privacy and data protection, access to justice, healthcare, defence, intellectual property, cybersecurity, or transport). The issues are sometimes cross domains (manifested in one or more sector/field

[4] Report of Stanford University on AI (2018). Online: https://ai100.stanford.edu. Quoted from: ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018, p. 8. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (quoted 1.11.2021).

[5] SURDEN, H.: Machine Learning And Law. In: *Washington Law Review*, Volume 89, No. 1, 2014, p. 88, ISSN 1942-9983. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2417415 (quoted 1.11.2021).

[6] ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018, p. 8. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (quoted 1.11.2021).

[7] *Ibidem*, p. 9–10.

of application (e.g., privacy and data protection), or inter-related (e.g., transparency, fairness, accountability).[8]

Predictive, classifying, and profiling algorithms (from decision trees to deep neural networks) increasingly impact our lives and societies. The use of algorithms, in recent years, has raised major socio-ethical challenges, such as discrimination, unjustified action, privacy infringement, spread of discrimination, job market effects or safety issues. The development and the use of AI systems are consequential to fundamental rights and freedms, and the moral and social values they embody (e.g., human dignity, freedom, equality).[9]

The paper aims to outline the impact of the deployment and use of AI systems on the fundamental rights and freedoms and the impact of fundamental rights and freedoms protection norms on the development of AI systems. First, we will outline the benefits and concerns the deployment and use of AI systems encompass, then we will proceed to the threats of violation of certain fundamental rights and freedoms by the deployment and use of artificial intelligence systems and the role the fundamental rights and freedoms protection legal norms can play in effective application and enforcement of ethical values and principles. Finally, we will outline vulnerabilities of AI systems and well-designed AI approach as an effective tool for overcoming AI systems' vulnerability of technical nature.

3.1 Benefits and concerns of AI systems

The AI systems are used in almost each field of society life in order to improve it or make it easier or more effective. Therefore, many benefits can be attributed to the deployment and use of AI systems, e.g., improvements in services, safety, lifestyle, helping solve problems, improving access to healthcare and predicting disease, making government services more efficient and accessible, making life easier for the visually impaired, organizing agriculture and helping farmers adapt to change, mitigating climate change, predicting natural disasters and conserving wild life or speeding up the judicial

[8] Compare: RODRIGUES, R.: Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. In: *Journal of Responsible Technology*, Volume 4, 2020, ISSN 2666-6596, p. 2. Online: https://www.sciencedirect.com/science/article/pii/S2666659620300056#bib0117(quoted 1.11.2021).

[9] Compare: AIZENBERG, E., van den HOVEN, J.: Design for human rights in AI. In: *Big Data and Society*, Volume 7, issue 2, 2020, p. 1, ISSN 2053-9517. Online: https://journals.sagepub.com/doi/10.1177/2053951720949566 (quoted 5.11.2021).

proceeding by helping to search for various legal sources and making their analysis and composing the decisions.[10]

Along with benefits, the deployment and the use of AI systems faces many concerns, e.g., perpetuating bias in criminal justice, facilitating mass surveillance, enabling discriminatory profiling, assisting the spread of disinformation, perpetuating bias in job market or driving financial discrimination against the marginalized. In principle, the deployment and use of the AI systems impact human autonomy, privacy and fundamental rights and freedoms.[11]

3.2 Fundamental rights and freedoms touched by the deployment and use of AI systems

Deployment and use of AI impact almost each fundamental right and freedom. Civil and political rights, economic, social and cultural rights as well as the third-generation human rights (right to peace, right to economic and social development, right to a healthy environment, right to natural resources, rights to intergenerational equity and sustainability or right to participation in cultural heritage). Widely prevalent in AI legal discussions are the non-discrimination right, the right to privacy and data protection, the right to access to justice and the right to equality.[12] One of the main questions arising with respect to the deployment and use of AI systems is whether contemporary legislation on fundamental rights and freedoms protection is sufficient to face all fundamental rights and freedoms' issues of AI.

Quite comprehensive report scoping the impact of the deployment and use of various AI systems on the fundamental rights and freedoms was elaborated by non-profit association Access Now, which focuses to uphold fun-

[10] See: ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018, p. 14–15. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (quoted 5.11.2021).

[11] For the details, see: Artificial intelligence in society. Paris: OECD Publishing, 2019. Online: https://doi.org/10.1787/eedfee77-en (quoted 1.11.2021).

[12] Compare: RODRIGUES, R.: Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. In: *Journal of Responsible Technology*, Volume 4, 2020, p. 6, ISSN 2666-6596. Online: https://www.sciencedirect.com/science/article/pii/S2666659620300056#bib0117 (quoted 1.11.2021).

damental rights in the digital age, by leading author Lindsey Andersen.[13] The authors of the report, called Human rights in the age of artificial intelligence, focuses on how different AI systems are used in the current world and ways in which they can both help, or harm society. They engage in the interference of current and foreseeable uses of AI with a broad range of fundamental rights and freedoms.

For example, a recidivism risk-scoring AI software used across the U. S. criminal justice system to assist judge in their sentencing decisions, led to more black defendants falsely labelled as high risk and thus requiring them as kept in pre-trial detention or sentence to longer prison terms. By rating a defendant as high or low risk of reoffending, criminal risk scoring system attribute a level of future quilt, which may interfere with the discrimination, presumption of innocence and a fair trial. The system can build its outputs in existing police bias through the use of past data. Or the analysis of data by AI system may reveal private information about individuals (personal data), which should be protected according to respective law[14] even if being derived from big data sets utilizing publicly available information. The processing of personal data by AI systems represents a high potential risk of interference with the right to privacy and data protection. With the growth of the internet and the new technologies AI is enabling more invasive surveillance tools. For example, facial recognition system threatens the anonymity and the fear of being watched can obstruct people to exercise their other rights as the freedom of association or the freedom of religion which includes freedom, either alone or in community with others and in public or in private, to manifest religion or belief, in worship, teaching, practice and observance.[15] The AI systems may be used by the authoritarian undemocratic governments in monitoring and targeting members of persecuted religious groups. The AI systems may also provide a detailed picture of individuals' movements as well as predict future location and thus lead to infringement of freedom to movement.

[13] ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (quoted 1.11.2021).

[14] Within the European Union, contained mainly in 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). OJ L 119, 04/05/2016, p. 1–88.

[15] See: Article 10 para 1 of the Charter of Fundamental Rights of the European Union. OJ C 202, 7.6.2016, p. 391–407.

The use of AI systems by social media companies in order to remove a wide range of content comprising terrorist propagation, hate speech or so called "fake news" from their sites within a very short time (e.g., 24 hours) after the content has been posted, may cause the interference to freedom of expression if the content is remove in error.[16] Similar AI technology may be used by authoritarian governments to identify "politically sensitive" content and to increase censorship. The censorship can be used to restrict the freedom of association by removing groups, pages and content that facilitate organization gathering and collaboration. Social networks' algorithm determining the content of a user's newsfeed and influences how widely and to whom content is shared poses an indirect threat to freedom of thought because it shapes the type of information a user has access to.[17] Using AI systems in creating and spreading disinformation during the election campaign challenges the fair election and the right to political participation. Moreover, it may result in lowering participation in elections, if voters lose trust in the legitimacy of elections. Although the right to work, which requires states to try to achieve full employment, may be influenced by the use of AI systems when they are deployed in the automation of jobs. The AI systems may result in job loss in certain sectors and prevent some people from accessing to labour market. Despite the fact, that the AI applications in healthcare are of valuable help in diagnosing the disease and providing individualized patient treatment, AI-powered systems in healthcare may be programmed in a way that places cost reduction over the wellbeing of the patient or be designed to recommend different treatments depending on the insurance status of the patient and thus result in discrimination. Or the use of autonomous weapons AI systems can result in the death or injury of innocent civilians.[18]

[16] For example, YouTube removed over 100,000 videos documenting cruelties in Syria after they were posted because of being violent content. Such videos often serve as the only evidence of serious crimes and human rights violations. See: ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018, p. 22. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (quoted 1.11.2021).

[17] People of course have the ability to access other sources of information or find different opinions, but limited time and attention obstruct most people to do this. In countries without the free press and limited access to internet, social networks like Facebook may often be the only sources of unregulated free information.

[18] For the more detailed overview of the direct and indirect threats to violations of fundamental rights and freedoms, see: ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018, p. 18–30. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf. For another

Some of the mentioned threats to violations of referred fundamental rights and freedoms, caused by the deployment and use of AI systems in various fields of society life, are of direct and some of indirect character. While some of the threats are possible to be eliminated without huge effort, the elimination of the others requires finding more complex and far-reaching solutions. From the previous text we can also conclude that the interference of AI system with the respective fundamental right or freedom may be caused by the deficiencies in design or nature of AI itself or by their misuse by public power authorities, if not taking into account the interference cause by the careless and irresponsible behaviour of touched individuals.

3.3 The interplay of fundamental rights and freedoms norms with the ethics

At the beginning, the ethics discourse has largely dominated the discussion about "good" and "bad" AI, considering ethical concepts like justice, fairness, transparency, and accountability.[19] Ethics have helped those researching and developing AI to define boundaries for themselves (ethical principles to guide the AI initiatives have been developed for example by Google,[20] Microsoft,[21] and DeepMind[22]).

The norms of fundamental rights and freedoms protection interplay with the norms of ethics. Contrary to the ethical norms and principles, fundamental rights and freedoms are more universal, but mainly well-defined and binding, codified in a body of international and domestic law. They also provide for accountability and redress. International treaties on the protection of fundamental rights and freedoms lay down obligations which the contract-

comprehensive study on the issues and their significance, proposed solutions, gaps and challenges relating to the breach of fundamental rights and freedoms within the deployment and use of AI systems, see: RODRIGUES, R.: Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. In: *Journal of Responsible Technology*, Volume 4, 2020, p. 12, ISSN 2666-6596. Online: https://www.sciencedirect.com/science/article/pii/S2666659620300056#bib0117 (quoted 1.11.2021).

[19] See: ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018, p. 17. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (quoted 1.11.2021).

[20] See: AI at Google: Our Principles, June 7, 2018. Online: https://www.blog.google/technology/ai/ai-principles/.

[21] See: Microsoft AI Principles. Online: https://microsoft.com/en-us/ai/our-approach-to-ai.

[22] See: Exploring the Real-World Impact of AI. Online: https://deepmind.com/applied/deepmind-ethics-society/principles/.

ing states are bound to respect and fulfil. States must abstain from intervention in the rights and take positive actions to fulfil their effective enjoyment.

Fundamental rights and freedoms norms come from the ethics. Fundamental rights and freedoms norms define in more detail ethical principles and offer the remedy in case of their breach. If the use of AI is unethical, it most probably also breaches certain fundamental right or freedom.[23] Fundamental rights and freedoms protection systems thus might be used to enforce the ethical use of AI.

Therefore, fundamental rights and freedoms protection norms play an important role within the development, deployment and use of AI systems. Their effective application and enforcement decisively contribute to effective application and enforcement of ethical values and principles such as justice, fairness, transparency, and accountability in the development, deployment and use of AI systems.

3.4 Vulnerabilities of AI systems

The above-mentioned issues of direct and indirect threats to violations of fundamental rights and freedoms, by the deployment and use of AI systems in various fields of society life, are attached to the vulnerabilities of AI system. Vulnerability generally refers to: *"the quality or state of being exposed to the possibility of being attacked or harmed, either physically or emotionally."*[24] It can be characterized as a *"weakness that can be exploited by one or more threats or a pre-disposition to suffer damage"* or as the *"diminished capacity to anticipate, cope with, resist and recover from the impact"*.[25] Focusing on vulnerabilities of AI systems may help to overcome and consolidate the threats of AI systems to fundamental rights and freedoms protection and guide the development, deployment and use of AI system towards better protection of fundamental rights and freedoms.

Vulnerabilities of the deployment and use of AI systems (weaknesses of AI systems to face the threats of violation of fundamental rights and freedoms)

[23] Compare: ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018, p. 17. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (quoted 1.11.2021).

[24] RODRIGUES, R.: Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. In: *Journal of Responsible Technology*, Volume 4, 2020, ISSN 2666-6596, p. 7. Online: https://www.sciencedirect.com/science/article/pii/S2666659620300056#bib0117 (quoted 1.11.2021).

[25] *Ibidem*.

depend on diverse factors like that of: (i) physical/technical nature, e.g., inadequate (poor) design or development of AI system, or its inadequate security (protection) against the attacks, or inadequate safety measures; (ii) social nature, e.g., lack of public information and awareness about the AI systems and their impact, or lack of literacy, education, skills training; (iii) political nature, e.g., limited political will to recognize AI risks and pursue a strategy to address them, or limited systems of good governance; (iv) regulatory nature, inadequate legislation, or monitoring and enforcement mechanism, or ineffective remedies for harms; and (v) economic nature, e.g., insufficient resources to cope with the harmful effects of AI systems, or insufficient investments in safe and ethically compliant AI systems, or insurance.[26]

Within the next lines we will focus on presenting a possible solution for AI systems' vulnerabilities of physical/technical nature. The technical shortfalls of AI systems related to their technology design are subject of interest of the so called "well-designed AI" approach. It calls for embedding such requirements in the AI systems which could perfectly address human rights and ethical issues of their use.

3.5 Well-designed AI (design for values in AI)

The fundamental rights and freedoms protection legal norms may play an important role in the design of AI systems and thus in the development of AI systems. The AI systems unquestionably provide the society with many benefits, but at the same time confront the society with unjustified or discriminatory decisions wrongly assumed to be accurate because they are made automatically and quantitatively. In order to reach the transparency, explainability and fairness of the use of AI well-designed AI approach emerged, which calls for the design of algorithms and AI that addresses the needs consistent with fundamental rights and freedoms. Well-designed AI approach is based on the process-oriented approach, rather that solutions-oriented approach. It translates fundamental rights and freedoms protection into context-dependent design requirements of AI through a structured, inclusive, and transparent process.[27]

[26] *Ibidem*.

[27] See: AIZENBERG, E., van den HOVEN, J.: Design for human rights in AI. In: *Big Data and Society*, Volume 7, issue 2, 2020, p. 1–14. ISSN 2053-9517. Online: https://journals.sagepub.com/doi/pdf/10.1177/2053951720949566 (quoted 1.11.2021).

Well-designed AI approach tries to ground the design process of AI in the values of human dignity, freedom, equality and solidarity, being inspired by the values on which the EU is founded according to the Preamble of the Charter of fundamental rights of the EU.[28] All rights, freedoms and principles recognized by the Charter have their origin in one or more of these values. The design of AI for values turns the focus from technological artefacts and technical requirements of AI to include a proactive consideration of the societal context in which the technology is embedded, and how societal needs and values can be translated into socio-technological design requirements of AI.[29]

It builds on the gradual translation of abstract values into design requirements referred as value specification. First, values are expanded into norms which support desired values, e.g., to value privacy may in a certain context imply informed consent, confidentiality or the right to erasure. Value specification is about specifications of what the value (like that of privacy) means in a specific context of use. Then, each of the norms is then specified further into socio-technical design requirements, e.g., informed consent to processing of personal data is to be implemented as positive opt-in (requires explicit permission). The relationship between higher levels and lower levels in the hierarchy is not deductive in the sense that the norms in the example above may be insufficient or entirely irrelevant interpretations of privacy value in a different context of use. Value specification provides an explication of an abstract value and is always context dependent.[30]

Such a framework as described above in general proceed to moral discussions about engineering AI design which can contribute to moral learning and technical improvement of AI systems and raise their social legitimacy and trust towards them.[31]

To what extend an AI system can replace the actions of a human being in performing certain task, taking into account the social, normative, and institutional purpose of that task, is left open. For example, using care robot

[28] Charter of Fundamental Rights of the European Union. *OJ C 202, 7.6.2016, p. 391–407.* According to the Preamble of the Charter: *"Conscious of its spiritual and moral heritage, the Union is founded on the indivisible, universal values of human dignity, freedom, equality and solidarity; it is based on the principles of democracy and the rule of law."*

[29] See: AIZENBERG, E., van den HOVEN, J.: Design for human rights in AI. In: *Big Data and Society*, Volume 7, issue 2, 2020, p. 2, ISSN 2053-9517. Online: https://journals.sagepub.com/doi/10.1177/2053951720949566 (quoted 5.11.2021).

[30] *Ibidem*, p. 3.

[31] *Ibidem*, p. 3–4.

for the activity of lifting a patient from a hospital bed. The activity has a goal directed aspect consisting in safety raising the patient out of bed at certain angle and speed and safety placing her/him in a wheelchair. If the activity is done by a human being it has also the practice-oriented aspect consisting in the development of the bond and relationship between the caregiver and the patient. It has a social value for the patient as well, what might be important for the patient's long-term care.[32] Therefore, one of the challenges of the well-designed AI approach is to overcome the socio-technical gaps of the use of AI by revealing valuable human-to-human interactions, that may otherwise go unnoticed, and incorporate them into the design of AI system.

Conclusion

The development, deployment and use of AI systems impact the effective application of almost each fundamental right and freedom. Despite of many benefits, the deployment and use of AI systems raise many concerns. They are linked with the vulnerabilities of AI systems caused by various factors. Taking the well-designed AI approach into account within the development process of AI systems should provide the cornerstone for researchers in AI to conduct their further studies. Furthermore, the development, deployment and use of AI systems face the general legal challenges such as the lack of AI specific legal regulations at national as well as at international level, lack of potential new regulatory and monitoring bodies if the existing ones are failing, lack of clarification on the application of existing laws, lack of academic debates in some countries, lack of judicial knowledge and training, or the vagueness in the legal status of automated AI systems.[33] The deployment and use of AI systems should be emanating from the fundamental rights and freedoms impact estimations and from the principles of transparency and accountability. There is a need to create the safeguards against the misuse of AI systems and the actions violating fundamental rights and freedoms. In that respect it is important to distinguish between the deficiencies of AI system itself and its intentional use leading to violation of fundamental rights and freedoms.

[32] *Ibidem*, p. 4.

[33] RODRIGUES, R.: Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. In: Journal of Responsible Technology, Volume 4, 2020, ISSN: 2666-6596, p. 8–9. Online: https://www.sciencedirect.com/science/article/pii/S2666659620300056#bib0117 (quoted 1.11.2021).

Bibliography

1. AIZENBERG, E., van den HOVEN, J.: Design for human rights in AI. In: *Big Data and Society*, Volume 7, issue 2, 2020, 14 p., ISSN: 2053-9517. Online: https://journals.sagepub.com/doi/10.1177/2053951720949566 (quoted 5.11.2021).
2. ANDERSEN, L. et al.: Human Rights in the Age of Artificial Intelligence. Report of the Access Now, November 2018, 38 p. Online: https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf. (quoted 5.11.2021).
3. Artificial intelligence in society. Paris: OECD Publishing, 2019, 152 p. Online: https://doi.org/10.1787/eedfee77-en. (quoted 5.11.2021).
4. Charter of Fundamental Rights of the European Union. *OJ C 202*, 7.6.2016, p. 391–407.
5. McCARTHY, J.: What is AI? / Basic Questions. Online: http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html. (quoted 5.11.2021).
6. 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). *OJ L 119*, 04/05/2016, p. 1–88.
7. Report of Stanford University on AI (2018). Online: https://ai100.stanford.edu. (quoted 5.11.2021).
8. Report of the World Commission on the Ethics of Scientific Knowledge and Technology (COMEST) on Robotics Ethics from the 14th of September 2017, 64 p. Online: https://unesdoc.unesco.org/ark:/48223/pf0000253952 (quoted 5.11.2021).
9. RODRIGUES, R.: Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. In: *Journal of Responsible Technology*, Volume 4, 2020, 12 p., ISSN: 2666-6596. Online: https://www.sciencedirect.com/science/article/pii/S2666659620300056#bib0117 (quoted 5.11.2021).
10. SURDEN, H.: Machine Learning And Law. In: *Washington Law Review*, Volume 89, No. 1, 2014, p. 87–115, ISSN 1942-9983. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2417415 (quoted 5.11.2021).

Author details

JUDr. Radoslav BENKO, PhD., LL.M.
Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of International law and European law
radoslav.benko@upjs.sk

4 ARTIFICIAL INTELLIGENCE AS AN UNDUE INFLUENCE IN CRIMINAL TRIALS: ISSUING THE USE OF ALGORITHMS UNDER THE PRINCIPLE OF INDEPENDENCE OF JUDGES IN EUROPE

Theo Antunes

Abstract:
This paper will focus its approach on the extent under which the use of AI based algorithms in criminal courts could be qualified as an undue influence under European human rights law. However, this paper will also point at some leading evolution that could render a compatibility between the principle of independence and the use of algorithms as legitimate influences within the criminal trial. It will focus on the following question: At what extent the use of AI can be qualified as an undue influence under European human rights law?

Introduction

The principle of independence stands at the crossway of the criminal trial and the guarantees of the rule of law, its main aim is to render justice through a neutral, law-based and fact-based decision without any interference from other branches of power nor from any other undue influence and interference.[1] This paper would thus aim at investigating the impact and the challenges of the use of such advanced technology on the principle of independence of the judiciary as a cardinal principle of the rule of law and criminal justice.

Firstly, this paper will focus on the question of what an undue influence under European human rights law is, and to what extent such notion would reach. This qualification would be made through the common work of both the European Court on Human Rights and the European Court of Justice. This approach would lead to assess the common features of the courts toward identifying an undue influence, but also where divergences can be found, especially when focusing on the *rationae personae* extent of the notion. This work would also be led with the textual materials on the principle of inde-

[1] COX, A.: The Independence of the Judiciary: History and Purposes. In: *University of Dayton Law Review,* vol. 21, no. 3, 1996, p. 566, ISSN 0162-9174.

pendence in Europe such as the consultative council of judges or the Venice commission; in order to better grasp the notion of undue influence under European human rights law.

Secondly, this paper will focus on whether such algorithms would meet this qualification while referencing what type of algorithms exist in such trials nowadays (reoffending probabilities, analytic justice…). This part will be divided in two questions. To what extent the algorithm constitutes an influence to the judge, as a direct influencer (the outcome of the algorithm to the judge) and as an indirect influencer (the developer of the technology to the judge). And the second question would focus on the extent of the "undue" character of the algorithm especially in cases where the variables of the algorithm are weighted in a manner that could alter direly the decision of the judge if followed.

Thirdly and finally, this paper will focus on the emerging legislative solutions that could render the AI based algorithm as a legitimate interference in the criminal trial, on the use of such technology but also regarding the legislative framework. These solutions would be divided between the technology itself, the developers of such technology and the involved actors, and finally the attitude of the judge toward the use of such technology. More than these *rationae personae* approaches to regulations, this paper will also focus on the overall legislative framework needed to both guarantee the effectiveness of the technology while preserving judges from being submitted to an undue influence under European human rights law.

This paper would thus serve as a better understanding of the potential threats of the use of such technology in the criminal trials in Europe but also the benefits such technology could bring in these trials without, however, undermining the principle of independence of the judiciary by incorporating an undue influence within the criminal administration of justice.

4.1 The notion of undue influence: Defining the extent and the content of a complex notion under European law

The use of AI in criminal trials represents a revolution in judicial decision-making. Such technology, as developed today, could be an asset in helping the judge assessing a decision. The most prominent form of advice given by such a technology lies in the prediction whether a given person is likely

to reoffend.[2] Such a use however poses major challenges under the principle of independence of the judiciary, one of the major cornerstones of the rule of law in Europe. Indeed, the use of AI could pose a challenge, especially in the case of protecting judges against undue influences. The protection against undue influences is one of the four aspects of the principle of independence. However, it remains a blurry notion and its framework of application remains unsure. As such, this section of the paper would focus on framing the concept of "undue influence" in Europe. It will first demonstrate that both the EU and the Council of Europe are relevant in order to assess such notion, especially in assessing whether AI could qualify as one in criminal trials (A). Then, this section will focus of what kind of influences are permitted and which ones are not under European law (B). Finally, it will be necessary to assess what would characterize an "undue" influence *per se* (C).

4.1.1 The relevance of the council of Europe and European Union Law in assessing undue influence and the working of AI

The regime of protection of the independence of the Council of Europe is mainly embodied in Article 6 of the European Convention on Human Rights (ECHR).[3] Thus, in order to assess the relevance of such protection, it is necessary to assess the protection of human rights in the Convention framework as a whole. Under Article 1 of the ECHR, it is provided that: "The High Contracting Parties shall secure to everyone within their jurisdiction the rights and freedoms defined in Section I of this Convention."[4] Then, the application of human rights under such a framework is conditioned under the jurisdiction notion. Through the interpretation of the case-law of the European Court on Human Rights (ECtHR), one can see that the notion of jurisdiction is deeply linked with the territorial approach to responsibility. This means

[2] OSWALD, M., GRACE, J, URWIN, S., BARNES, G. C.: *Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality.* In: *Information & Communications Technology Law*, vol. 17, no. 2, 2018, p. 223–250, ISSN 1469-8404; Committee on Legal Affairs and Human Rights. Justice by algorithm – the role of artificial intelligence in policing and criminal justice systems. Provisional version. 2020, p. 8, 10.

[3] Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by protocols Nos. 11 and 14, adopted on 4 November 1950, ETS 5, Article 6 (1).

[4] Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by protocols Nos. 11 and 14, adopted on 4 November 1950, ETS 5, Article 1.

that States must ensure human rights within their territory, or any other area they have an effective and direct control upon.[5] Hence the approach of the ECtHR is of a general application of human rights of persons, pending that States have jurisdiction on them.

The protection of the independence in the EU and the application of human rights in general is more complex than the ECtHR one. Indeed, the protection of independence of the judiciary is enshrined at Article 47 of the Charter of the Fundamental Rights of the European Union.[6] However, the conditional application of human rights is found at Article 51 that provides "the provisions of this Charter are addressed to the institutions and bodies of the Union with due regard for the principle of subsidiarity and to the Member States only when they are implementing Union law."[7] When it comes to the Member States insurance of human rights, one can see it is only of a conditional approach of implementing EU law. In this perspective, it thus represents a special legal framework compared to the ECtHR one. Thus, the question: Is the use of AI as part of criminal trials part of the implementation of EU law? The answer is twofold.

First of all, the General Data Protection Regulation oversees the regulation of "automated decision making"[8] under which AI is one of the many aspects through the working of algorithms.[9] Moreover, the will of the EU to directly regulate AI can be demonstrated through the project of regulation of AI by the Commission as a special regulation of this technology, directly mentioning the use of AI as a criminal trial tool.[10] Secondly, under Arti-

[5] *Al-Skeini and Others v. United Kingdom*, no. 55721/07, § 131, ECHR 2011; *Soering v United Kingdom*, no. 14038/88, § 86, ECHR 1989.

[6] Article 47 of the Charter of Fundamental Rights of the European Union, OJ EU C 326/391, 26.10.2012 (hereinafter as "EU Charter"), Article 47.

[7] Article 51 of the EU Charter.

[8] Article 22 of the 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), OJ EU L 119/1, 4.5.2016 (hereinafter as the "GDPR Regulation").

[9] Data protection working party. Guidelines on Automated individual decision-making and profiling for the purposes of Regulation 2016/679, adopted on 3 October 2017, p. 8.

[10] European Commission. Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. COM (2021) 206 final. 21.4.2021, recital 6.

cle 288 of the TFEU, regulations are of direct effect on Member States.[11] This means that both the GDPR and the proposition of regulation by the Commission is and will be of direct effect, thus giving way for the application of the human rights provisions of the CFEU. And by extension, this means that the principle of independence would apply in such context.

Hence, even if the framework of the EU is of a special one, it would apply in such context of AI in the criminal trials frame. In this perspective, both the ECtHR and the CFEU would be relevant in order to specifically determine what an undue influence is.

4.1.2 Undue influence: What influences threaten the principle of independence?

The question of undue influence ultimately targets what influences can threaten the principle of independence. If some influences are regarded as legitimate (experts for instance), some are regarded as non-legitimate in the criminal trial. The question thus arises, what is the extent of influences that would be regarded as non-legitimate under European law. This then leads to assess what point of origins of the influence would not be compatible with the principle of independence. In order to answer such a question, it would be necessary to assess such an extent under both European frameworks and to draw whether they share the same approach of what would constitute an illegitimate influence.

The Council of Europe has developed two main approaches under the qualification of what an influence under the principle of independence would entail. Through the work of the organs of the Council of Europe, by recommendations and non-binding instruments; but also, through the work of the European Court of Human Rights when interpreting Article 6 of the Convention.

Regarding the textual basis of the Council of Europe, one can see most of the texts regarding the point of origins of the influence in a broad manner. Indeed, this approach does not pose any limit to where this influence may come from and depart away from the separation of powers approach.[12] This

[11] Consolidated versions of the Treaty on European Union (TEU) and the Treaty on the functioning of the European Union (TFEU), OJ EU C 326, 26.10.2012, Article 288 TFEU.

[12] European Commission for Democracy Through Law: European standards on the independence of the judiciary; Rule of Law Checklist, 2016 (Vienna); Consultative Council of the European Judges, Opinion N°1; Consultative Council of European Judges, Magna Carta of Judges (2010); Consultative Council of European Judges, opinion N°18/2015: Special

"extensive" approach, also shared by the work of Tescher, seems to encompass all kind of influences possible stemming from States and private companies alike.[13] Moreover, this approach is the most enshrined at the Council of Europe level. This leads to assess that under this non-binding approach of European law, the priority is given to an extensive approach of the influence.

However, the challenge also lies in the case-law of the ECtHR, the main adjudication body of the Council of Europe. The view of such body is less clear than the textual approach. Indeed, considering the extensive case-law of the ECtHR on what would constitute an illegitimate influence over a trial it had a long-term approach of the principle. This case law is of an irregular one, focusing on a plural approach of influences but not sticking to one approach in particular, rending the overall cases confusing as to defining the scope. In defining the extent of the influence, the latest ECtHR case law varies in four different approaches: Either it approaches illegitimate interferences from the executive and legislative;[14] from the parties to the case;[15] from "other entities"[16] (the press was considered as such an entity); or does not specify the point of origin of an illegitimate influence.[17] These different approaches taken by the ECtHR are rather confusing, for it seems to be of an irregular approach and there are no explicit reasons for using one specific terminology. This leads to a certain confusion when it comes to analyzing if the ECtHR sees the notion as an extensive one or rather stuck to the principle of separation of powers. However, due to the consideration the ECtHR has for the protection of human rights against non-state entities,[18] it could be led to assess that the protection against undue influence by non-state entities could be protected in

text regarding other branches but still extensive and similar to previous texts; European Charter on the Statutes for Judges (1998); Committee of Ministers, Judges, Independence, Efficiency and responsibility (2010).

[13] TRECHSEL, S.: *Human Rights in Criminal Proceedings*, Oxford University Press, 2005, p. 54.

[14] *Beg S.P.A v Italy*, no. 5312/11, § 128, ECHR 2021; *Guðmundur Andri Ástráðsson v Iceland*, no. 26374/18, § 219, 2020.

[15] *Anzelika v Lithuania,* no. 36093/13, § 78, ECHR 2020.

[16] *Khrykin v Russia,* no. 33186/08, § 28, ECHR 2011.

[17] *Industrial Financial Consortium Investment Metallurgical Union v Ukraine,* no. 10640/05, § 148, ECHR 2018.

[18] AKANDJI-KOMBE, J. F. *Positive Obligations under the European Convention on Human Rights: A Guide to the Implementation of the European Convention on Human Rights,* Human Rights Handbook, Volume 7, p. 5; WILDHABER, L.: The European Court of Human Rights in Action. In: *Ritsumeikan Law Review,* no. 21, 2004, p. 84, ISSN 2434-2424.

two ways. Firstly, by the positive obligations framework of States when they have to "secure" human rights in their jurisdictions as a preventive measure but also concerning investigative measures after the violation was committed.[19] Hence, one could see that under such a framework, non-state entities could be regarded as potential illegitimate influencers. However, apart from the influence of medias on a case, there are no concrete cases of non-state entities influencing the outcome of a judgement under ECtHR law.[20] It thus appears that the Council of Europe approach is neither regular nor clear when it comes to assessing the illegitimate influence notion.

The European Union approach of the notion is different from the Council of Europe one, however. For decades, the EU had taken a similar approach as the CoE departing from the *Johnston* case, where it assimilated the right to fair trial as a general principle of EU law referencing to the case law of the ECtHR on the matter.[21] Hence, there was a certain uniformity in both jurisdictions. However, as it seems that the EU built its own approach of human rights through their own textual instrument, the interpretation of the principle of independence has also evolved. From a general principle of EU law, it became primary law with the enactment of the Lisbon Treaty and became a milestone in the values of the EU. Since such enactment, the CJEU has held its own approach of the principle of independence and by extension, the notion of undue influence. The CJEU has thus held that the principle of independence should be erected against those influences no matter where they come from, either direct or indirect, that can alter the judge's decision and threaten the autonomy of the judicial decision.[22] In its latest case law, the CJEU has maintained a robust and continuous approach of the principle of independence.[23] In this perspective any external or internal influence whether direct or indirect that could alter the judge's decision should be regarded

[19] STOYANOVA, V.: *Fault, knowledge and risk within the framework of positive obligations under the European Convention on Human Rights.* In: Leiden Journal of International Law, vol. 33, no. 3, 2020, p. 606, ISSN 1478-9698.

[20] *Paulikas v Lithuania,* no. 57435/09, § 62, ECHR 2017; *Craxi v Italy,* no. 34896/97, § 104, ECHR 2002; *G.C.P v Romania*, no. 20899/03, § 48, ECHR 2011.

[21] Court of Justice of the European Union (hereinafter as "CJEU"), Judgment of 15.5.1986, *Johnston v Chief constable of the Royla Ulster Constabulary,* C-222/84, para. 18.

[22] CJEU, Judgment of 19.11.2019, *A.K v Krajowa Rada Sądownictwa, and CP and DO v Sąd Najwyższy,* C-585/18, C-624/18, C-625/18, para. 153.

[23] CJEU, Judgment of 4.5.2017, *Banco de Santander SA,* C-274/15, para. 57; CJEU, Judgment of 27.2.2018, *Associação Sindical dos Juízes Portugueses v Tribunal of Contas,* C-64/16, para. 41.

as illegitimate.[24] From this perspective, one could see that this constitutes a clearer approach than the ECtHR but also more extensive in its mechanisms. This approach tends to maximize the reach and extent of the application of the protection against any undue influence, one can see that the CJEU decided to focus on a consequence approach of the influence, since it does not matter where it comes from, but such an influence must not alter judge's decision and threaten their decisional autonomy.

One can see that the overall European approach of what is an illegitimate influence is not clear, whereas the EU system seems to have established a robust and continuous approach, the Council of Europe approach is still fragmented on such a complex notion between extensive and restrictive approach, between the ECtHR and the non-binding instrument. However, more than approaching the question of the influence; it is necessary to assess what constitutes an undue influence under European law.

4.1.3 Undue Influence: What qualifies an influence as undue?

If the question of the influence focuses on the point of origin of the influence, meaning an illegitimate actor of the judiciary decision-making, one also has to focus on the "undue" character of an influence in order to grasp the full understanding of the notion. For non-judiciary actors can actually participate at some extent in the decision making (medical experts, psychological experts…) and influence the judge on the decision that should be taken.[25] But what if the decision reached by these non-judicial actors was inappropriate? Or completely irrelevant to the case at hand? Could it be qualified as a legitimate influence on one hand, but of an undue character?

Through the case-law of both jurisdictions and the non-binding instruments, one could grasp the meaning of what undue is. The European Union through the case law of the CJEU seems to have approach the undue influence by a purposive goal. Meaning that the influence in order to be undue must have had for purpose to "alter" the judge's decision having for effect to rip the judge of its "autonomy" during the judicial decision-making.[26] In this approach then, it is the effect of the influence that will render it undue,

[24] CJEU, Judgment of 24.6.2019, *European Commission v Republic of Poland,* C-619/18, para. 112.

[25] *Constantinides v Greece*, no. 76438/12, § 41, ECHR 2016; *Muller v Germany,* no. 54963/08, § 51–52, ECHR 2014.

[26] CJEU, Judgment of 24.6.2019, *European Commission v Republic of Poland,* C-619/18, para. 112.

because it strikes the judge in the core of its functions, the judicial decision making, by making lose to the judge the power to judge. This approach would aim at preserving the core function of the judge, but one might ask what of the influences that would not take away the autonomy of the judge? It seems that the European Union approach thus embraces a comprehensive approach with these influences that could help a judge reaching a decision as long as it does not make lose the judge its autonomy in the decision-making.

The CoE approach of what is an undue influence is however more complex, fragmented between the textual approach and the ECtHR. Under the textual approach, it seems that the main approach aims at protecting any influences deemed as "inappropriate".[27] But the question remains, what would inappropriate entail under the human rights regime? Pursuant to the Oxford dictionary definition of appropriate, the definition aims to encompass something that would be "suitable, correct or acceptable"[28] depending on the circumstances of the event at hand; something inappropriate would thus be the opposite of such a definition. But how to adapt such a definition to the human rights regime. Such an inappropriateness could thus take the form of an influence whose content would have a discriminatory component for instance or that would undermine other human rights in a disproportionate way. This could lead to believe that any influence that would undermine the protection of another human rights during the criminal trial could be inappropriate. This approach is however non-satisfactory as to the psychological expertise for instance, that may constitute a breach of privacy under Article 8 but would still be appropriate at some extent due to the limitative regime found in paragraph 2 of the same provision. Thus, is the undue character of an influence submitted to the same derogatory regime than the human rights general framework? This answer is not provided under the CoE framework but makes one think about the extent of the inappropriateness of an influence and whether it would have been qualified as undue.

The ECtHR approach is however silent on the issue of what an undue could be qualified like. But hints of what it is framing upon can be determined. Indeed, one of the few times where the ECtHR had to assess on whether an "ex-

[27] European Commission. Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. COM (2021) 206 final. 21.4.2021, recital 6.

[28] Oxford Dictionary. *Appropriate*, Online: https://www.oxfordlearnersdictionaries.com/definition/english/appropriate_1 (quoted 8.11.2021).

ternal undue influence"[29] was found in the case; it was cases implying media influence on the case that could have led to an undermining of the presumption of innocence of the person on trial.[30] The fact that such an influence could have undermined such protection might have been the reason why it was called "undue", for many media can comment on a case without, however, been qualified as undue. Therefore, one could assess that more than the point of origin of the influence, it is the potential undermining of another human right protection could make an influence undue.

Therefore, as for the question of what constitutes an undue influence under European human rights law, this paper would adopt the approach of: firstly, an influence that it deemed illegitimate for the judicial decision making, meaning an actor that can actually influence a decision; secondly, it would be the capacity to undermine the autonomy of the judge's decision when it is deciding upon both the culpability and the sentence of an individual.

4.2 AI in criminal trials as an undue influence? Identifying the major points of tensions between such use and the principle of independence under European law

As mentioned above, an undue influence must have the capacity to influence a decision in an undue manner, thus the question lies: To what extent can we qualify AI in criminal trials as potential undue influence?

Artificial intelligence can represent an asset for the Judge when deciding a case. However, as above-mentioned, such technology should not influence the judge in an "undue" manner. The question thus lies, what components would qualify AI as such? In order to assess this question, it will be necessary to approach it with two major axes. The first axe will focus on the influence paradigm that the technology brings, in both direct and indirect means (A). It will then focus on the "undue" potential aspect of the AI and assess what components of the technology could qualify it as such (B).

[29] STOYANOVA, V.: *Fault, knowledge and risk within the framework of positive obligations under the European Convention on Human Rights.* In: Leiden Journal of International Law, vol. 33, no. 3, 2020, p. 606, ISSN 1478-9698.

[30] *Ibidem.*

4.2.1 AI as an influencer over the judicial decisions: Drawing the extent of the influence paradigm

The potential influence of an AI can be established through two different vectors. Firstly, through the direct working of the algorithm that provides for an outcome to a legal issue (for instance the probability of reoffending of an individual), and the extent under which this would constitute an influence on the judge. Secondly, more of an indirect influence, from the developers of the AI to the judge, where AI is perceived a mere token of the developers reasoning and values.

The algorithmic technology can be one that can directly influence the decision of the judge. However, it is necessary to assess these different vectors of influences to better determine whether such technology could be qualified as undue influence. One could see that there are three different ways under which such technology can direct influence the judge. These influences however share a certain degree of extent on its reach on the judge.

The influencing process known as technological anchoring is of major importance when considering a use by the judge of the AI technology during the decision making phase. This term reflects upon two features. The anchoring effect and the high level of technology of the AI, the anchoring effect is a decision process under which a person would likely be anchored by one piece of information over others.[31] In the judicial decision making, it was demonstrated that a judge would likely be anchored regarding the first piece of evidence held for or against an individual;[32] the judge's decision would likely be anchored from it. This means that the judge's decision would start from this piece, and it would be hard for the judge to go against the mindset the first piece of evidence created. However, the technological anchoring is based not on the chronology of the evidence, but on the argument that such technology is designed and promoted as being a highly precise and expert technology in its functioning.[33] This component of high technology leads to

[31] GOLDSZLAGIER, J.: L'effet d'ancrage ou l'apport de la psychologie cognitive à l'étude de la décision judiciaire. In: *Les cahiers de la Justice*, no. 4, 2015, p. 524, ISSN 1958-3702.

[32] TVERSKY, A., KAHNEMAN, D.: Judgment under Uncertainty: Heuristics and Biases. In: *Science*, vol. 185, no. 4157, p. 1124–1131, ISSN 1095-9203; TVERSKY, A., KAHNEMAN, D.: Subjective probability: A judgment of representativeness. In: *Cognitive Psychology*, vol. 3, no. 3, 1972, p. 430–454, ISSN 1095-5623; KAHNEMAN, D., TVERSKY, A.: On the psychology of prediction. In: *Psychological Review*, vol. 80, no. 4, 1973, p. 237–251, ISSN 0033-295X.

[33] KITCHIN, R.: Thinking critically about researching algorithms. In: *Information, Communication & Society,* vol. 20, no. 1, 2017, p. 14, ISSN 1369-118X; WESSLEN, R. et al.: Anchored

assess that decision makers can let themselves influenced more easily than if it was a less advanced means of advising. Hence the technological anchoring could be a direct influencer in such a way, where the algorithm, because of its functioning, could lead to subconsciously influencing the judge and it would be hard for the judge to depart from the result of the algorithm.

The glistening simulacrum is the pervasive effect of the AI, it is the capacity of the technology to best or to fool humans in the decision-making process.[34] In this perspective it could influence the judge to take a decision that would be unethical or totally disproportionate to the situation at hand.

The "sheep effect" is one that poses a major challenge to the independence of the judges.[35] Indeed, this type of influence is one that transforms the judge from an actual decision-maker to a token of the algorithm outcome.[36] This means that the judge is only applying what the algorithm found without necessarily explaining the reason for such decision. In this case, the algorithm implicitly takes over the decision process, which represents real danger for the principle of judicial independence in European human rights law.

One can see that the judge could thus be influenced at different extent by the use of the algorithm. Such direct influence pose vital issues to the principle of independence, especially in cases where the judge would not dare go against the recommendations of the algorithm. In order to lower such an influence, it would be necessary to make such an algorithm better explainable and transparent as it will be demonstrated in part III.

More than a direct influence, AI could also represent an indirect means for illegitimate actors to play a role in the development of the AI. As such, AI could represent a conduit of their values and reasoning that could influence

in a Data Storm: How Anchoring Bias can affect user strategy, confidence, and decisions in visual analytics. University of North Carolina, 2018, p. 2. Online: https://researchain.net/archives/pdf/Anchored-In-A-Data-Storm-How-Anchoring-Bias-Can-Affect-User-Strategy-Confidence-And-Decisions-In-Visual-Analytics-3214328 (quoted 1.12.2021)

[34] SPAULDING, N. W.: Is Human judgement necessary?: Artificial Intelligence, Algorithmic Governance, and the Law. In: DUBBER, M. D., PASQUALE, F., DAS, S. (eds.): *The Oxford Handbook of Ethics of AI,* Oxford University Press, 2020, p. 377, ISBN 9780190067397.

[35] GARAPON, A.: Le numérique est un remède à la lenteur de la Justice. DALLOZ. Lefebvre Dallox. Actualité, 4.5.2018. Online: https://www.dalloz-actualite.fr/interview/antoine-garapon-numerique-est-un-remede-lenteur-de-justice#.YYjmCWDMJPY (quoted 8.11.2021).

[36] Data protection working party. Guidelines on Automated individual decision-making and profiling for the purposes of Regulation 2016/679, adopted on 3 October 2017, p. 21.

the judge on the judicial decision making.[37] The challenge would regard both traditional actors of the principle of independence, such as the executive, but also "legal techs", private entities that have the potential to interfere directly with the judicial decision making. Indeed, despite the name, artificial intelligence does not act like a real intelligence, it is submitted to the programming of its developers and the data it is fed with.[38] In this perspective, the judgement of the developers to consider some variables instead of others is displaced to the algorithm. Under such a theory, the algorithm would only be the extension of the developers, way of thinking and judgement.

The public actors, meaning the executive, may use such algorithms as a conduit of their own values and influence the outcome of a trial. Whereas under European law, the executive is forbidden to do so, whether through direct or indirect means,[39] such algorithms could be another means for the executive to threaten judicial independence; for these algorithms are often developed by private companies and deployed under executive monitoring.[40]

The private actors, meaning the legal techs, also represent an indirect threat to the judicial independence through the use of algorithms. This challenge emerges as more and more private companies are being delegated public powers and services by the state, such as military and security for instance.[41] However, the power to render justice has always been perceived as one that is non-delegable.[42] The fact that algorithms developed by these private entities and use by the judge in criminal trials can however lead to tremendous consequences on the rendering of criminal justice. These algorithms are indeed developed and built by these private companies; but even

[37] CONTINI, F.: Artificial Intelligence: A New Trojan Horse for Undue Influence on Judiciaries? UNODC. Online: https://www.unodc.org/dohadeclaration/en/news/2019/06/artificial-intelligence_-a-new-trojan-horse-for-undue-influence-on-judiciaries.html (quoted 8.11.2021).

[38] KATSUYA ENDO, S.: Technological Opacity & Procedural Injustice. In: *Boston College Law Review,* vol. 59, no. 3, 2018, p. 851–857, ISSN 0161-6587.

[39] *Agrokomplex v. Ukraine,* no. 23465/03, § 135–141, ECHR 2011; CJEU, Judgment of 24.6.2019, *European Commission v Republic of Poland,* C-619/18, para. 112.

[40] LEROUX. O.: Justice pénale et algorithme. In: HUBIN, J. B., JACQUEMIN, H., MICHAUX, B. (eds.): *Le Juge et l'algorithme : Juges augmentés ou justice diminuée*, LARCIER, 2019, p. 58, 68, ISBN 9782807911161.

[41] KRENT, H. J.: *The private performing the Public: Delimiting Delegations To Private Parties.* In: University of Miami Law Review, vol. 65, p. 523; CRAWFORD, J.: *State Responsibility,* Cambridge University Press, 2013, p. 123, ISBN 9781139033060.

[42] EMMERSON, B. et al.: *Human Rights and Criminal Justice*, 3rd Edition, London Sweet and Maxwell, 2012.

if the private sector can be hired to develop products that would help the judicial sector, these products were targeting only ancillary activities of the judicial power that did not have an impact on the core decision as to alter the judgement outcome. Such new product however would target the core of the judicial decision, by advising the judge on a potential outcome concerning an individual during the criminal trial. As mentioned above, the judge can suffer from the sheep effect, meaning turning into a mere token of the algorithm outcome, this would mean that the developers are indirectly the one deciding on a judicial decision. This process of influence is dire when it comes to saving judicial autonomy from undue influence.

4.2.2 AI as an undue component: Defining the undue in the AI

Moreover, even if the Judge is somehow free from the influence of the AI, such an AI should not be built in a way that would render it an "undue influence". As mentioned above, what constitute the undue of an influence could be the content of such an influence. In this instance, it is vital for the algorithm upon which the AI is built to not contain any variables that could lead the judge to an inappropriate decision. Here lie two main approaches: the first is the choice of variables and data would be important in order to avoid any discriminatory outcome from the algorithm. Secondly, the very functioning of the algorithm and the fact that it is shrouded in a "black box" could qualify it as an undue influence.

What would make an AI undue in the criminal trial? This question lies at the very core functioning of the AI and one potential led to answer this question would be the question of the variables, meaning the data by which the processing would lead the algorithm to provide a solution.[43] However, one has to be careful with the choice of these data, under many considerations:

Firstly, the choice of some data that are biased from the start. Indeed, it was demonstrated that some of these algorithms used in criminal trials relied upon data gathered by the state and especially from the police forces.[44] How-

[43] BERK, R.: An impact assessment of machine learning risk forecasts on parole board decision and recidivism. Working Paper No. 2016-4.0, University of Pennsylvania, Department of Criminology, 2016, p. 4–5. Online: https://crim.sas.upenn.edu/sites/default/files/WP2016-04_Berk_MachineLearningParole_08.03.2016%281%29.pdf (quoted 1.12.2021); BENSOUSSAN, A., BENSOUSSAN, J.: *IA, Robots et droit*, 1re edition, BRUYLANT, 2019 Brulyant, p. 102, ISBN 9782802763673.

[44] RHUE, L.: *Anchored to Bias: How AI-Human scoring can induce and reduce bias due to the anchoring effect'*. University of Maryland, 2019.

ever, for some of these algorithms it was demonstrated that the data were biased as they were collected, thus when the normal functioning of the algorithm, by processing these data, was already biased from the start and led to discriminatory outcome as the bias were perpetuated and repeated.[45]

Secondly, the challenge lies on what kind of data can be setup in the algorithm and are adapted to the programming of this algorithm. However, when one takes into account too much data or irrelevant data, this might create a disparate outcome, as irrelevant factors are taken into account to provide a solution. This was established through the working of the COMPAS algorithm that was taking too many factors into account, leading to give discriminatory outcomes on Afro-American population.[46]

Thirdly, data that can indirectly lead to a discriminatory outcome. These are data that as such are not discriminatory (family name, residential city...), but taken together as a group of data the algorithm would deduce that a certain person belongs to a certain minority and as a result, might take a decision that would lead to a discriminatory outcome for these minorities.[47]

The functioning of the AI can also be a source of undue influence under European law. More than the data it is using, one should also focus on the functioning of the algorithm, meaning how does the algorithm comes to an outcome, from the data to the solution. Such undue character can be exposed in two different approaches, firstly through the weight of the variables when the functioning is known: and secondly through the "black box" paradigm.

It was demonstrated through the use of certain algorithms, that some of the calculations made by it could be flawed and disproportionate. It is the case of the HART algorithm, that was deployed in England as a means for the local police to assess whether a person would be likely to reoffend. Individuals were then categorized as a high probability, medium probability and low probability reoffender.[48] This outcome is determined through a weight of the

[45] BARABAAS, C.: Beyond bias: "Ethical AI" in Criminal Law. In: DUBBER, M. D., PASQUALE, F., DAS, S. (eds.): *The Oxford Handbook of Ethics of AI,* Oxford University Press, 2020, p. 745–758, ISBN 9780190067397.

[46] ANGWIN, J. et al.: Machine bias: there's software used across the country to predict future criminals. And it's biased against blacks. Propublica, 23.5.2016. Online: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (quoted 1.12.2021).

[47] OSWALD, M., GRACE, J, URWIN, S., BARNES, G. C.: *Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality.* In: *Information & Communications Technology Law*, vol. 17, no. 2, 2018, p. 228, ISSN 1469-8404.

[48] OSWALD, M., GRACE, J, URWIN, S., BARNES, G. C.: *Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality. In: Infor-*

different variables that were installed during the development phase. However, it was demonstrated that the algorithm was developed in a way that was minimizing the threat posed by individuals. In order to do so, the algorithm overestimated the risk of some individuals in order to incapacitate threats more effectively.[49] However, such overestimation could lead to decide a unbased harsher sentence on an individual due to an overestimation of a risk by the algorithm. This sentence, based only on the theory of incapacitation and without concrete basis other than overestimated an individual could be qualified as undue.

However, most of the time, the algorithmic process is not known to the user or the individual whom the decision will concern; and sometimes to the developers themselves it is called the "black box".[50] In this perspective the user, meaning the judge, would not know the process under which the algorithm would decide of a solution; adding to this the paradigm of the direct influence it can create on the judge, would create a decision that can be based on a technology that is not transparent thus based on information not available.[51] Thus, considering the independence of the judge, how can one judge explain the decision if the reasoning of one of his means is unknown to him and the developers.

Hence the functioning of the algorithm could be qualified as undue, because of the high potential of bias it can create but also through its unknown functioning that could lead to a disproportionate sentence without the individual to know what led the algorithm to decide so.

4.2.3 AI in criminal trials as a legitimate influence? Drawing a compatibility scheme between the principle of independence and the use of AI in criminal trials.

Artificial intelligence thus can be qualified as an undue influence under European law if no guarantees or principles are applied to its inner working and its use by the judge. In this perspective, if a use of AI tends to be generalized

mation & Communications Technology Law, vol. 17, no. 2, 2018, p. 227, ISSN 1469-8404.

[49] *Ibidem,* p. 236.

[50] YEUNG, K.: Algorithmic regulation: A critical interrogation. In: *Regulation & Governance*, vol. 12, no. 4, 2018, p. 516, ISSN 1748-5991.

[51] European Commission for the Efficiency of Justice. European ethical charter on the use of Artificial Intelligence in judicial systems and their environment, adopted on 3–4 December 2018, para. 137, 159, 160. Online: https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (quoted 1.12.2021).

in some legal systems, then this technology would have to be adapted to the principle of independence. In this perspective, it will be necessary to assess such compatibility under two perspectives, one considering the technology itself (A); while the other would focus on the legal framework that would render the use of such technology compatible under European law (B).

4.2.4 What principles for the governance of algorithms in criminal trials: Addressing the unveiling of the undue character of AI in criminal trials?

The principle of transparency is one that is often promoted by human rights bodies as part of the principle governing a trustworthy AI.[52] Such a principle expects the algorithm to unveil the inner working of its functioning, from the data used, to the model developed and what happens between the introduction of data and the outcome of the algorithmic solution. This publicity would allow for more clearance in the path of the algorithm solution and witness what path he did take in order to reach its conclusion. However, such a transparency is problematic considering the proprietary nature of the algorithm as a private company product and thus leading to be protected under intellectual property law.[53] Such an issue has already been presented in front of U.S courts in the cases of *Kansas v Walls* and *Loomis v Wisconsin* where the intellectual property was confronted to due process rights.[54] Such intellectual property protection is also recognized in the GDPR recital 61, where it is provided forth that intellectual property law should not be breached considering the publicity of the algorithms.[55] Such provision was also held in the Loomis case and prevented any disclosure of the inner working of the algorithm; however, the Walls case has also presented an exception to this rule, if the algorithm is the sole piece of evidence under which the judge gives a de-

[52] *Ibidem*, principle 4; LESLIE, D. et al.: Artificial intelligence, human rights, democracy, and the rule of law: a primer. The Council of Europe, 2021, p. 21. Online: https://www.turing.ac.uk/sites/default/files/2021-03/cahai_feasibility_study_primer_final.pdf (quoted 1.12.2021).

[53] YEUNG. K., HOWES, A., POGREBNA, G.: AI Governance by Human Rights–Centered Design, Deliberation, and Oversight: An End to Ethics Washing. In: DUBBER, M. D., PASQUALE, F., DAS, S. (eds.): *The Oxford Handbook of Ethics of AI,* Oxford University Press, 2020, p. 77–108, ISBN 9780190067397.

[54] *Loomis v Wisconsin,* [2016], Wisconsin Supreme Court, N° 16-6387; *Kansas v Walls* [2017], Court of Appeal of the State of Kansas, N°116,027.

[55] Recital 61 of the GDPR Regulation.

cision, then the intellectual property of the algorithm cannot be invoked, and the algorithm must be disclosed to the defendant.[56] Such transparency thus remains vital in order to make sure that no undue functioning could have influenced the judge in its decision.

However, the complexity of the algorithm and its functioning requires more than transparency. Indeed, judges are no computer scientists, and even if the algorithm is transparent to them, there will be an issue on how to explain the outcome of the algorithm.[57] The reasoning of the judge can represent an assurance against a potential undue influence, for the reasoning explains the reason of the decision. However, if a judge based a decision upon the algorithm outcome but cannot explain neither the reason it relied upon it nor how the algorithm came to such a conclusion, this could represent a problem of assurances against undue influence. Then, more than transparency such algorithms should answer to the principle of explainability in order to unveil the inner functioning and to assess whether such algorithm reasoning was tainted in an undue manner.[58] It could also lead the judge to intervene and not consider the algorithm if it did act in such a manner. Therefore, taken together the principle of transparency and explainability could represent a barrier to the undue character of the influence in order to preserve the integrity of the principle of independence against such influences. Such publicity in both the knowledge and the understanding of the functioning would be the first step toward building trustworthy algorithm, upon which the EU has planned its strategy for such technology.[59] Indeed, the knowledge and understanding of such technology would lead to assess what data might cause an undue component in the influence scheme. The

[56] *Kansas v Walls* [2017], Court of Appeal of the State of Kansas, N°116,027; ZAVRŠNIK, A.: Criminal Justice, Artificial intelligence systems, and human rights. In: *ERA Forum,* vol. 20, 2020, p. 574, ISSN 1863-9038.

[57] BODDINGTON, P.: Normative modes: Codes and Standards. In: DUBBER, M. D., PASQUALE, F., DAS, S. (eds.): *The Oxford Handbook of Ethics of AI,* Oxford University Press, 2020, p. 135, ISBN 9780190067397.

[58] DIGNUM, V.: Responsibility and Artificial intelligence. In: DUBBER, M. D., PASQUALE, F., DAS, S. (eds.): *The Oxford Handbook of Ethics of AI,* Oxford University Press, 2020, p. 225, ISBN 9780190067397; LESLIE, D. et al.: Artificial intelligence, human rights, democracy, and the rule of law: a primer. The Council of Europe, 2021, p. 37.

[59] High-level expert group on Artificial intelligence set up by the European Commission, Ethic Guidelines for Trustworthy AI, European commission, published on 8 April 2019; European Commission. Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. COM (2021) 206 final. 21.4.2021, p. 1.

irrelevant and not necessary data could thus be corrected in further implementation of the algorithm.

Thus, these two principles remain essential in order to build fair algorithms and for safekeeping the protection of the judge against undue influences under European law; nevertheless, in order to further protect such vital component of the rule of law, not only it is essential to limit the undue; but also, the influence.

4.2.5 What extent of the influence of AI in the criminal trials: Drawing a European perspective?

In order to ass the influence of AI regulation under the principle of independence in Europe, one could see that two sorts of frameworks can aim to regulate such influence. The first one would be the general framework of protection of human rights adapted to this paradigm; the second one would be the specialized framework specifically adapted to the undue influence paradigm.

The general framework of protection of human rights and AI offers the ground basis protection that can adapt to the undue influence paradigm. In this approach one has to focus on two different approaches these regimes aim to target: the state and the developers.

In the general framework of human rights, states hold positive obligations in order to "secure" human rights in their jurisdictions.[60] In the matter of new technology, it was held that states hold a "special responsibility" when they introduce new technology that have the potential to harm human rights in their criminal justice system.[61] Therefore, when introducing such a technology one could see that states would have to provide for special guarantees in order to limit the harm to human rights. From this general perspective, comes another regime of human rights obligations for states: The positive obligations. Whereas the traditional perspective of human rights protection focused on a negative approach of protection; meaning that states should not harm their citizens human rights (a vertical protection of human rights from the state to the citizen) the perspective of positive obligations tend to not only provide a vertical protection but also the obligation for states to use the

[60] *Cevrioğlu v. Turkey*, no. 69546, § 50, ECHR 2016; *Assanidze v. Georgia*, no. 71503/01, § 144, ECHR 2004; AKANDJI-KOMBE, J. F. *Positive Obligations under the European Convention on Human Rights: A Guide to the Implementation of the European Convention on Human Rights,* Human Rights Handbook, Volume 7, p. 5.

[61] *Marper v. The United Kingdom*, no. 30562/04 and 30566/04, § 112, ECHR, 2008.

means at disposal in order to secure them in a horizontal perspective (from citizens to citizens).[62] These two sets of obligations aim at providing a general protection of human rights in the jurisdictions of states. One could ask however, what is the material scope of such positive obligations under the principle of independence in Europe. The European courts have yet to statute on such topic as only a negative aspect has been witnessed in the case-law, due to the state interfering in the principle of independence; and even in cases where non-states actor could have had an influence over the criminal trial (press and media), the ECtHR only assessed that: judges being professional have a presumption against such influences; and that since courts have given reasoned decision, it is enough to affirm that judges were not influenced.[63] However, the ECtHR did not undergo any positive obligations for the State. One could theorize that States have to take the guarantees in order not to let any undue influences affect criminal justice by adopting a legal framework capable of protecting judges, through continuous training of judges and satisfactory legal safeguards by criminalizing behavior that could undermine such independence such as in France where such influences are punished under various criminal provisions.[64] But on the specific topic of AI, it is more complex. For a criminal intent on the AI would be difficult to demonstrate in cases where it would exercise an influence on the judge; the question lies whether the developers could be held criminally liable by negligence in cases where data infused in the algorithm lead to a discrimination of the individual concerned. This remains purely hypothetical at the moment, however. Therefore, the challenge of AI as an undue influence would also lie in the context where it could thrive in a legal void where an undue influence could occur without criminal intent.

The specialized framework directly targeting the elimination of undue influence can be deduced from both the principle of independence regime and the regulation regarding algorithm and AI.

Indeed, under these two regimes it seems that AI would be deployed on the first instance of criminal trials. Two arguments can be held to demon-

[62] *Ireland v. United Kingdom*, no. 530/71, § 239, ECHR 1977.

[63] STOYANOVA, V.: *Fault, knowledge and risk within the framework of positive obligations under the European Convention on Human Rights.* In: Leiden Journal of International Law, vol. 33, no. 3, 2020, p. 606, ISSN 1478-9698.

[64] Article 434-9 Code pénal Français; Article 434-9-1 Code Pénal Français ; Cour de cassation, chambre criminelle, 14 mars 1972, 71-91.077; Cour de cassation, Chambre criminelle, 30 juin 2004, 03-85019; Cour de cassation chambre criminelle, 23 janvier 1919, 18-82.833.

strate such point; firstly under the principle of independence, in order to guarantee that judge was not influenced in a undue manner there should be a control of higher courts on lower courts as to assess whether the principle of independence was secured;[65] secondly under the GDPR and the regulation of automated data processing under which algorithms thrive, individuals have the right to appeal the algorithmic decision to a human controller.[66] So, one could see that under these two regimes, it is necessary to have a control over the decision of whether lower court or the algorithm, in order to secure the independence or individuals rights. Hence in the case where algorithms are implemented in the criminal justice system only in first instance, this would allow for higher courts to both control whether the algorithm did influence the judge in an undue manner and could also exercise its scrutiny on the decision.

Moreover, the question lies on the extent of the place of the algorithm in the criminal trial. One could see that there seems to be a common extent between numerous legal systems under which the algorithm cannot be used as the sole basis for the judicial decision. In order to demonstrate such extent, the French, the U.S. and the European framework would be studied.

The U.S. system is relevant for this study as algorithms are incorporated in the criminal justice system for some years and has revealed the flaws of such use. Two cases demonstrate such extent of the use of the algorithm; *Loomis v Wisconsin* demonstrated that the algorithm did not have to be disclosed to the parties of the proceedings as long as other materials supported the judge's decision;[67] on the opposite *Kansas v Walls* demonstrated that when the algorithm is the main or sole piece under which the judge took its decision, it must be disclosed.[68] One could argue that the extent under which the algorithm is used as part of larger materials and support these materials could lead to believe that the judge did not face an undue influence. Whereas when it was the sole material that led to the decision, the disclosing must be effective in order to firstly, allow the defendant to respond to the argument and secondly to make sure whether the algorithm did not unduly influence the Judge, through its functioning.

The European way focuses on such extent too, especially with the current regulation contained in the GDPR in Article 22; under such provision the an

[65] *Gatt v Malta,* no. 46466/16, § 88, ECHR 2019.
[66] Article 22(3) of the GDPR Regulation.
[67] *Loomis v Wisconsin,* [2016], Wisconsin Supreme Court, N° 16-6387.
[68] *Kansas v Walls* [2017], Court of Appeal of the State of Kansas, N°116,027.

individual cannot solely be subjected to an automated decision making that would have legal effects on this individual;[69] the project of regulation of AI by the Parliament and the council also focuses on an AI as an assistant and on the danger of over-reliance of AI as a decision helper.[70] Therefore, this European approach of the AI leads to assess that even though it has a clear positive side on using it for criminal justice, it should not take an overreached place in the decision making. Such approach is also shared in the French system, where expert reports that could influence the sentencing of an individual (psychiatric, risk…) cannot be the sole basis on which the decision is made.[71]

Hence, one can see that the place of the algorithm could be one among other materials, but not a sole decider. This would prevent any "token" approach of the relationship between the algorithm and the judge and limit the extent of the influence in order for the judge to consider the AI as a piece that would provide guidance for the decision, rather than taking over the decision of the judge.

Conclusion

One could see that the artificial intelligence could pose a dire challenge to the principle of independence in criminal trials. The developing, the functioning and the influence it could have on the judge can qualify it as undue if no proper safeguards are taken in order to prevent an "algorithmic justice". Such safeguards must however be weighted considering other relevant features such as intellectual property and the transparency of the algorithm. In order for justice to remain independent, such algorithms must however comply with the principles that have already been erected in the field of independence; but also, on the emerging principles for a trustworthy AI focusing on the necessity of transparency and explainability in order for AI to become a legitimate influence in the field of criminal trials.

[69] Article 22 of the GDPR Regulation.

[70] Recital 40 of the GDPR Regulation.

[71] Cour de cassation, 2ème chambre civile, arrêt du 13 avril 2008, n° 07-16824; Cour de cassation, 2ème chambre civile, arrêt du 8 septembre 2011, n° 19919; Cour de cassation, 2ème chambre civile arrêt du 7 septembre 2017, n°16-15531.

Bibliography

1. *Agrokomplex v. Ukraine,* no. 23465/03, ECHR 2011.
2. AKANDJI-KOMBE, J. F. *Positive Obligations under the European Convention on Human Rights: A Guide to the Implementation of the European Convention on Human Rights,* Human Rights Handbook, Volume 7.
3. *Al-Skeini and Others v. United Kingdom*, no. 55721/07, ECHR 2011.
4. ANGWIN, J. et al.: Machine bias: there's software used across the country to predict future criminals. And it's biased against blacks. Propublica, 23.5.2016. Online: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (quoted 1.12.2021).
5. *Anzelika v Lithuania*, no. 36093/13, ECHR 2020.
6. *Assanidze v. Georgia*, no. 71503/01, ECHR 2004.
7. *Beg S.P.A v Italy*, no. 5312/11, ECHR 2021.
8. BENSOUSSAN, A., BENSOUSSAN, J.: *IA, Robots et droit*. 1re edition, BRUYLANT, 2019 Brulyant, ISBN 9782802763673.
9. BERK, R.: An impact assessment of machine learning risk forecasts on parole board decision and recidivism. Working Paper No. 2016-4.0, University of Pennsylvania, Department of Criminology, 2016. Online: https://crim.sas.upenn.edu/sites/default/files/WP2016-04_Berk_MachineLearningParole_08.03.2016%281%29.pdf (quoted 1.12.2021).
10. *Cevrioğlu v. Turkey*, no. 69546, ECHR 2016.
11. Charter of Fundamental Rights of the European Union, OJ EU C 326/391, 26.10.2012.
12. CJEU, Judgment of 15.5.1986, *Johnston v Chief constable of the Royla Ulster Constabulary,* C-222/84.
13. CJEU, Judgment of 19.11.2019, *A.K v Krajowa Rada Sądownictwa, and CP and DO v Sąd Najwyższy*, C-585/18, C-624/18, C-625/18.
14. CJEU, Judgment of 24.6.2019, *European Commission v Republic of Poland,* C-619/18.
15. CJEU, Judgment of 27.2.2018, *Associação Sindical dos Juízes Portugueses v Tribunal of Contas,* C-64/16.
16. CJEU, Judgment of 4.5.2017, *Banco de Santander SA,* C-274/15.
17. Code pénal Français.
18. Committee of Ministers, Judges, Independence, Efficiency and responsibility (2010).
19. Committee on Legal Affairs and Human Rights. Justice by algorithm – the role of artificial intelligence in policing and criminal justice systems. Provisional version. 2020.
20. Consolidated versions of the Treaty on European Union (TEU) and the Treaty on the functioning of the European Union (TFEU), OJ EU C 326, 26.10.2012.
21. *Constantinides v Greece*, no. 76438/12, ECHR 2016.

22. Consultative Council of European Judges, Magna Carta of Judges (2010).
23. Consultative Council of European Judges, opinion N°18/2015: Special text regarding other branches but still extensive and similar to previous texts.
24. Consultative Council of the European Judges, Opinion N°1.
25. CONTINI, F.: Artificial Intelligence: A New Trojan Horse for Undue Influence on Judiciaries? UNODC. Online: https://www.unodc.org/dohadeclaration/en/news/2019/06/artificial-intelligence_-a-new-trojan-horse-for-undue-influence-on-judiciaries.html (quoted 8.11.2021).
26. Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by protocols Nos. 11 and 14, adopted on 4 November 1950, ETS 5.
27. Cour de cassation, chambre criminelle, 23 janvier 1919, 18-82.833.
28. Cour de cassation, 2ème chambre civile, arrêt du 7 septembre 2017, n°16-15531.
29. Cour de cassation, 2ème chambre civile, arrêt du 13 avril 2008, n° 07-16824.
30. Cour de cassation, 2ème chambre civile, arrêt du 8 septembre 2011, n° 19919.
31. Cour de cassation, chambre criminelle, 14 mars 1972, 71-91.077.
32. Cour de cassation, chambre criminelle, 30 juin 2004, 03-85019.
33. COX, A.: The Independence of the Judiciary: History and Purposes. In: *University of Dayton Law Review,* vol. 21, no. 3, 1996, p. 566, ISSN 0162-9174.
34. CRAWFORD, J.: *State Responsibility.* Cambridge University Press, 2013, ISBN 9781139033060.
35. *Craxi v Italy,* no. 34896/97, ECHR 2002.
36. Data protection working party. Guidelines on Automated individual decision-making and profiling for the purposes of Regulation 2016/679, adopted on 3 October 2017.
37. DUBBER, M. D., PASQUALE, F., DAS, S. (eds.): *The Oxford Handbook of Ethics of AI.* Oxford University Press, 2020, ISBN 9780190067397.
38. EMMERSON, B. et al.: *Human Rights and Criminal Justice.* 3rd Edition, London Sweet and Maxwell, 2012.
39. European Charter on the Statutes for Judges (1998).
40. European Commission for Democracy Through Law: European standards on the independence of the judiciary.
41. European Commission for the Efficiency of Justice. European ethical charter on the use of Artificial Intelligence in judicial systems and their environment, adopted on 3–4 December 2018. Online: https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (quoted 1.12.2021).
42. European Commission. Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. COM (2021) 206 final. 21.4.2021.

43. *G.C.P v Romania*, no. 20899/03, ECHR 2011.
44. GARAPON, A.: Le numérique est un remède à la lenteur de la Justice. DALLOZ. Lefebvre Dallox. Actualité, 4.5.2018. Online: https://www.dalloz-actualite.fr/interview/antoine-garapon-numerique-est-un-remede-lenteur-de-justice#.YYjmCWDMJPY (quoted 8.11.2021).
45. *Gatt v Malta,* no. 46466/16, ECHR 2019.
46. GOLDSZLAGIER, J.: L'effet d'ancrage ou l'apport de la psychologie cognitive à l'étude de la décision judiciaire. In: *Les cahiers de la Justice*, no. 4, 2015, p. 507, ISSN 1958-3702.
47. *Guðmundur Andri Ástráðsson v Iceland*, no. 26374/18, § 219, 2020.
48. High-level expert group on Artificial intelligence set up by the European Commission, Ethic Guidelines for Trustworthy AI, European commission, published on 8 April 2019.
49. HUBIN, J. B., JACQUEMIN, H., MICHAUX, B. (eds.): *Le Juge et l'algorithme : Juges augmentés ou justice diminuée*. LARCIER, 2019, ISBN 9782807911161.
50. *Industrial Financial Consortium Investment Metallurgical Union v Ukraine*, no. 10640/05, ECHR 2018.
51. *Ireland v. United Kingdom*, no. 530/71, ECHR 1977.
52. KAHNEMAN, D., TVERSKY, A.: On the psychology of prediction. In: *Psychological Review*, vol. 80, no. 4, 1973, p. 237, ISSN 0033-295X.
53. *Kansas v Walls* [2017], Court of Appeal of the State of Kansas, N°116,027.
54. KATSUYA ENDO, S.: Technological Opacity & Procedural Injustice. In: *Boston College Law Review,* vol. 59, no. 3, 2018, p. 823, ISSN 0161-6587.
55. *Khrykin v Russia*, no. 33186/08, ECHR 2011.
56. KITCHIN, R.: Thinking critically about researching algorithms. In: *Information, Communication & Society,* vol. 20, no. 1, 2017, p. 14, ISSN 1369-118X.
57. KRENT, H. J.: *The private performing the Public: Delimiting Delegations To Private Parties.* In: University of Miami Law Review, vol. 65, p. 507.
58. LESLIE, D. et al.: Artificial intelligence, human rights, democracy, and the rule of law: a primer. The Council of Europe, 2021, p. 21. Online: https://www.turing.ac.uk/sites/default/files/2021-03/cahai_feasibility_study_primer_final.pdf (quoted 1.12.2021).
59. *Loomis v Wisconsin,* [2016], Wisconsin Supreme Court, N° 16-6387.
60. *Marper v. The United Kingdom*, no. 30562/04 and 30566/04, ECHR, 2008.
61. *Muller v Germany*, no. 54963/08, ECHR 2014.
62. OSWALD, M., GRACE, J, URWIN, S., BARNES, G. C.: *Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality.* In: *Information & Communications Technology Law*, vol. 17, no. 2, 2018, p. 223, ISSN 1469-8404.
63. *Paulikas v Lithuania,* no. 57435/09, ECHR 2017.

64. 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), OJ EU L 119/1, 4.5.2016.

65. RHUE, L.: *Anchored to Bias: How AI-Human scoring can induce and reduce bias due to the anchoring effect*'. University of Maryland, 2019.

66. Rule of Law Checklist, 2016 (Vienna).

67. *Soering v United Kingdom*, no. 14038/88, ECHR 1989.

68. STOYANOVA, V.: Fault, knowledge and risk within the framework of positive obligations under the European Convention on Human Rights. In: *Leiden Journal of International Law*, vol. 33, no. 3, 2020, p. 601, ISSN 1478-9698.

69. TRECHSEL, S.: *Human Rights in Criminal Proceedings*. Oxford University Press, 2005.

70. TVERSKY, A., KAHNEMAN, D.: Judgment under Uncertainty: Heuristics and Biases. In: *Science*, vol. 185, no. 4157, p. 1124, ISSN 1095-9203.

71. TVERSKY, A., KAHNEMAN, D.: Subjective probability: A judgment of representativeness. In: *Cognitive Psychology*, vol. 3, no. 3, 1972, p. 430, ISSN 1095-5623.

72. WESSLEN, R. et al.: Anchored in a Data Storm: How Anchoring Bias can affect user strategy, confidence, and decisions in visual analytics. University of North Carolina, 2018, p. 2. Online: https://researchain.net/archives/pdf/Anchored-In-A-Data-Storm-How-Anchoring-Bias-Can-Affect-User-Strategy-Confidence-And-Decisions-In-Visual-Analytics-3214328 (quoted 1.12.2021).

73. WILDHABER, L.: The European Court of Human Rights in Action. In: *Ritsumeikan Law Review,* no. 21, 2004, p. 83, ISSN 2434-2424.

74. YEUNG, K.: Algorithmic regulation: A critical interrogation. In: *Regulation & Governance*, vol. 12, no. 4, 2018, p. 505, ISSN 1748-5991.

75. ZAVRŠNIK, A.: Criminal Justice, Artificial intelligence systems, and human rights. In: *ERA Forum,* vol. 20, 2020, p. 567, ISSN 1863-9038.

Author details

Théo ANTUNES
PhD student
University of Luxembourg and University of Strasbourg
Faculty of Law
theoantunes97@gmail.com

5 ROBOT AS A JUDGE? POSSIBILITY OF THE USE OF AN ARTIFICIAL INTELLIGENCE IN COURT PROCEEDINGS IN SLOVAKIA

Silvia Brnčalová

Abstract

Robotization is nowadays present in many areas. Robots and AI technologies simplify our day-to-day activities. Rumor has it that some professions might become obsolete in the future due the robotization. Is it a case in judiciary as well? Will artificial intelligence decide on our right and duties? These are the questions, that the author is trying to answer in this article, while highlighting possible advantages and disadvantages connected to the possible use of the artificial intelligence throughout the decision-making process in courts in Slovakia.

Introduction

We are on the brink of the new industrial revolution. This one is faster and more dynamic than its predecessors.[1] This industrial revolution doesn't bring only the questions and challenges regarding to the sector of technologies. Many legal questions are arising as well.[2] Questions, which are posing as challenges, which were not imaginable for the lawyers in the past few centuries. Robots with artificial intelligence were in the past considered just a science fiction, however today they are more or less part of our everyday life. Just look at the robotic vacuum cleaner. It is able to function without any human intervention and thanks to the built-in AI it can detect the obstacles on the floor, and instead of senselessly hitting them – it bypasses them. But nowadays there are many more sophisticated robots than automatic vacuum cleaners, mostly in the sector of medicine, cosmonautics etc.

These days, robots replacing humans e.g., in the industrial production, are common. Tasks done by ten people in the past are done efficiently by one robot. It is estimated that in a few years or decades some human professions

[1] See also MAŘÍK, V.: *Průmysl 4.0 – výzva pro Českou republiku*, Praha: Management press, 2016, ISBN 978-80-7261-440-0.

[2] SMEJKAL, V.: *Kybernetická kriminalita*, 2. rozšírené a aktualizované vydanie. Praha: Aleš Čeněk, 2018, p. 817, ISBN 97-8807-380-720-7.

might become obsolete due to the advancements in the robotics. Should lawyers be concerned about losing their job?

The aim of this article is not to spread panic and to create dystopic scenarios, but to present some ideas regarding to the possible future of the judiciary (in Slovakia) in the times, where AI is slowly becoming present in every part of our society. These ideas are compared to the actual situation at courts. The author currently works as a higher judiciary officer, thus some thoughts presented above were inspired by her legal practice. Our approach was mainly inspired by the legal framework of the European Union and the Council of Europe regarding to the possible use of an AI at the decisioning process at courts in Slovak republic. Scientific methods used here are mainly analysis, synthesis, and deduction.

5.1 Decision-making process on courts nowadays (in Slovakia)

Regarding the Constitution of the Slovak republic the judiciary shall be administered by independent and impartial courts. Justice at all levels is administered independently of other state bodies. Courts are deciding mainly on rights and duties of subjects in civil law and criminal law matters. If laid down by the law, they also examine the lawfulness of public administration bodies decisions and lawfulness of decisions, measures, or other acts of the public authority bodies.[3]

The main person which is connected to every court is a judge. Judge[4] is a person with the law degree, nominated by the president of the Slovak republic after completing all the requirements laid by the law (mostly after successfully completing the selection procedure). Judge is the authority in the field of justice, who has to decide impartially, fairly, lawfully, according to the facts of the case and without undue delays in the decision-making process. Judges are the main actors on hearings, they are also giving judgements and resolutions, but they aren't the only actors with a ruling power within the judiciary. These other subjects are partially mentioned in the art. 142 sec. 2. of the constitution, which states that some matters may be decided also by

[3] Article 141 and 142 of the Constitution of the Slovak Republic, published under No. 460/1992 Coll.

[4] Framework for the role of a judge is set out in the act no. 385/2003 Coll. on judges and accessors.

a court's employee authorized by a judge. Such employees are judicial offices, mainly higher judiciary officers.

The position of the higher judiciary officer is governed by a specific legal act.[5] This function may be held only by a person with a law degree and without being a convicted criminal for committing an intentional criminal offence. Higher judicial officer is serving mainly the field of the civil law, criminal law, administrative law and insolvency law. In some cases, they may operate only on the basis of authorization given by a judge. In the civil law matter, the higher judicial officers mostly decide on the costs of the proceedings, on court fees, on removing errors in the claims, on jurisdiction of the court etc., but also carry out requests and make concepts of decisions for the judge. In the criminal law matters judicial officers decide also about providing a defending attorney, about setting of the period of the detention and the sentence, but also make concepts of decision for the judge etc. Specific acts of law are providing other tasks for higher judiciary officers, such as making decisions about objections regarding to the discontinuation of the enforcement procedure.[6] Exhausting sum of tasks of the higher judiciary officer is not our goal. But in general, we may say, that higher judiciary officer is providing mostly supportive tasks, which are inevitable for the final decision in cases held by a court.

5.2 Artificial intelligence and the judiciary

5.2.1 AI in the legal practice

Artificial intelligence (abbreviation "AI") has become part of the legal practice of some companies and individuals. In the past years, various software running within the AI, which is used by the lawyers, have been created.[7] The modern software is able to do various operations e.g., create a draft of the contract. Such software is LISA, which is considered the first legal software with the AI. LISA is able for example to draft a non-disclosure agreement within fifteen minutes.[8] Some software is also used for checking contracts before concluding these contracts. Example in this way can be a software cre-

[5] Act no. 549/2003 Coll. on judiciary officers.

[6] See § 10 para. 2 of the Act no. 233/2019 Coll. on the discontinuation of the old enforcement procedures.

[7] FUNTA, R.: Zmení umelá inteligencia budúcnosť právnického povolania? In: *Justičná revue*, vol. 71, n. 6–7, 2019, p. 620, ISSN 1335-6461.

[8] Robot Lawyer LISA. Online: https://www.f6s.com/robotlawyerlisa (quoted 31.10.2021).

ated by the company Lawgeex, which is able to check contracts much faster, than by a human-lawyer.[9]

It is worth mentioning a Technology assisted review (TAR) too. TAR is based on computer learning and is used to the database searching relevant electronic legal information (so called e-Discovery). This data shall be used in the legal practice for resolving a real case. TAR can facilitate practice of lawyers, because without it lawyers would have to search through gigabytes of data (e.g., searching in the past judgements).[10] In 2012, the use of TAR in e-discovery was for the first time upheld by the court (in the USA).[11] For searching through abundance of the legal documents software ROSS created by IBM is used as well, which can look for relevant information among legal acts and case-law.

For research of legal documents there are also software products running within AI by company LexisNexis, used by many law offices.[12] According to the survey done by the International Legal Technology Association, majority of the large law offices (with over 700 employees) uses some kind of AI tools in their practice.[13]

In the USA, AI is being applied in the court. Although it doesn't stand behind the judicial decisions itself, it serves as a supportive element in the process. Software called COMPASS is used to assess the risk, whether the convicted person is likely to commit another crime in the future. Basis for such assessment is information gained on a basis of an interview with the convict and information concerning the crimes committed by this person in the past. This prediction of the possible recidivism was a base element in the Erik L. Loomis v. Wisconsin case. In this case, Mr. Loomis was given a maximum sentence by the judge, because the software COMPAS has flagged him as an individual with high possibility of recidivism of committing a crime in the future. Mr. Loomis has objected, that his right to a fair trial in this proceed-

[9] Lawgeex. Online: https://www.lawgeex.com/ (quoted 31.10.2021).

[10] What Is Technology Assisted Review (TAR)? Online: https://blog.pagefreezer.com/what-is-technology-assisted-review-tar (quoted 31.10.2021).

[11] First judicial decision approving use of predictive coding (TAR) is Da Silva Moore v. Publicis Groupe, 287 F.R.D. 182 (S.D.N.Y. 2012) ruled by the U. S. Magistrate Judge for the Southern District of New York Andrew J. Peck. Online: https://www.fedarb.com/professionals/judge-andrew-j-peck/ (quoted 31.10.2021).

[12] LexisNexis. Online: https://www.lexisnexis.co.uk/ (quoted 31.10.2021).

[13] VANNI, R.: How Artificial Intelligence Is Transforming the Legal Profession. Kira, 08.05.2020. Online: https://kirasystems.com/learn/how-artificial-intelligence-is-transforming-the-legal-profession/ (quoted 31.10.2021).

ing has been violated, because the court decision was based on the results of the aforementioned analyses. However, the Supreme court in Wisconsin has overruled his objection, because Mr. Loomis was given an opportunity to comment the results of analyses of his person but did not perform it in due time.[14]

It is necessary to say, that these software products are finding their use in legal practice predominantly in the western countries such as the USA, but in Slovakia they are rarely found in use.

5.3 Framework for the AI in the judiciary

Currently, the use of AI in decision making process in the courts in Slovakia's legal system is not regulated by any law. However, it is not prohibited as well. But the potential use of an AI in this area is a hot topic not only in the Slovak republic, but in other countries as well.[15] There's also no binding framework regarding the AI within the European Union, but same steps for the possible legal framework of the AI have been taken.

In July 2018 European commission appointed a group named High-level expert group on artificial intelligence (AI HLEG) as an independent expert group on the use of AI. This group has created a document called Ethics guidelines for trustworthy AI. These guidelines were meant to provide a guidance to AI practitioners.[16] The main objective of the guidelines is a human-centric approach on AI to achieve trustworthiness. It should meet 3 main conditions of which is one legality, in a sense that trustworthy artificial intelligence (TAI) should be a subject to legislation. Others two are ethics, which should guarantee the accordance between various ethical principles and robustness on technological and social sides. These three components – legality, ethics and robustness should work together regarding TAI.[17] Ethical guidelines for TAI in its part C called "Examples of opportunities and critical concerns

[14] FUNTA, R.: *Zmení umelá inteligencia budúcnosť právnického povolania?* op. quoted, p. 620–621.

[15] See PSHAVA, V., BABENKO, I., PLETNEV, V., SOKOLOV, A.: The possibility of using artificial inteligence in legal proceedings of contemporary Russia. In: *Revista inclusiones*, vol. 7, 2020, p. 396–405, ISSN 0719-4706.

[16] SMUHA, N. A.: The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence. In: *Computer Law Review International,* vol. 20, no. 4, 2019, p. 3, ISSN: 2194-4164.

[17] Ethical guidelines for the trustworthy artificial intelligence. Online: https://op.europa.eu/sk/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-sk/format-PDFS. 6 (quoted 31.10.2021).

raised by AI" offer non-exhaustive list of the areas, where the use of an AI might be highly valuable (e.g., education, sustainable infrastructure etc.), as well as list of areas, where the use of an AI might have a negative impact (e.g., tracking individuals, lethal autonomous weapon systems etc.). Sector of judiciary is not included in any of those lists, but that's understandable since those are only non-exhaustive examples of possible benefits and risks regarding to the use of an AI.

The proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (2021/0106(COD)) stipulates which AI systems should be deemed as a high-risk AI systems. Such systems are also AI systems intended to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts.[18] High-risk AI systems shall comply with the requirements established in the Chapter 2 of the Proposal. Among other requirements, there's "human oversight", which stands for the possibility to oversee the AI system while it is in use.

In European parliament, the Committee on Civil Liberties, Justice and Home Affairs have released the Report on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)) on 13 July 2021. Committee highlights importance of possible human control over the output created or co-created by the AI.[19]

In 2018, the Council of Europe adopted the European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment.[20] This charter represents a framework of basic principles, which due to the dynamic advancement of AI, should be accepted by the lawmakers and persons working in the field of the judiciary. While using the AI in the courts, these five principles should be followed:

- Principle of respect for fundamental rights – using the AI during the deciding process should be in accordance with fundamental rights.

[18] Art. 6 and Annex III of the The proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (2021/0106(COD)).

[19] Report on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)) from 13.07.2021. Online: https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_EN.html (quoted 31.10.2021).

[20] European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment. Online: https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (quoted 31.10.2021).

- Principle of non-discrimination – data processing done by the AI shall not be discriminatory for any individuals or groups of individuals.
- Principle of transparency, impartiality and fairness – basis for the machine learning should be taken from relevant justice system professionals (judges, prosecutors, lawyers, etc.) and researchers/lecturers in the fields of law and social sciences. Such software shall be stored and executed in secure environments to ensure its system integrity and intangibility.
- Principle of transparency, impartiality and fairness: data processing should be controllable by authorized external audits.
- Principle "under user control" – professionals in the field of justice should be at any moment able to review the judicial decision made by an AI, or a data used by the AI, while such decision made by an AI shall not be binding at any case for the professional in the field of justice. Users must be informed, whether in their case the AI was used. They may also have a right to object such use of an AI in their proceedings which leads to the case being directly handled by a court within the meaning of art. 6 of the European convention on human rights (right to a fair trial).

Scientists are also active in giving their opinions regarding to the possible laws applicable for the AI. In 2017 a conference called *Future of Life Institute 2017 Asilomar Conference* was visited by experts in programing, lawyers, philosophers, engineers etc. concerned with the Artificial Intelligence. The output of the conference was 23 principles for the AI (AI Asilomar Principles).[21] The scientist Stephen Hawking, or the founder of the SpaceX company Elon Musk were among the people who agreed with the 23 principles presented by this conference. The principles are divided into 3 segments – research principles, principles concerning ethics and values and long-term principles. One of the ethics and value principles is applicable to justice system – decision making process with the help of AI. This principle number 8 is saying that it is required to legitimately justify any use of AI in the decision-making process in judiciary, while these decisions should be controlled by a person.

Integrating the AI and helping with decisions is not excluded by authors of the principles, however they find it necessary for decisions about the rights

[21] Future of Life Institute 2017 Asilomar Conference. Online: https://ai-ethics.com/2017/08/11/future-of-life-institute-2017-asilomar-conference/ (quoted 31.10.2021).

and obligations of companies and individuals wasn't in a full competence of programed artificial intelligence, without the review by a human factor.[22]

5.4 Possible use of the artificial intelligence during the decision-making process in courts (in Slovakia)

The idea of a robot with an AI, which is present in a court room in the role of a judge deciding about the guilt and a sentence of a convicted person, seems a little bit far fetched, and for now has a place only in a science fiction novels. It is true, that this idea can seem somewhat dystopic, as lives of individuals would end up in the metaphorical hands of some robot respectively an AI. In real life scenario, it wouldn't have to be an android (humanoid) type of a robot, which performs all judicial tasks. Let's take for example such a robot, which is used in medicine in surgeries and this robot doesn't lead the operation by itself without any human personnel, however it allows for a less invasive operation under the supervision by human surgeon.[23]

According to the approach of the European union and the Council of Europe to the use of an AI in the field of justice, it seems that the AI might be used in court in the future. Sooner or later, it is inevitable, that an AI will be used in the field of a justice.

At first, we must understand, that proceedings at court are not happening only in the public hearings. Important part of a judicial decision-making process happens outside of a courtroom, e.g., deciding about complaints against the decision issued by a higher judiciary officer, deciding about urgent measures etc. The main role of an AI could be supporting and helping judicial officers and judges during their decision-making process.

Before issuing any kind of a decision, it is essential to get to know the factual circumstances of any case. Making the final decision is often not simple, and that's why it is necessary to look for a relevant court law. Browsing through a relevant Slovak and European (or international) court law might take a considerable amount of time. If an AI might be able to identify such a relevant court law, it would make the work of judges and judicial officers faster and easier. They could in return use the saved time on more compli-

[22] Asilomar's 23 principles: Researchers attempt to establish guidelines around Artificial Intelligence safety, 2017. Online: https://analyticsindiamag.com/asilomars-23-principles-researchers-attempt-guidelines-around-artificial-intelligence-safety/ (quoted 31.10.2021).

[23] See, e.g., the medical robot Da Vinci. Online: https://www.davincisurgery.com/ (quoted 31.10.2021).

cated cases, where the case analysis and resolving it is much more time consuming.

For the software to be able to identify and browse in the relevant court law, it must have access to such court law. This access has to be permanent with the regular updating of this database of court law, including the court law of the court of Justice of the European Union and the European Court of Human Rights. But creation of such database is not enough by itself. For the real ease of the deciding process, it is needed that such AI will be able to evaluate the case under given criteria for it to be able to find proper court law for a case.

A judge or a higher judiciary officer should then deal with the case law found by an AI. Next they should evaluate whether such case law is truly applicable for their case or not. This way, the human control over the outputs made by an AI would be ensured.

More controversial seems to be delegating a deciding power from a judge (or a higher judiciary officer) to on AI. In some cases, some part of a decision-making process could be handled by an AI. What we have in mind are so called "form applications and decisions". Generally, every case should be decided on its specific features and grounds. In some cases, applicants are filing almost the same applications in various similar (almost the same) cases, in which they demand same output from the judge. Current example might be filling the objections based on the act no. 233/2019 Coll., where applicants are objecting the same circumstances in similar cases. In such cases, court has no other choice than to issue practically the same decisions, where the only difference is the name of the defendant. If such decisions were issued by an AI, court officials might have more time to focus on individual cases, which do not belong to this "form applications and decisions". But human control should be present hereby as well. Such decisions created by an AI should be accepted and might be (if needed) also reviewed by a judge (or higher judiciary officer).

Another field where the AI might find its use, is creating analysis on individual cases with potential solution on the case, which would be only a supportive measure for the judge. It wouldn't be binging for him in any way in his decision-making process. Resort for such on outcome could be mainly existing decision from the courts in similar cases, but also legislative acts or published opinions and articles of legal researchers and scholars. The goal of an AI would be to analyze these sources and process, whether some of them might be applicable in an actual case.

We don't see the aim of the use of an AI in the decision-making process in the field of justice in the replacing human judges and other judiciary officers. But their work might be simplified by the use of an AI so they might have more time for complicated and more complex cases. Overall, the decision-making process could be more effective and precise, because judge might profoundly inspect the case in the saved time, instead of spending hours on the research on existing court law including international and European court law. If such analysis was done by a software with an AI, it would take the software a fraction of a time to select relevant documents and to propose them to the judge. Judge would not have to lose his time on studying irrelevant to the case and time-consuming court law, which is not applicable on his actual case. Finding that the court law is irrelevant is usually discovered after reading some part of it and thus spending time on it.

5.5 Advantages and disadvantages from the potential use of an AI in the judiciary

5.5.1 Financing

It is often spoken, that money is not everything, but when it comes to the development of the newest technologies, financing is its crucial element. Without sufficient financial resources it is impossible to expect, that a sophisticated software might be developed. And this doesn't stand only for the software in the field of justice, but in any other field too.

On one hand, the market with legal technologies is relatively small, which is a negative factor. On the other hand, there's e.g., market with the medical technologies, where there is a bigger competition but also bigger demand and thus a chance for bigger profit.[24] But if we focus on possible demand for a software aimed solely for the field of justice, the market is getting much smaller. Potential "customer" here is a state, represented by a Ministry of Justice of Slovak republic, because judiciary is a part of a state administration. Relying on a financing aimed for the development of and AI useful for the deciding process on court, coming solely from a state budget seems quite optimistic.

Possible solution might be offering this software to law offices, universities, and other corporations, where such software might find its use, e.g.,

[24] FUNTA, R.: *Zmení umelá inteligencia budúcnosť právnického povolania?* Op. quoted, p. 617–626.

faster research in legal documents and in a case law etc. That's how such software might be placed on a market and not relying only on a potential interest and financing from the state. But on the other hand, a software that will be publicly offered might be only one with the AI for research and not for the creation of the "form decision" or any other kind of decisions. According to us, such software (able to create a judicial decisions) shall be solely in the use of courts.

5.5.2 Citizen's trust

In the aforementioned text, we presented an opinion, that the AI should have only a supportive role in the decision-making process, while total replacement of a human factor is not welcomed. Fact is, that (in Slovakia) there's very low trust of our citizens in a system of justice (only 30% citizens have trust in our courts), and that's why Slovakia stands at the end at the ranking about the trust in the courts among member states of the European union.[25]

It isn't clear whether usage of the AI might have the same impact on our citizens, or a total opposite – that the trust in the courts might increase. Essential factor here is without any doubt a level of usage of the AI in the decision-making process, i.e., whether AI issues the concepts of decisions, or prepares research for the judge or does something else. The possibility that citizens might have a bigger trust in the AI in the court rather than into humans in the courts cannot be denied or confirmed for now. With the AI, there might be no suspect of bias of the AI, but since the presence of a human control is inevitable, this means that this person should be able to prove to have no bias in a case. On the other hand, another group of citizens might miss the contact with humans or might negatively feel that artificially created software is deciding about their rights and duties. But these are just presumptions for now because there is not enough data regarding to the opinion of the Slovak public to the (possible) use of the AI on courts. But the answer for this will be known when the AI will be used in the decision-making process in Slovak courts.

[25] Standard Eurobarometer 92 – Autumn 2019, European parliament. Online: https://europa.eu/eurobarometer/surveys/detail/2255 (quoted 31.10.2021).

5.5.3 Fastening of the deciding process and simplifying the tasks for humans

Some judicial files consist of hundreds or thousands of pages. A judge or a judiciary officer, who works with such a file, has to get to known it as a whole. Studying such file takes some time. Following research of legal acts, including the secondary law of the EU and a relevant case law of national and international courts, takes another time. This is the time, which the judge (or a judicial officer) doesn't spend on the deciding process itself, because prior to making decision, studying plenty of legal documents that might be relevant to the case is inevitable. During this research, he has to select between the relevant and irrelevant documents, but for "marking" the document as an "irrelevant" at least partial study of the document is needed. If there was an AI used for such research, it could save the judge hours of his time, which could be spend on other cases. In majority of cases, he would have to study only the documents, that were marked as "relevant" by the AI. Thus, putting AI into the process of research might cause faster and more effective judicial decisioning process.

Using the AI for the "form decisions", where almost the same decision in almost the same cases is needed, might fasten the deciding process too. AI might mark the application as "suitable" for the form decision. Then the draft of such decision might be prepared by AI itself, while adding relevant information to the form decision, such as name of the defendant or the number of the contract etc. Then the judge will sign the decision after checking (if needed) whether the decision is truly correct, so there would be none deciding by the AI which is impossible to supervise.

Conclusion

Development in the field of technologies, mostly AI, is inevitable. Ignoring it is not a solution. AI will sooner or later become (or has already become) part of our daily lives. We should learn how to use it for our benefit and not see it as a danger, but as an opportunity. The same stands for the field of justice as well.

AI has a potential to create a basis for more effective judicial decision-making process on our courts. But that doesn't mean, that the whole decision-making process should depend on the AI. Nowadays, it might be possible to use tools running within the AI for an effective research of relevant

case law and other legal documents, including the secondary EU law and decision of international courts. Such usage of the AI into the research process has a potential to make the work of judges (and judiciary officers) more effective, so they could spend their time only on relevant sources found by the AI and the time they saved when they didn't need to do the research themselves, they could use on another cases. It is highly presumable, that using the AI at the courts might cause that the proceedings would be handled faster and effectively. And fast (and effective) decision-making at courts is a dream of every applicant, isn't it? However, as stipulated in the aforementioned documents of the EU and the Council of Europe, it is necessary to have a human control over the outputs created by the AI in the decision-making process on courts. Totally autonomous AI deciding on rights and duties of humans stays for now only in the science fiction literature.

Bibliography

1. Act no. 233/2019 Coll. on the discontinuation of the old enforcement procedures.
2. Act no. 385/2003 Coll. on judges and accessors,
3. Act no. 549/2003 Coll. on judiciary officers,
4. Asilomar's 23 principles: Researchers attempt to establish guidelines around Artificial Intelligence safety, 2017. Online: https://analyticsindiamag.com/asilomars-23-principles-researchers-attempt-guidelines-around-artificial-intelligence-safety/ (quoted 31.10.2021).
5. Constitution of the Slovak republic, published under no. 460/1992 Coll.
6. Ethical guidelines for the trustworthy artificial intelligence. Online: https://op.europa.eu/sk/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-sk/format-PDFS. 6 (quoted 31.10.2021).
7. European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment. Online: https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (quoted 31.10.2021).
8. FedArb, Judge Andrew J. Peck. Online: https://www.fedarb.com/professionals/judge-andrew-j-peck/ (quoted 31.10.2021).
9. FUNTA, R.: Zmení umelá inteligencia budúcnosť právnického povolania? In: *Justičná revue*, vol. 71, no. 6–7, 2019, p. 617–626, ISSN 1335-6461.
10. Future of Life Institute 2017 Asilomar Conference. Online: https://ai-ethics.com/2017/08/11/future-of-life-institute-2017-asilomar-conference/ (quoted 2021-10-31).
11. Lawgeex. Online: https://www.lawgeex.com/ (quoted 31.10.2021).
12. LexisNexis. Online: https://www.lexisnexis.co.uk/ (quoted 31.10.2021).

13. MAŘÍK, V.: *Průmysl 4.0 – výzva pro Českou republiku*. Praha: Management press, 2016, ISBN: 978-80-7261-440-0.
14. PSHAVA, V., BABENKO, I., PLETNEV, V., SOKOLOV, A.: The possibility of using artificial intelligence in legal proceedings of contemporary Russia. In: *Revista inclusiones*, vol. 7, 2020, p. 396–405. ISSN 0719-4706.
15. Report on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)) from 13.07.2021. Online: https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_EN.html (quoted 31.10.2021).
16. Robot Lawyer LISA. Online: https://www.f6s.com/robotlawyerlisa (quoted 31.10. 2021).
17. SMEJKAL, V.: *Kybernetická kriminalita*. 2. rozšírené a aktualizované vydanie. Praha: Aleš Čeněk, 2018, ISBN: 97-8807-380-720-7.
18. SMUHA, N. A.: The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence. In: Computer Law Review International, vol. 20, no. 4, 2019, p. 97–106, ISSN: 2194-4164.
19. Standard Eurobarometer 92 – Autumn 2019, European parliament. Online: https://europa.eu/eurobarometer/surveys/detail/2255 (quoted 2021-10-31).
20. The medical robot Da Vinci. Online: https://www.davincisurgery.com/ (quoted 31.10.2021).
21. The proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (2021/0106(COD)).
22. VANNI, R.: How Artificial Intelligence Is Transforming the Legal Profession. Kira, 08.05.2020. Online: https://kirasystems.com/learn/how-artificial-intelligence-is-transforming-the-legal-profession/ (quoted 31.10.2021).
23. What Is Technology Assisted Review (TAR)?. Online: https://blog.pagefreezer.com/what-is-technology-assisted-review-tar (quoted 31.10.2021).

Author details

Mgr. Silvia BRNČALOVÁ
Pan-European University
Faculty of Law
Institute of Public Law
silvia.brncalova@paneurouni.com

6 URGENT JUDICIAL PROTECTION AGAINST DECISIONS OF ROBOTS

Martin Serfőző

Abstract

There are, in fact, no difficulties when analyzing artificial intelligence as a subject matter of legal regulation. Generally, it is just a computer program as a subject matter of intellectual property. But, in comparison with other "things", human society allows artificial intelligence to autonomously interfere with individual's rights increasingly. One of methods, by which legal system reacts to this phenomenon is regulation of protection against decisions based solely on automated processing of personal data under article 22 of General data protection regulation. Although this regulation is, primarily, setting up limitations, it is also a legal basis that tolerates artificial intelligence in a process of making decisions with considerable influence on individual's legal position. There is no reason to deny protection of individual's rights, which are guaranteed by General data protection regulation, from impact of artificial intelligence by a classical judicial institutes including necessary measures taken by court in a civil proceeding under procedural law of Slovak republic.

Introduction

Artificial intelligence, as a topic, is receiving more and more attention of human society. Artificial intelligence is a common part of life for majority of people, and it becomes impossible (even for opponents of modern technology and conspirators) to avoid its influence. Common things such as assessing of a loan application in bank, increasing and decreasing the credit limits on a credit card, paying at a self-service cash register in a hypermarket, employee's attendance records, permission to exit from a paid car parking place, targeting of advertisement and many of other banalities can be managed by artificial intelligence autonomously without people even realizing it. This is a natural consequence of technological progress and any efforts to eliminate or suppress artificial intelligence are doomed to fail, not only because of their regressive nature, but especially due to absence of relevant factual and logical arguments.

Of course, people can misuse artificial intelligence to attack an individual's legal position, just like they can misuse any other thing. Nor can we ignore possible errors in the code of artificial intelligence, which can have a serious impact on individuals. However, the fear of the unknown is not a suitable ba-

sis for thinking about the further direction of artificial intelligence. The role of legal systems is to create mechanisms that, on the one hand, will make it possible to reap the benefits of artificial intelligence on an exceptionally large scale and, on the other hand, will provide protection for individuals and society against the misuse or imperfection of artificial intelligence.

The aim of this paper is to outline an analysis of artificial intelligence as a subject of legal regulation, as well as to analyze legal limits of using artificial intelligence in the context of personal data processing and subsequently to analyze the possibilities of urgent judicial intervention under civil procedural law of Slovak republic, if there is a threat or negative effects of using artificial intelligence when personal data processing.

The basic research method is the analysis of legal norms and opinions of experts in the field of information and communication technology law and personal data protection.

6.1 Artificial intelligence

Anyone can consider artificial intelligence as one of the key elements of modern society and its further development. In principle, there is no area of life in human society where artificial intelligence is not present or where there is no potential for its use. It is also a phenomenon that evokes a full range of reactions, from fear through resistance to enthusiasm and fascination. Precisely, because of its ubiquity and diversity of use, in the first place I consider it necessary to define the very essence of this phenomenon. For the purposes of the analysis that forms the subject of this paper, I perceive two important levels: a general understanding of artificial intelligence and its legal regulation.

6.1.1 General meaning of artificial intelligence

There is no doubt that the study of artificial intelligence falls primarily within the remit of the natural sciences. Even the Short dictionary of the Slovak language defines artificial intelligence as "a department of computer science dealing with the modeling of human perception and thinking in solving complex tasks".[1] However, the general perception of artificial intelligence is

[1] Slovník súčasného slovenského jazyka A – G, H – L, M – N, 2006, 2011, 2015. Online: https://slovnik.juls.savba.sk/?w=inteligencia&s=exact&c=Rc6b&cs=&d=kssj4&d=psp&d=ogs&d=sssj&d=orter&d=scs&d=sss&d=peciar&d=ssn&d=hssj&d=bernolak&d=noundb&d=orient&d=locutio&d=obce&d=priezviska&d=un&d=pskcs&d=psken# (quoted 05.10.2021).

based more on the accent of the feature of intelligence (i.e., not as a department of science), which can be defined as "the ability to understand, think independently and solve problems or situations by intellectual processes, intellectual talent, understanding",[2] which is added to an inanimate, artificial object, i.e. one that "was created by human intervention, not in a natural way; manufactured".[3] It is exactly this (second mentioned) understanding of artificial intelligence evoking diverse reactions of laypeople, but also of scientists and experts. Even in the field of legal theory and practice, such a perception of artificial intelligence is "mycelium" for a number of questions, doubts, but also for challenges in the context of classical legal institutes. This is probably because legally protected values must now resist the attacks not only of the classic enemy – man, but (as something new) also man-made intelligence, i.e., robots with human-like properties. And, of course, it is hard for lawyers to react to something, that is new and unknown for them (in its essence). Lawyers discover the nature of robots subsequently, usually long after robots integrate into common social processes.

In other words, the regulation of human behavior is relatively easy, because man has been a natural addressee of legal norms since the beginning of their creation, while the regulation of artificial intelligence requires a completely new, hitherto unknown approach. As M. Husovec and M. Mesarčík note, "lawyers are rarely the authors of technological change. The development of the information society is a fact for them";[4] this opinion captures the essence of the reaction of the legal system to artificial intelligence. If legal norms were to create artificial intelligence, these norms would undoubtedly take care of incorporating artificial intelligence into existing legal institutes, or new ones (institutes) would be actively created so that artificial intelligence could be effectively regulated. On the other hand, the role of lawyers is not to create social changes, but to recognize the changes and then respond to them in a prompt and appropriate manner. This is also the case of penetration of artificial intelligence into human life.

This contribution is therefore based on the understanding of artificial intelligence as a thing created by human, manifested by properties analogous

[2] *Ibidem*.

[3] Synonymický slovník slovenčiny, 2004. Online: https://slovnik.juls.savba.sk/?w=umel%C3%BD&s=exact&c=Mcab&cs=&d=kssj4&d=psp&d=ogs&d=sssj&d=orter&d=scs&d=sss&d=peciar&d=ssn&d=hssj&d=bernolak&d=noundb&d=orient&d=locutio&d=obce&d=prieziska&d=un&d=pskcs&d=psken# (quoted 05.10.2021).

[4] HUSOVEC, M., MESARČÍK, M., ANDRAŠKO, J.: *Právo informačných a komunikačných technológií 1*, Bratislava: TINCT, 2020, p. 14, ISBN 978-80-973837-0-1.

to human intelligence. At the same time, in this contribution, the terms artificial intelligence and robot will be used as synonyms.

6.1.2 Artificial intelligence as a subject of legal regulation

No matter how sophisticated a robot can be, it is still created by human. The quality of the robot's ability to learn new things and use new knowledge in decision-making is directly proportional to the quality of human contribution to the robot's creation. In other words, the better a person writes an algorithm, the better a robot will work, even when creating algorithms by the robot itself or when correcting existing algorithms by the robot. This aspect makes the robot just a thing, an intangible asset; *de lege lata* robot is computer program. Opinions of professionals confirm this. According to the Handbook on european data protection law developed by the EU Agency for fundamental rights together with the Council of Europe (together with the European court of human rights office) and the European data protection supervisor, "artificial intelligence (AI) means the intelligence of machines acting as 'intelligent agents'."[5] S. Demčák says that "it is basically a work covered by copyright protection".[6]

However, the robot also differs from the classic computer program, as its use can significantly (but not completely) replace a human, because based on the same input robot can create unique decisions. We can use an e-shop order as an example. Steps of ordinary computer program would be the same for each order of the same goods. However, the robot can consider facts that are not related to the order and based on these facts the robot can make a decision that will affect the position of the customer. E.g., if I order shoes, according to the weather forecast on the day of delivery of the shoes (it seems it will rain), the robot will put umbrella to a package as a gift for me (robot makes marketing decision that is benefit for customer). Contrary, if I order shoes and there have been many returned orders in my city over the last three months, the robot will not provide me with free shipping, although it has always provided it to me (robot makes decision based on behavioral prediction. Robot wants to be cost efficient. But for the customer it is disadvantageous). We would find countless examples.

[5] Handbook on European data protection law, 2018 edition, p. 351. Online: https://fra.europa.eu/sites/default/files/fra_uploads/fra-coe-edps-2018-handbook-data-protection_en.pdf (quoted 06.10.2021).

[6] DEMČÁK, S.: Právne postavenie umelej inteligencie. In: *Bulletin slovenskej advokácie,* no. 11, 2020, ISSN 1335-1079.

The approach to robots as objects of legal relations and not subjects *de lege ferenda* will probably remain unchanged. We can see e.g., Art. III par. 1 of the proposal for a regulation of the European parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial intelligence act) and amending certain union legislative acts, which defines a robot as "software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with".[7]

Finding the legal basis of artificial intelligence is not self-serving. Based on it, it is possible to consider the application of specific legal norms to legal relationships in which the robot interacts. In this context, at least two opinions on the legal regulation of artificial intelligence appear in professional literature. First opinion is based on the need for regulation the IT industry (and thus also artificial intelligence) by specific standards or the need for a separate legal branch, the second opinion is based on the application of existing legal institutes also in the field of IT.[8] I am convinced that every robot is *de facto* made by human and *de jure* a computer program. I therefore take the stand, that there is no need to create a separate legal branch for the regulation of artificial intelligence (of course, importance and contribution of ICT law as a scientific and pedagogical discipline is unquestionable). As well as e.g., use of the features of weapons, the use of robots already has its legally set limits, which represent barriers not for things (robots), but for their users. If we were thinking about robot, that can get the status of a unique being by learning and it would cause harm to the rights of human, I am convinced that the responsibility for this intervention would lie with the person that used the robot and did not prevent unwanted interference with human. It is his matter and his responsibility to treat the potentially dangerous robot like any other potentially dangerous thing. No legal norm contains exact rules for handling e.g., a knife, nor an explicit ban on using a knife to attack the body integrity

[7] Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial intelligence act) and amending certain union legislative acts, 21.04.2021. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 05.10.2021).

[8] M. Husovec a M. Mesarčík refer to the opinions of prof. Lawrence Lessig as a supporter of the opinion on the need to create a specific legal branch, and the opinions of prof. Frank Easterbrook, who would consider the creation of such a legal branch to be "horse law". See HUSOVEC, M., MESARČÍK, M., ANDRAŠKO, J.: *Právo informačných a komunikačných technológií 1*, Bratislava: TINCT, 2020, p. 18, ISBN 978-80-973837-0-1.

of another, however, its use for such an attack constitutes a more severe criminal sanction for the offender than an attack without a knife. Also e.g., no legal norm prevents a user from using a robot to decrypt wireless communication in his own (home) network to obtain a password, but using the same robot to find a password to access another network (i.e., overcoming a security measure using a robot) can lead to criminal sanctions.[9] We can see approach based on legal responsibility of human when using robots also in the European parliament resolution of 6 October 2021 on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters, where the European parliament "underlines that legal responsibility and liability must always rest with a natural or legal person, who must always be identified for decisions taken with the support of AI".[10]

Of course, in the context of robots development, as well as in terms of their use, it is appropriate to define rules ensuring a high level of security for the fundamental rights and freedoms of individuals that can be affected by use of robots (which is ultimately one of the objectives of the proposal for an Artificial intelligence act, whereas point 1.1 of the explanatory memorandum to this proposal states that "predictable, proportionate and clear obligations are also placed on providers and users of those systems to ensure safety and respect of existing legislation protecting fundamental rights throughout the whole AI systems' lifecycle"[11]), but this does not change the fact that the addressee of these standards is always human, and thus that the robot is still perceived as an object, not a subject of legal relations.

6.2 Legal basis for protection against artificial intelligence

As mentioned several times, a robot can influence an individual's legal position. Both positive and negative. At first sight, there is no legal norm in legal

[9] According to § 247 par. 1 of the Criminal Code, whoever overcomes a security measure and thus gains unauthorized access to a computer system, or its part shall be punished by imprisonment for up to two years.

[10] Par. 13 European Parliament resolution of 6 October 2021 on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)), 06.10.2021. Online: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html (quoted 09.10.2021).

[11] Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial intelligence act) and amending certain union legislative acts, 21.04.2021. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 09.10.2021).

system of Slovak republic that would explicitly regulate the rights of individuals to protect against the use of robots. But I believe that such norm exists. It is incorporated in directly applicable Art. 22 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).[12] This norm sets *a priori* a general ban on the application of decisions based on the results of the use of artificial intelligence. However, this prohibition is relativized and therefore it (prohibition) does not apply if:

- human enters the decision-making process
- the decision does not have legal effects with significant influence on person
- operator fulfilled at least one of the explicitly stated exceptions under Art. 22 par. 2 GDPR and in the case of processing of a special category of personal data also conditions under Art. 22 par. 4 GDPR.

The conditions for eliminating the ban on the use of robots are set quite broadly, which encourages us to think more in the direction that although Art. 22 GDPR systematically begins by setting the ban, or in better words by the right of every person not to be subject to robotic decisions (regardless of whether the decision is preceded by robotic profiling or not. It should be noted that profiling "is only a basis for decision-making"[13]), it is on the contrary a legal basis allowing the use of artificial intelligence, together with setting fundamental conditions of its use. Generally speaking, a robot can always be used to make decisions (we are considering a decision falling within the scope of GDPR). However, a robot may be used for decision-making with legal effects or with effects similar to a direct legal effect on person, only if a human intervenes in the decision-making process or if one of the exceptions under par. 2 Art. 22 GDPR is fulfilled.

In other words, robots can make decisions autonomously and without the participation of a human element only if the decision does not have legal ef-

[12] See also for example ULÍČEK, M., DONÁT, J., NONNEMANN, F., LICHNOVSKÝ, B., TOMÍŠEK, J.: GDPR. Obecné nařízení o ochraně osobních údajů. Praktický komentář, 30.04.2017. Online: https://www.smarteca.cz/my-reader/19048_20170607_0?fileName=Text%2FGDPR_-_Obecne_narizeni_o_ochrane-9.xhtml&location=pi-5684 (quoted 11.10. 2021).

[13] CYPRICHOVÁ, A. In: HUDECOVÁ, I., CYPRICHOVÁ, A., MAKATURA, I. a kol.: *Nariadenie o ochrane fyzických osôb pri spracúvaní osobných údajov/GDPR, 1. zväzok*, 2. vydanie, Žilina: Eurokódex, 2020, p. 467, ISBN 978-80-8155-094-2.

fects concerning person, or effects that significantly affect person, similar to legal effects[14] (e.g., automated targeted advertisement for sportswear when visiting website) or such a decision is permitted by European Union law or by the law of the member state to which the operator is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests).[15] In other cases, human intervention is required either *ex post* (if an autonomous decision of the robot was necessary to conclude or perform the contract between the data subject and the operator[16] or based on the explicit consent of the data subject[17]), i.e., decision of robot can effect person even if the person concerned does not request human intervention or *ex ante*, which *de facto* means that it will not be an autonomous decision of the robot.

Exactly, *ex ante* human participation is one of the conditions which application will, in my opinion, be questionable in practice. The Working Party (WP29) set up under article 29 of Directive 95/46/EC of the European parliament and of the Council of 24 October 1995 interprets the requirement of human element in automated decision-making with legally significant effects in way, that "the controller cannot avoid the Article 22 provisions by fabricating human involvement. For example, if someone routinely applies automatically generated profiles to individua ls without any actual influence on the result, this would still be a decision based solely on automated processing. To qualify as human involvement, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be conducted by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the relevant data. As part of their DPIA, the controller should identify and record the degree of any human involvement in the decision-making process and at what stage this takes place."[18] In my opinion, this interpretation needs to be clarified to avoid going

[14] Art. 22 par. 1 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) (hereinafter as the "GDPR Regulation").

[15] Art. 22 par. 2 letter b) of the GDPR Regulation.

[16] Art. 22 par. 2 letter a) of the GDPR Regulation.

[17] Art. 22 par. 2 letter c) of the GDPR Regulation.

[18] Guidelines of WP29 on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 03.10.2017, as last Revised and Adopted on 6 February 2018, p. 21. Online: https://ec.europa.eu/newsroom/article29/items/612053 (quoted 11.10.2021).

counterproductive. Namely, this opinion of WP29 can also be interpreted as a requirement to supervise the decision, not as a requirement to take a decision; this is also confirmed by the fact that the WP29 document implies, that human participation does not need to be replaced, but only increased.[19] In my opinion, however, the legal requirement of Art. 22 par. 1 GDPR is absolutely clear. Human must decide (cause consequences), not supervise. The process of creating any decision (final stage) begins where the phase of obtaining and evaluating the basis for the decision (preparing stage) ends. It seems unrealistic to demand effective "human supervision" at the decision-making process in a final stage. The only way to ensure human's participation is that human must "make a decision", not supervise. Of course, the entire process (preparing stage) until the decision-making phase can be exclusively automated. In my opinion, even fully automated formalization of the decision (in the sense of drawing up a draft of decision by robot) and placement of a facsimile is not even ruled out. However, human must decide whether the decision thus prepared will be used to produce the effects it envisages.

6.3 Legal measures to protect against artificial intelligence

The legal system offers several means of protection against robots' decisions. Finally, the General data protection regulation has become more widely known thanks to (or rather, unfortunately) the sanctions that serve as a consequence of violating the rights or obligations regulated by GDPR. In particular, the infringement of the data subjects' rights pursuant to Art. 22 GDPR shall be subject to administrative fines up to 20 000 000 EUR, or in the case of an undertaking, up to 4% of the total worldwide annual turnover of the preceding financial year, whichever is higher.[20] However, it is an administrative sanction which, from the point of view of the individual's legal position, may not be of real significance. Removal of the negative effects of robots' decisions is therefore also possible, for example, through compensation for damage or non-pecuniary damage, as envisaged in Art. 82 GDPR. In many cases, however, that may not be enough; there may be a need to reverse the robot's

[19] This approach is confirmed by WP29 elsewhere in the document in question, saying that it is sufficient to significantly increase the level of human intervention. See: Guidelines of WP29 on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 03.10.2017, as last Revised and Adopted on 6 February 2018, p. 30. Online: https://ec.europa.eu/newsroom/article29/items/612053 (quoted 14.10.2021).

[20] Art. 83 par. 5 letter b) of the GDPR Regulation.

decision immediately, because later compensation for damage or non-pecuniary damage may not be sufficient (very typical is a dismiss of a bank loan application due to a malformed profile /let's assume that this is an online process/, when e.g., the robot drew data from databases from the expired obligations of which were not removed. Another example might be a notice of breach of work discipline made by a robot that watches the time employees spend online during working hours, and the robot incorrectly evaluates that an employee did not visit the site to perform work tasks). In such cases, there must be a means for effective protection, which also presupposes Art. 79 par. 1 GDPR. A necessary measure issued by a court in civil proceedings may also be one such means.

6.3.1 Necessary measure

One of the most important means of judicial protection of affected or endangered rights of individuals (in general) in legal system of Slovak republic is the institute of necessary measure. It belongs to the group of acute judicial measures, and it has replaced the previously applicable institute of preliminary measure. The theory ranks necessary measure among the means of procedural security, which are a traditional part of the right to a fair trial. Procedural security means are characterized by priority of time, when need to protect endangered or infringed rights is seemingly so acute that it is not appropriate to insist on full finding of the facts and thus on a thorough legal assessment of the case.

It should be noted that the application of the necessary measure also falls within the competence of the Office for personal data protection of the Slovak republic. But regarding the material scope of Act no. 18/2018 Coll. on protection of personal data and the amendment of certain laws, as amended (which is legal base for competence of this Office), necessary measure regulated by this act is not an institute designed to protect individual rights. In addition, it is expressly said in that act that "personal data protection supervision does not involve disputes arising out of contractual or pre-contractual relationships between the controller or intermediary and the person or other persons concerned, which are subject to the courts or other authorities under specific regulations".[21] Although interference with non-contractual relationships (e.g., fictive invoice for services based on the monitoring of activities on the Inter-

[21] § 81 par. 5 Act no. 18/2018 Coll. on protection of personal data and the amendment of certain laws, as amended.

net) is not explicitly excluded in this act, in my view, seeing the substantive scope of the act, such disputes also fall within the authority of courts.

I consider necessary measure, as regulated in the third part, third head of Act no. 160/2015 Coll. Civil proceedings code for adversarial proceedings, as a relevant tool for urgent judicial protection of individuals rights affected by decision of robots. Its fundamental importance lies in the fact that there is no need to develop the civil court proceedings in its entirety[22] and, also, in the speed of intervention.[23] However, that advantage is balanced by the condition that there must exist objective need for urgent regulation of the situation or that there is a risk that the later execution proceedings will be spoiled;[24] from the point of view of protection against the robot's decision, the first of these conditions will probably be applied in most cases, i.e., the need for regulation of the situation without delay.

The mandatory requirements of proposal to order a necessary measure by court are:

- general requirements of the action in civil proceedings
- a description of the decisive facts justifying the need to immediate regulation of situation or justifying the fear that the execution proceedings will be spoiled
- a description of the facts credibly proving the merits and duration of the claim to be protected
- content of the requested necessary measure.[25]

It is therefore essential, first of all, that there exists a claim to be protected. In the case of a robot decision, at least two categories of claims can be considered:

- the first is the subjective right itself affected by the robot's decision
- the second is the right not to be subject to a decision based solely on automated data processing, as guaranteed by Art. 22 par. 1 GDPR.

[22] According to § 329 par. 1 first sentence of the Civil proceedings code for adversarial proceedings, the court may decide on an application for necessary measure without hearing and giving the opinion of the parties and also without ordering a hearing. According to § 330 par. 2 of the Civil proceedings code for adversarial proceedings, if the nature of the case so permits, the court may order necessary measure, the content of which would be identical to the statement in the main proceedings.

[23] According to § 328 par. 2 of the Civil proceedings code for adversarial proceedings the court shall decide about proposal to order necessary measure no later than within 30 days from the delivery of the proposal, which meets the requirements pursuant to § 326.

[24] § 325 par. 1 of the Civil proceedings code for adversarial proceedings.

[25] § 326 par. 1 of the Civil proceedings code for adversarial proceedings.

As an example, we can use a situation where a robot (on behalf of its employer) notifies an employee about a less serious breach of work discipline. Notice will be based on robot-made analysis of use of an internet browser by employee during working time. The employee may, by proposal for necessary measure, demand that the employer will not take this decision of robot into account. However, it will be crucial which category of entitlement will employee decide to protect. If he applies the first category, he must credibly certify that he has not committed a breach of work discipline and that he applies for protection against arbitrary dismissal from work. In this case, the employee will also have to describe the facts from which it will be clear, that the robot incorrectly evaluated the use of the Internet browser, whereas the employee is at risk of dismissal for repeated, less serious breaches of work discipline, as the internet browser continues to be used in his work in a way that the robot has evaluated as private use.

But if the employee applies the second category, the threat of his dismissal for repeated, less serious breaches of work discipline will still be relevant, but it will not be necessary to examine the correctness or incorrectness of the robot's decision. It is sufficient for the employee to seek protection of his right not to be subject to the robot's decision.

At first sight, therefore, the application of the protection of law under Art. 22 par. 1 GDPR seems to be more appropriate and faster alternative. However, I do not consider that the application for necessary measure can be regarded as complete, until the applicant has credibly established that the robot 's decision was taken without legally significant human intervention or that the exception under Article. 22 par. 2 GDPR.

It is requested to add that due to the legal nature of the robot (i.e., an intangible asset), a necessary measure must be ordered against the entity that used the robot.

6.3.1.1 Prerequisites for a successful certification of a protected claim in a proposal for necessary measure

One of the main principles of the General data protection regulation is the transparent access to the person whose personal data are processed. It also includes the operator's obligation to notify the data subject of the existence of fully automated decision-making.[26] Naturally, these are cases of allowed solely automated decision-making, i.e., one that is based on the exceptions

[26] Art. 13 par. 2 letter f) and Art. 14 par. 2 letter g) of the GDPR Regulation.

under Art. 22 par. 2 GDPR or it is a decision without legal effects. It is therefore possible to assume that the operator will notify person only if it is a permitted use of the robot.

However, the right to seek court protection through necessary measure, based on the person's right not to be subject to a decision made solely by automated data processing, requires the controller to behave contrary to the content of the notification in such a way that he deviates from the permitted use of the exception (e.g., the data subject will be informed that the decision is made by a person on the basis of a profile created by a robot, but in reality the person does not influence the decision-making process in any way) or operator violates an absolute ban (without fulfilling the notification obligation). The role of the person concerned seeking protection by necessary measure will be to certify exactly these facts (credibly enough for court), which requires a detailed knowledge of the controller's process of decision making.

I consider that in the vast majority of cases this task will be impossible for the person concerned, especially in cases where the decisions are just formally signed by the human (e.g., if the human signs a decision prepared by a robot without any consideration, i.e., when signature is only a symbolic gesture, warned by WP29 in the opinion mentioned in point 2 of this paper). Even more so when there is a breach of the operator's notification obligations, from which the person concerned could know of existence of a fully automated decision.

Conclusion

In my opinion, the legal system should not be a cookbook. It is okay for private law to regulate the behavior of individuals by setting principles and boundaries. This style of regulation allows the legal system to respond very flexibly and quickly to societal changes. Of course, this requires an elevated level of legal culture (including the culture of lawyers). I think it is good that the general data protection regulation regulates the use of robots (which can affect a person's legal status) in this way.

Nevertheless, artificial intelligence, its legal status and responsibility for its actions are the subject of controversy, and the legal system will undoubtedly have to face these issues. I believe that today there already exists regulation to protect individuals rights against decisions of robots. It is the personal data processing legislation. At the same time, it is possible to use traditional

means of judicial protection for this purpose, including means of procedural security, which also include a necessary measure issued by a court in civil proceedings. In this context, artificial intelligence can be perceived as an object, not a subject of legal relations, and it is a computer program as an object of intellectual property. The conduct of artificial intelligence is the responsibility of the entity which used artificial intelligence (or allowed artificial intelligence to act) and the means of protection against the effects of the use of artificial intelligence must be applied against that entity.

Despite theoretical admissibility, in practice it will not be easy, and in many cases not even real, to seek urgent court protection against a robot's decision because the person concerned will not be able to credibly certify (in proposal for necessary measure), that the ban on use of robot was violated. The person concerned remains to seek urgent protection just like in case of a human's decision, but this changes the material substance of protection – it will not be protection against artificial intelligence, but protection against the content of the decision.

Bibliography

1. Act no. 160/2015 Coll. Civil proceedings code for adversarial proceedings, as amended.
2. Act no. 18/2018 Coll. on the protection of personal data and on the amendment of certain laws, as amended.
3. Act no. 300/2005 Coll. Criminal Code, as amended.
4. DEMČÁK, S.: Právne postavenie umelej inteligencie. In: *Bulletin slovenskej advokácie,* no. 11, 2020, p. 26, ISSN 1335-1079.
5. European Parliament resolution of 6 October 2021 on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)), 06.10.2021. Online: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.html (quoted 14.10.2021).
6. Guidelines of WP29 on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 03.10.2017, as last Revised and Adopted on 6 February 2018. Online: https://ec.europa.eu/newsroom/article29/items/612053 (quoted 14.10.2021).
7. Handbook on European data protection law, 2018 edition. Online: https://fra.europa.eu/sites/default/files/fra_uploads/fra-coe-edps-2018-handbook-data-protection_en.pdf (quoted 14.10.2021).

8. HUDECOVÁ, I., CYPRICHOVÁ, A., MAKATURA, I. a kol.: *Nariadenie o ochrane fyzických osôb pri spracúvaní osobných údajov/GDPR. 1. zväzok*, 2. vydanie, Žilina: Eurokódex, 2020, ISBN 978-80-8155-094-2.
9. HUSOVEC, M., MESARČÍK, M., ANDRAŠKO, J.: *Právo informačných a komunikačných technológií 1*. Bratislava: TINCT, 2020, 262 p., ISBN 978-80-973837-0-1.
10. Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial intelligence act) and amending certain union legislative acts, 21.04.2021. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 14.10.2021).
11. 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).
12. Slovník súčasného slovenského jazyka A – G, H – L, M – N, 2006, 2011, 2015. Online: https://slovnik.juls.savba.sk/?w=inteligencia&s=exact&c=Rc6b&cs=&d=kssj4&d=psp&d=ogs&d=sssj&d=orter&d=scs&d=sss&d=peciar&d=ssn&d=hssj&d=bernolak&d=noundb&d=orient&d=locutio&d=obce&d=priezviska&d=un&d=pskcs&d=pskeп# (quoted 14.10.2021).
13. Synonymický slovník slovenčiny, 2004. Online: https://slovnik.juls.savba.sk/?w=umel%C3%BD&s=exact&c=Mcab&cs=&d=kssj4&d=psp&d=ogs&d=sssj&d=orter&d=scs&d=sss&d=peciar&d=ssn&d=hssj&d=bernolak&d=noundb&d=orient&d=locutio&d=obce&d=priezviska&d=un&d=pskcs&d=psken#.
14. ULÍČEK, M., DONÁT, J., NONNEMANN, F., LICHNOVSKÝ, B., TOMÍŠEK, J.: GDPR. Obecné nařízení o ochraně osobních údajů. Praktický komentář, 30.04.2017. Online: https://www.smarteca.cz/my-reader/19048_20170607_0?fileName=Text%2FGDPR_-_Obecne_narizeni_o_ochrane-9.xhtml&location=pi-5684 (quoted 14.10.2021).

Author details

JUDr. Martin SERFŐZŐ
external PhD. student
Pavol Jozef Šafárik University in Košice
Faculty of Law
Department of Civil Law
martin@serfozo.sk

7 AN IDENTITY PROBLEM OF AI. WHO OR WHAT IS AI?

Renáta Kišoňová

Abstract

Submitted paper focuses on philosophical and legal consideration about artificial intelligence as a subject. As an increasingly frequent participant in legal relations, artificial intelligence must be the bearer of rights and obligations. But how can we define artificial intelligence in legal codes and what identity do we assign to it? The Civil Code deals with the terms natural and legal person and presupposes a person with them (a legal person also acts through natural persons-people). In the presented text, we will rely on the concept of Hans Kelsen, who perceived the term "person" (whether natural or legal) only as a general term for a bundle of legal obligations and rights. The term person or man is obscured by anthropomorphizing and personifying language. I will state its limits in the case of the subject of copyright, which can be artificial intelligence (today AI composes music, writes poetry, creates visual works and other). Copyright law attributes authorship to a natural person (which, however, is not AI) and thus raises the question of how to define the authorship of artificial intelligence. Who, or what is actually artificial intelligence and how can one think about its identity, freedom, emotionality, responsibility?

Introduction

Submitted paper[1] focuses on the problem of AI identity. Who or what is AI? I will discuss mentioned topic in the context of Hans Kelsen´s law philosophy and in the context of Copyright Act of Slovak Republic. Dominantly I will use the hermeneutic method (considering Kelsen´s conception, his terminology and investigation), etymology (considering the terms author, artificial intelligence, identity), analytical method, historical-philosophical approach. The text is interdisciplinary questioning about AI identity, considering scientific empirical data (mostly from evolutionary and social psychology)results and contemporary Slovak law. The main assumption of my work is that Slovak Civil Law does not deal with the concept of AI in any form, neither in copyright law nor in the interpetation of citizen and personality. I will also focus on Act no. 185/2015 Coll. on Copyright and Related Rights (Copy-

[1] This work was supported by the Slovak Research and Development Agency under the Contract no. APVV-19-0166.

right Act). According to my hypothesis, mentioned Act no. 185/2015 Coll. On Copyright and Related Rights does not consider and does not include AI among the authors defined in the law. The goal of submitted text is related to my consideration about identity of AI and with the need to amend the mentioned Act no. 185/2015 Coll. On Copyright and Related Rights.

7.1 Artificial Inteligence as an author

I am going to discuss the relationship between AI as an author, creator of art and ability to create, ability of new creation. The ability to create art[2] has long been attributed exclusively to humans. It has been understood as a exceptionality which differentiates us from the rest of the animal kingdom. And it also distinguished us from machines and robots. Objective observation does not bring much information as, for instance, depicted in a documentary Le Mystére Picasso of 1955, in which a viewer can see a birth of a creative intention and a process of the work of a painter. A camera shoots all Picasso´s motions in details which are from time to time accompanied by director´s questions and Picasso´s replies. However, we observe only a way how the artist sketches particular compositions, throws away sketches, draws new ones and permanently redraws something and so on. The documentary shows improvisatory spontaneity and Picasso´s virtuosity, yet neither the working author himself, nor a viewer knows what will arise from the creative activity. Picasso himself commented on this issue: "When one is working, he does not know what the result will be. It is not indecisiveness. A change takes place no sooner than during work." A substance of artistic creation remains veiled in mystery also due to the fact that plenty of its moments are irrational, intuitive and random. As several artists (Kafka, Kundera, Wagner, Eco, Goethe and others) assume, a coincidence cooperates with a creator only if he is prepared. Artistic creativity is a matter of a whole personality. It is not shaped only by a few exceptional abilities. If the artist´s personality is to bring new aesthetic values, it has to be manifold, sensitive and original. The most fre-

[2] Artistic creation represents a certain type of creativity. According to Blažeková´s typology we can distinguish *existential* and *instrumental* creativity. BLAŽEK, B., OLMROVÁ, J.: *Krása a bolest*, Praha: Pyramída, 1985, p. 414, ISBN 11-101-85. Existential creativity has the following characteristics: it interferes into all spheres of human existence, cannot be taught, remains non-professional in its expression and focuses on a sense of life. On the other hand, instrumental creativity concerns rather specific features of a human being, can be taught and leads to professionalization and improvement of performance.

quently mentioned properties connected with artistic creativity are as follows: sensitivity (perceptiveness, problem sensitivity), fluency (ability to produce quickly and heterogeneously, finding of various alternatives of a solution), flexibility (elasticity of thinking), originality (ability to produce new ideas), redefinition (ability to repeatedly determine and remake), elaboration (ability to develop thoughts), analytical character (ability of being good at analysing and using existing experience, dividing a whole into parts), synthetic character (ability to apply old elements and experience in new means of creation) and improvisation (ability to produce without preparation, according to a current mood). The introduced creative abilities are quite complex. Except for the abovementioned, the following creator´s characteristics are also presented in research as decisive: self-madness and self-reliance, independence, self-control, self-development, assertiveness, contemplativeness, variability, immediacy, spontaneity, creativity, dreaminess, fantasy, openness, and mental wealth (i.e. rich inner life,[3] processed painful experience, etc.).

We can see that an artistically creative individual lives rich inner life, works on himself, is open to world and change, while at the same time remains himself, suffers from uncertainty, fear and pain.[4] Artistic creativity is connected with learning, observation, imitation. As Démuth says, artists discover and learn new algorithms and algorithms represent a methodological set of steps that need to be taken in order to achieve a goal. When we learn the algorithm, we can use it repeatedly to create the same or similar works.[5] And "what works for education of young artists works also for artificial intelligence".[6] Companies like Google, or Tesla use mimicking the functioning of neural networks. AI has recently become capable of produc-

[3] And rich emotional life also, regarding moral, social and aesthetical emotions. Let us remind Aristotle, who presented the basic emotions of anger, love, hate, meekness, fear, courage, shame, kindness, compassion, envy and rivalry in the 2nd book of Rhetoric in relation to the speaker's personality considering the personality of speaker, author, or creator, or, Jean Paul Sartre who has described in his book Nausea deep and existential experience of disgust, which has been gushing from his main character´s feelings towards people, things around him, his own face and body. *Nausea*, one of the canonical works of existentialism is the story of Antoine Roquentin, a writer who is horrified at his own existence. In an impressionistic, diary form he ruthlessly catalogs his feelings and experiences. His thoughts culminate in a pervasive, overpowering feeling of nausea.

[4] KULKA, J.: *Psychologie umění,* Praha: Grada, 2008, p. 392–398, ISBN 9788024723297.

[5] DÉMUTH, A.: Umenie a umelá inteligencia – výzvy a nebezpečenstvá. In: *Espes. The Slovak Journal of Aesthetics,* vol. 9, no. 1, 2020, p. 28, ISSN 1339-1119. Online: https://espes.ff.unipo.sk/index.php/ESPES/article/view/169/201 (quoted 2.11.2021).

[6] *Ibidem*.

ing ever-more complex creations which are becoming increasingly indistinguishable from works made by human beings. For example there are algorithm-created paintings the complexity and unconventional style of which were anonymously judged to be superior to human efforts. Therefore the question of changes in legal regulations in criminal law, copyright law and so on arises. Existing laws were not built to address the issues created by artificial inteligence dynamic growth. One of the most intriguing and relevant question is whether AI may be considered an 'author' in the context of copyright law. The Act protects the link between the author and his or her work. As a result, if we were to assume that artificial intelligence may not be an author, and that its creations do not constitute works, then these creations are not subject to the Act's copyright protections. This, in turn, would preclude the existence of any copyrights for any such creation. As Katarzyna Szczudlik says in the light of Polish law, and there is the same approach in Slovak law: "If we are to assume that for the purposes of the Act, AI may not be an author, we must then determine whether any other person owns the copyrights to the AI's creations. This potentially includes the AI's owner or the person who created it. On the other hand, if we assume that neither the owner, creator nor the AI itself may be considered an 'author' then no one can claim copyright protection for the AI's creations. This raises the question of why anyone would develop AI capable of creating paintings or writing novels if these potentially copyright-eligible creations could not generate profits for its owners?"[7] In 1940, during the World War II, the electronic computer designed by Alan Turing successfully cracked encrypted messages that had been sent by the Nazi command to its naval fleet. Why am I mentioning Turing? Turning back to artist as a student, his idea was that we might have a computer that could learn beyond its original instructions. Fast forward sixty, seventy years, we have seen machines displaying much of what Turing described.[8] This consideration led me to the question of AI identity. Is there

[7] SZCZUDLIK, K.: Can Artificial Intelligence be the Author? newtech.law 15.2.2018, Online: https://newtech.law/en/can-artificial-intelligence-be-the-author/ (quoted 5.11.2021).

[8] Gary Kasparov, world chess champion, lost a match to IBM's Deep Blue computer, remarking that he could "smell" a new form of intelligence across the table; and in 2016, Lee Sedol, one of the best players of the Chinese game Go, was beaten 4-1 by Google's AlphaGo program. PEARLMAN, R.: *Recognizing Artificial Intelligence (AI) as Authors and Inventors Under U.S. Intellectual Property Law.* In: Richmond Journal of Law & Technology, no. 2, 2018, p. 1–37, ISSN 1091-7322. Online: https://jolt.richmond.edu/recognizing-artificial-intelligence-ai-as-authors-and-inventors-under-u-s-intellectual-property-law/ (quoted 3.11.2011).

even any identity? We are recently far from the stage, when AI will be capable of empathy, or will experience emotions. Even though AI is able to imitate emotions according to facial expressions, it is not able to experience moral, social (or even aesthetical) emotions of disgust, admiration, anger and so on. And I suppose that identity also includes experiencing and realizing identity. One of the most generous definitions of identity says, that identity means: "condition or character as to who a person or what a thing is; the qualities, beliefs, etc., that distinguish or identify a person or thing".[9] In this sense, AI includes identity[10] in the meaning of its qualities, characters, in other words algorithms.

7.2 Hans Kelsen´s approach to a person, personality and subject of law

When we take into account all the mentioned ideas, the concept of philosopher Hans Kelsen[11] (1881–1973) seems inspiring. He labelled his theory of positive law "the pure theory of law". Kelsen explained the nature of its purity: "It seeks to preclude from the cognition of positive law all elements foreign thereto. The limits of this subject and its cognition must be clearly fixed in two directions: the specific science of law, the discipline usually called jurisprudence, must be distinguished from the philosophy of justice, on the one hand, and from sociology, or cognition of social reality, on the other."[12] Kelsen has opened the possibility of understanding the person in the meaning of auxiliary or artificial concept created by legal knowledge. In his concept, the juristic person represents a legal substance to which only rights and duties belong as its legal quality. "The idea, that 'the person has' duties and rights, involves the relation of substance and quality."[13] According to Kelsen,

[9] Dictionary.com. *Identity.* Online: https://www.dictionary.com/browse/identity (quoted 6.11.2021).

[10] Etymology of the term identity has appeared in 1600, in the meaning "sameness, oneness, state of being the same," from French *identité* (14c.), from Medieval Latin *identitatem* (nominative identitas) "sameness," ultimately from Latin *idem* (*neuter*) "the same". Online Etymology Dictionary. *Identity.* Online: https://www.etymonline.com/word/identity (quoted 6.11.2021).

[11] Hans Kelsen was a leading legal positivist. His major works on legal positivism were The General Theory of Law and State andThe Pure Theory of Law.

[12] KELSEN, H.: *The Pure Theory of Law,* Berkeley: University of California Press, 1967, p. 266, ISBN 1-58477-206-9.

[13] *Ibidem*, p. 93–94.

this assumption maintained by traditional jurisprudence is inaccurate. The legal rule can not determine the whole existence (even identity) of a human being. A human being may be a person "at law" only with respect to a certain extent of acting and when the conduct of a human being is not the subject of legal rights, the human being is not in relation to the legal order. He says that the mistake of traditional legal thinking subsists in seeking something to which the duties and rights belong. As Kelsen said, "the person exists only insofar as he or she 'has' duties and rights; apart from them the person has no existence whatsoever. To define the physical (natural) person as a human being is incorrect, because man and person are not only two different concepts but also the results of two entirely different kinds of consideration. Man is a concept of biology and physiology, in short, of the natural sciences. Person is a concept of jurisprudence, of the analysis of legal norms."[14] And "that the human being is a legal subject (subject of rights and obligations) means nothing else, as has been emphasized, but that the human behaviour is the content of legal obligations and legal rights – nothing else than that a human being is a person or has personality".[15] In this sense the term "person" may be extended to the AI concept. And for Slovak law it represents a challenge for future adjustments.

Conclusion

In the submitted paper I presented broader interdisciplinary view on the issue of AI, creative process, art, author, person, Hans Kelsen´s theory of person. Artificial intelligence as a creator (of art for example) should be included into the legislation, which represents a challenge for possible changes in Slovak civil and criminal law as well.

Bibliography

1. BLAŽEK, B., OLMROVÁ, J.: *Krása a bolest*. Praha: Pyramída, 1985, ISBN 11-101-85.
2. DÉMUTH, A.: Umenie a umelá inteligencia – výzvy a nebezpečenstvá. In: *Espes. The Slovak Journal of Aesthetics,* vol. 9, no. 1, 2020, p. 26, ISSN 1339-1119. Online: https://espes.ff.unipo.sk/index.php/ESPES/article/view/169/201 (quoted 2. 11. 2021).
3. Dictionary.com. Online: https://www.dictionary.com/ (quoted 6. 11. 2021).

[14] *Ibidem*, p. 94.
[15] *Ibidem*, p. 94–95.

4. KELSEN, H.: *The Pure Theory of Law.* Berkeley: University of California Press, 1967, ISBN 1-58477-206-9.
5. KULKA, J.: *Psychologie umění.* Praha: Grada, 2008, ISBN 9788024723297.
6. Online Etymology Dictionary. Online: https://www.etymonline.com/word/identity (quoted 6.11.2021).
7. PEARLMAN, R.: *Recognizing Artificial Intelligence (AI) as Authors and Inventors Under U.S. Intellectual Property Law.* In: Richmond Journal of Law & Technology, no. 2, 2018, p. 1–37, ISSN 1091-7322. Online: https://jolt.richmond.edu/recognizing-artificial-intelligence-ai-as-authors-and-inventors-under-u-s-intellectual-property-law/ (quoted 3.11.2011).
8. SZCZUDLIK, K.: Can Artificial Intelligence be the Author? newtech.law 15.2.2018, Online: https://newtech.law/en/can-artificial-intelligence-be-the-author/ (quoted 5.11.2021).

Author details

Mgr. Renáta KIŠOŇOVÁ, PhD.
Comenius University in Bratislava
Faculty of Law
Department of Theory of Law and Philosophy of Law
renata.kisonova@flaw.uniba.sk

8 ISSUE OF LEGAL PROTECTION TO AI-CREATED OBJECTS

Yuliya S. Kharitonova

Abstract

New technologies make it possible to create and deploy new objects of digital art and digital innovations to put them into circulation for the use of new solutions of technological nature. Based on the analysis of the basic approaches to the definition of the properties of copyright and patentability of objects of intellectual property rights in the article formed the legal approaches to the definition of the nature and characteristics of the new objects of intellectual property rights, created by AI. Since strictly speaking, the objects created by AI are not the objects of the exclusive rights of authors and inventors, the thesis points out the need to develop a specific sui generis legal regime for AI outputs.

AI, Technological Convergence and Creativity

The most common approach to understanding AI is that artificial intelligence (AI) is the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. The term is frequently applied to the project of developing systems endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience.[1] So as a discipline of computer science aimed at developing machines and systems, artificial intelligence technology aims to solve problems deemed to require human intelligence with little or no human intervention.

Russia has adopted The National strategy for the development of artificial intelligence for the period up to 2030 – AI Strategy 2030, which reveals the concept of artificial intelligence. We weren't the first, that's why we had a chance to explore best practices to develop a strategy. National strategies have been launched in about 30 countries, while about 10 others are working on it. In February 2020, the European Commission issued the *White Paper on Artificial Intelligence (A European approach to excellence and trust)*, while in September 2020 the Committee of Legal Affairs issued the *Report on Intel-*

[1] COPELAND, B. J.: artificial intelligence. Britannica. 9.11.2021. Online: https://www.britannica.com/technology/artificial-intelligence (quoted 5.12.2021).

lectual Property Rights for the development of artificial intelligence technologies.[2]

The Russian AI Strategy 2030 reveals the concept of artificial intelligence as *"a set of technological solutions* that allows simulating human cognitive functions (including self-learning and searching for solutions *without a pre-set algorithm*) and obtaining results that are *at least comparable to the results of human intellectual activity* when performing specific tasks".

It turns out that the human brain and AI take data from the outside and on the basis of this data make a decision or create a new artwork or invention.

Therefore, whether the outputs of AI comparable to the outputs of human intellectual activity are received as a result of applying a set of technological solutions simulating human cognitive functions, the question arises if the regime of intellectual property objects can be extended to such outputs.

The convergence of technologies manifests itself in the fact that artificial intelligence interacts with Big Data collection and analysis technology and distributed registry systems for the tokenization of AI works.

8.1 AI and Copyright: Issues of Creativity

If we assume that AI can create new works, we need to determine whether AI-generated output would be protected by IP law. According to the Russian legislation works are objects of copyright regardless of their merits and purpose, and also the way of their expression (Article 1259 of Civil Code of Russian Federation). In Russian law there is no requirement for the originality of a work. However, we suppose that in countries, where the originality of work is necessary to provide the exclusive right to the copyright, AI will require to rethink the approaches to copyright.

Why?

AI is usually equated with „weak" or „narrow" AI, which are techniques and applications programmed to perform individual tasks. Weak AI simulates human cognition.

But even now, lawyers should consider that AI will be strong in its ability to be used for works that were not originally conceived by humans. A signifi-

[2] European Commission: White Paper On Artificial Intelligence – A European approach to excellence and trust. COM (2020) 65 final. 19.2.2020. Online: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (quoted 5.12.2021).

cant dilemma then arises with respect to works created entirely by artificial intelligence without human intervention (AI-created works).

The concepts of author are compared through three key elements: author's relationship with work, author's relationship with others and presumptions about author's personality and creative process.

Today, the position of the Court of Justice of the European Union is of prime importance. By defining the concept of originality as an autonomous concept of EU law, the CJEU has taken a human-centric approach through a series of decisions. In particular, the CJEU identifies originality as the result of the author's personal intellectual creation. The CJEU concluded that it means "subject matter which is original in the sense that it is its author's own intellectual creation."[3]

The *originality* of the work is linked to human creativity. The legal concept of originality, although it is a very important prerequisite for the definition of work, is not specified by the law. The conceptual framework comes from theory but is mainly provided by jurisprudence.

Let us just note that the dynamic concept of originality moves between a human-centric approach, which puts the individuality of the author at the core, and a work-centric approach, which focuses on the individuality of the work. According to the law, it is necessary to pay attention to the mode of expression only.

This point leads us to the fact that strictly speaking, application of AI can produce original outputs with a special mode of expression. That is, they are the same as those created by humans. Researchers have put forward the thesis that it is possible to reconcile the principles of copyright on AI works by reorienting the rationale for protection from "author" to "work" as such.[4]

P. Mezei points, unless paradigm shift in copyright law, the lack of direct human element of an AI-generated output shall lead to the unavailability of copyright protection for these outputs.[5] But it is well known that in many jurisdictions only humans are recognized as the author.

[3] Court of Justice of the European Union, Judgment of 16.7.2009, *Infopaq International A/S v. Danske Dagblades Forening,* C-5/08, para. 37.

[4] MAGGIORE, M.: Artificial Intelligence, computer generated works and copyright. In: BONADIO, E., LUCCHI, N.: *Non-Conventional Copyright*, Edward Elgar Publishing, 2018, chapter 18, ISBN 9781786434067.

[5] MEZEI, P.: "You AIn't Seen Nothing yet" – Arguments against the Protectability of AI-generated Outputs by Copyright Law. University of Szeged, 20.7.2021. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3890051 (quoted 5.12.2021).

8.2 AI Inventions: The New Patentability

In Russian law an invention is a technical solution that meets the requirements of novelty, inventive step, and industrial applicability is recognized as (Article 1350 of the Civil Code of Russian Federation). The list of patentability criteria to be confirmed by the Prior Act includes inventive step, novelty, and utility of the object in the US. Under the provisions of the German Patent Act (§1 PatG – Patentfähige Erfindung – patentability of the invention), the patentability (Patentfähigkeit) of a technical solution is assessed through compliance with the criteria of novelty (Neuheit), inventive activity (erfinderische Tätigkeit) and industrial applicability (Gewerblichkeit).

In general, the criteria enshrined in the national legislation are accepted at the international level. They are the most general categories exactly until there is no need to qualify such a new and peculiar object due to artificial intelligence technology. On 21 December 2019, the European Patent Office (EPO) announced its refusal to examine two patent applications, designating an AI system DABUS as the inventor, on the formal ground of failure to fulfil the requirement of the European Patent Convention that 'an inventor designated in the application has to be a human being, not a machine'.[6]

Do AI-generated-inventions require patent protection or a similar incentive system at all? Technically, such AI system's inventive activity produces an output worthy of a patent. But in contrast to the issue of copyright, in the field of patents, the problem of identifying the novelty and inventive step of the invention is more clearly seen.

Under Russian law, an invention is new if it is not known from relevant prior art. An invention has an inventive step if it does not explicitly follow the prior art for a person skilled in the art. State of the art standard for an invention includes any information which has become publicly available in the world before the priority date of the invention.

So we can see the inventive level of technical solutions created by artificial intelligence. P. Block's thesis that the European Patent Convention does not explicitly exclude the patentability of inventions created by artificial intelli-

[6] European Patent Office: EPO refuses DABUS patent applications designating a machine inventor. 20.12.2019. Online: https://www.epo.org/news-events/news/2019/20191220.html (quoted 5.12.2021); see also European Patent Office: Summary of the relevant facts and submissions. 27.1.2020, Application Nr. 18275174.3.Online: https://register.epo.org/application?documentId=E4B63OBI2076498&number=EP18275174&lng=en&npl=false (quoted 5.12.2021).

gence systems and that the method of creation of the invention does not matter seems more convincing.[7]

A condition of patentability is that the invention involves an inventive step or be non-obvious. The standard applied for assessing non-obviousness is whether the invention would be obvious to a person skilled in the relevant art to which the invention belongs.

In the aforementioned example of works created by DABUS (AI system was named as the inventor), the patent applications were rejected by the European Patent Office because of the argument: under English law and the European Patent Convention (EPC), the term *inventor* refers only to a natural person.

8.3 Trademark Image Recognition: Confusingly Similar Designations

It does not mean that a human being can no longer be an expert in identifying innovations as inventions. However, access to data and the ability of AI to algorithmically create innovations requires at least one more argument.

Let's turn to the field of trademarks. Trademarks are intended to distinguish the origin of goods and services and to prevent consumer confusion. It is the problem of Confusion in Trademark.

Of-confusion is the term used when the infringed trademark is capable of confusing the consumer or both marks are confusingly similar to each other.

Researchers have discovered a problem: the definition of a designation as a trademark depends on human perception. It is the person who conducts the examination. It is the person as a consumer who perceives the trademark as an identifier of the goods. There are ongoing discussions about the way that AI interacts with trademarks in the online environment.

However, on the consumer side, trademark recognition is increasingly handled by AI, in connection with the applications of the smart home technology and the industrial internet.

For example, AI assistants, search engines, customer service bots and online marketplaces play an important role in shaping the consumer decision-making process. The way that a consumer interacts with the online mar-

[7] BLOK, P. H.: The inventor's new tool: artificial intelligence – how does it fit in the European patent system? In: *European intellectual property review*, vol. 39, no. 2, 2017, p. 69–73, ISSN 0142-0461.

ketplace through AI may result in the presentation of only a limited number of brands to a consumer, or other alterations in the way that consumers make product selections.

8.4 *Sui Generis* System of IP Rights for AI-generated Outputs

It is possible to expand the range of so-called non-traditional objects of intellectual rights on the principle of *sui generis*.

That special legal regime differs from the regimes of copyright and patent protection. This approach is quite justified, since the above objects are distinguished by significant specificity, and it is not entirely logical to artificially "tie" them to one institution or another, it is much more effective to develop a special legal regime.

Some scholars, like Lauber-Ronsberg worries that giving copyright protection to AI-created works will fundamentally alter the concepts of copyright law and the underlying legislative structure. The researcher admits that developing an *ad hoc* right may be a preferable approach for coping with AI's uncertainty. Nevertheless, it is concerned about the standards for the provided protection as well as the allocation of related rights.[8]

Copyright or Patenting is not the only way to give legal protection to these results. It is quite possible to grant intellectual rights to the developer (and/or owner or user) of the specified system within the framework of the sui generis institution. In addition to the norms on objects, subjects and the scope of their rights, the norms of this institution should regulate the use of big data, since in many cases their collection, processing, storage involves the appeal to personal data. A mechanism for transferring rights to such objects is also needed.

Conclusion

The most stable approach today is the one according to which artificial intelligence as a technological solution is primarily a tool in the hands of a human inventor. However, the complexity of algorithms, their continuous improvement leads to the fact that more and more opportunities are opening up for (in a certain sense) "autonomy" of artificial intelligence technology regard-

[8] LAUBER-RÖNSBERG, A., HETMANK, S.: The concept of authorship and inventorship under pressure: Does artificial intelligence shift paradigms? In: *Journal of Intellectual Property Law & Practice*, vol. 14, no. 7, 2019, p. 570–579, ISSN 1747-1540.

ing the identification and search for technical solutions for specific tasks of human life.

The correct legal framework for protecting AI-generated inventions is still in question. Protection within the current legislative system of patent law or copyright is problematic.

The need to protect AI-generated inventions is rooted in the European Union's policy of strengthening and promoting technology and innovation.

On my opinion, regarding AI-generated outputs classified into works or inventions, legal exclusivity must be ensured through sui generis rights. The above option is a clear solution and does not force the existing legal framework to incorporate AI-generated outputs that have different structural characteristics from a work or from an invention.

Bibliography

1. COPELAND, B. J.: artificial intelligence. Britannica. 9.11.2021. Online: https://www.britannica.com/technology/artificial-intelligence (quoted 5.12.2021).
2. European Commission: White Paper On Artificial Intelligence – A European approach to excellence and trust. COM (2020) 65 final. 19.2.2020. Online: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (quoted 5.12.2021).
3. Court of Justice of the European Union, Judgment of 16.7.2009, *Infopaq International A/S v. Danske Dagblades Forening,* C-5/08.
4. MAGGIORE, M.: Artificial Intelligence, computer generated works and copyright. In: BONADIO, E., LUCCHI, N.: *Non-Conventional Copyright*, Edward Elgar Publishing, 2018, ISBN 9781786434067.
5. MEZEI, P.: "You AIn't Seen Nothing yet" – Arguments against the Protectability of AI-generated Outputs by Copyright Law. University of Szeged, 20.7.2021. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3890051 (quoted 5.12.2021).
6. European Patent Office: EPO refuses DABUS patent applications designating a machine inventor. 20.12.2019. Online: https://www.epo.org/news-events/news/2019/20191220.html (quoted 5.12.2021).
7. European Patent Office: Summary of the relevant facts and submissions. 27.1.2020, Application Nr. 18275174.3. Online: https://register.epo.org/application?documentId=E4B63OBI2076498&number=EP18275174&lng=en&npl=false (quoted 5.12.2021).
8. BLOK, P. H.: The inventor's new tool: artificial intelligence – how does it fit in the European patent system? In: *European intellectual property review*, vol. 39, no. 2, 2017, p. 69, ISSN 0142-0461.

9. LAUBER-RÖNSBERG, A., HETMANK, S.: The concept of authorship and inventorship under pressure: Does artificial intelligence shift paradigms? In: *Journal of Intellectual Property Law & Practice*, vol. 14, no. 7, 2019, p. 570, ISSN 1747-1540.

Author details

Yuliya S. KHARITONOVA, Doctor, Professor
Head of the Center for Legal Studies of Artificial Intelligence and Digital Economy, a scientific and educational center
Lomonosov Moscow State University, Moscow, Russia
Law School
sovet2009@rambler.ru

9 ARTIFICIAL INTELLIGENCE AND ITS IMPACT ON INTELLECTUAL PROPERTY LAW

Renáta Romanová

Abstract

Artificial intelligence brings new situations that the current regulation of intellectual property law, whether it is copyright or patent law, does not sufficiently regulate, despite the fact that publications related to artificial intelligence, as well as patent applications related to artificial intelligence are growing rapidly. At the moment, we can say that we are in the age of frontier technologies, to which we only have to respond promptly, because this time can already begin to evolve on its own. The article talks about the problems or pitfalls of artificial intelligence in intellectual property law and also points out possible solutions to these negative effects in the future. It points out the influence of artificial intelligence in the field of trademarks, patents, but also does not forget the authors with reference to the case law in this area.[1]

Introduction

When robots act like robots, humans have increased opportunities to act like humans.[2]

We nowadays usually come across the term artificial intelligence (AI), which is nothing special, because artificial intelligence is used today in various areas of our lives, such as autonomous cars, online shopping and advertising, factory robots, cyber security and more. Although this term is often used, we still do not have a universal definition that describes what is meant by artificial intelligence. While we can say that we are at the beginning of the definition of AI, we must not forget that AI is initially considered part of computer science, which deals with the design of intelligent computer systems, systems that exhibit the properties we associate with intelligence in hu-

[1] This paper was prepared within the framework of the research project VEGA no. 1/0643/20 "Legal and Ethical Perspective of Artificial Intelligence".

[2] AGARWAL, R.: Why Low-Skilled Workers Will Win In The Robot Revolution. 16. January 2019. Online: https://www.forbes.com/sites/washingtonbytes/2019/01/16/why-low-skilled-workers-will-win-in-the-robot-revolution/?sh=5dee7d187538 (quoted 9.11.2021).

man behavior, such as understanding. language, thinking, learning, problem solving and more.[3]

Dimiter Dobrev defined artificial intelligence as: *"AI will be such a program which in an arbitrary world will cope not worse than a human"*,[4] there are also many definitions, but we will lean towards the definition that the European Commission agrees with in its strategy papers, namely: *„AI systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal."*[5]

The development of AI brings new opportunities for our company, and the risks associated with it increase in direct proportion to the opportunities. As AI goes beyond the legislation of one country, efforts are being made to harmonize legislation in the European Union, where the European Parliament has already set up a Special Committee on Artificial Intelligence in the Digital Age (AIDA) to address the potential impact of AI on the European economy. The European Parliament adopted 3 reports on 20 October 2020 on how the European Union should regulate AI in order to promote innovation but also to protect ethical standards while ensuring confidence in new technologies. These 3 rules concern: 1. the balance between the protection of citizens and the promotion of technological development, 2. the protection of individuals and businesses through a new system of civil liability, and 3. an effective system of intellectual property protection and guarantees for developers.[6] In this paper, we will focus in particular on the third point, which concerns intellectual property rights ("IPR") in connection with AI. Of course, as far as AI legislation is concerned, there are even global efforts

[3] BARR, A., FEIGENBAUM, E.: *The handbook of artificial intelligence*, 1981, p. 397–409, ISBN 978-0-86576-089-9. Online: https://doi.org/10.1016/B978-0-86576-089-9.50013-2 (quoted 9.11.2021).

[4] DOBREV, D.: A Definition of Artificial Intelligence. 19. January 2004. Online: https://arxiv.org/pdf/1210.1568.pdf (quoted 9.11.2021).

[5] GESLEY, J.: Legal and Ethical Framework for AI in Europa: Summary of Remarks. In: *Proceedings of the ASIL Annual Meeting*, vol. 114, 2020, p. 240–242, ISSN 0272-5037. Online: https://doi.org/10.1017/amp.2021.46 (quoted 9.11.2021).

[6] European Parliament News: Parliament leads the way on first set of EU rules for Artificial Intelligence. 20. October 2020. Online: https://www.europarl.europa.eu/news/en/press-room/20201016IPR89544/parliament-leads-the-way-on-first-set-of-eu-rules-for-artificial-intelligence (quoted 9.11.2021).

to unify legislation in this area, but in this article we will focus mainly on the regulation of IPR in connection with AI in the European Union.

One of the reasons why the emphasis in AI is also specifically placed on the field of intellectual property law is that, as follows from the report on the analysis at the sector level, in the period 2014-2016, so 45% of the total economic activity (GDP) in the EU is attributable to IPR-intensive industries, worth EUR 6.6 trillion or 38.9% of all employment in the EU (83.3 million) can be attributed, directly or indirectly, to IPR- intensive industries.[7] In view of the above, we can conclude that IPR has a significant position in the EU, and in the context of the current COVID-19 pandemic, it may even help the EU's economic growth. An important question is whether the current IPR legislation in the EU is also sufficient in relation to AI and provides balanced protection for machine-made works and inventions, or whether new IPR legislation needs to be introduced to clearly define the status of artificial intelligence.

9.1 Patent law

The number of published patent applications for artificial intelligence has increased by 400%[8] in the last decade, with the United States being the world leader in AI patent applications.[9] IBM is one of the world's largest patents for machine learning and AI patents, followed by Samsung and Microsoft.[10] In the EU area, 180,250 applications were submitted by 2020, and it is not surprising that during the pandemic, most of them were in the field of Medical Technology, with Computer Technology in third place.[11] In order to speed up the patentability of AI, it is necessary to address the individual problems associated with the patentability of AI.

[7] For more details, see: EUIPO and EPO: Odvetvia intenzívne využívajúce PDV a ekonomická výkonnosť v Európskej únii. Správa o analýze na úrovni odvetví, september 2019. Online: https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observatory/documents/IPContributionStudy/IPR-intensive_industries_and_economicin_EU/summary/IP_Contribution_Report_092019_execsum_sk.pdf (quoted 5.12.2021).

[8] European Parliament News: AI rules: what the European Parliament wants. 21.10.2020. Online: https://www.europarl.europa.eu/news/sk/headlines/society/20201015STO89417/formovanie-europskej-legislativy-v-oblasti-umelej-inteligencie (quoted 9.11.2021).

[9] *Ibidem.*

[10] For more details, see: European Patent Office. European patent applications. Section "Top tech. fields". Online: https://www.statista.com/statistics/1032627/worldwide-machine-learning-and-ai-patent-owners-trend/ (quoted 5.12.2021).

[11] *Ibidem.*

In the territory of the Slovak Republic (SR), patent law is regulated by Act no. 435/1990 Coll. The Patent Act (hereinafter also "PA") and Decree no. 223/2002, which serves to implement the Patent Act. The Patent Act lacks a definition of the invention but regulates which inventions are patentable in §5 (1) PA "Patents are granted for inventions in all fields of technology which are new, involve an inventive step and are industrially applicable".[12] So each solution needs to be considered from 4 perspectives:

1. protection area
2. novelty
3. inventive step
4. industrial applicability.[13]

In the EU, we have The European Patent Convention (EPC), which also does not define the invention and at the same time in art. 52 requires the same 4 criteria for granting a patent. For the needs of this contribution, the third point is important, namely the inventive activity, respectively inventive step. The EPC does not define anywhere the need of a natural person in connection with an inventive step, so we can assume that such an inventive step can also be taken by AI. Article 56 of the EPC states: *„An invention shall be considered as involving an inventive step if, having regard to the state of the art, it is not obvious to a person skilled in the art."* Given this definition, AI cannot be ruled out as an inventor.

The problem arises at the time of filing a patent application in which AI is listed as the inventor. In the process of obtaining a patent, 3 entities act, namely: the inventor, the applicant for a patent application (national or European) and the patent owner. The inventor of the invention can only be a natural person – the creator of the technical solution, who fulfills the features of the invention mentioned above and will be eligible for legal protection in the form of a patent. This is the first problem why AI cannot be recognized as an author under current legislation.

The second problematic point is the part of intellectual property law that requires the creative intellectual activity of the creator.[14] Initial phase by

[12] Zákon č. 435/1990 Zb. o patentoch, dodatkových ochranných osvedčeniach a o zmene a doplnení niektorých zákonov.

[13] VOJČÍK, P. et. al.: *Právo duševného vlastníctva*, 2. ed., Plzeň: Aleš Čeněk, 2014, p. 494, ISBN 978-80-7380-527-2.

[14] ŠKREKO, A.: Priemyslené právo v informačnej a poznatkovo orientovanej spoločnosti. In: ŠVIDROŇ, J. (ed.): *Právo duševného vlastníctva v informačnej spoločnosti a v systéme práva*, Bratislava: VEDA, 2009, p. 519, ISBN 978-80-224-1033-5.

a person who can influence them to a large extent, either by uploading specific materials or by programming the AI itself, and also only a natural person has the opportunity to think, creative thinking, which is a basic prerequisite for any solution. The machine will not ask questions like this: why is it like this, just looking for answers to the questions or tasks we give it. Although the natural person inventor does not have to fulfill the condition of legal capacity, as is the case with copyright, this does not automatically mean that the inventor should be an AI, which is still a thing. Even if we achieve that AI is given the opportunity to think, which in the foreseeable future we consider this possibility to be highly unlikely, then we will still have someone behind the creation of this AI, so we believe that at least in connection with AI, as the inventors, could only talk about the so-called co-origin. If we were to allow AI to attribute the inventions of the invention, we could encounter several problems, such as e.g., the possibility of invoking the protection of one's origin or property interests, while we consider the most difficult to invoke liability related to the invention or the conclusion of contracts, such as patent license agreement.

Subsequently, the applicant or owner of the patent may already be, in addition to a natural person, a legal entity that is registered in the patent register, but even in this case there is no room for AI. If we want to grant any right to AI, a change in the law is necessary, because the current regulation did not provide for AI and a possible inventor.

In 2018, two patent applications were filed,[15] in which the artificial intelligence "DABUS" was mentioned as the inventor of the patent, while the applicant and owner of AI DABUS is Dr. Stephen Thaler (CEO of Imagination Engines), which was rejected by the European Patent Office (EPO) on the grounds that they did not comply with the EPC because of the requirement that the inventor named in the application must be a human being and not a machine. In addition, the EPC noted that the inventor as a natural person is considered an internationally applicable standard, as other national courts have similarly ruled, e.g., in Great Britain[16] or the USA.[17] It is worth noting

[15] EPO Patent application: EP 18 275 163 – food container; EPO Patent application: EP 18 275 174 – devices and methods for attracting enhanced attention.

[16] *Thaler v Comptroller General of Patents Trademarks And Designs* [2021] EWCA Civ 1374 (21 September 2021). Online: https://www.bailii.org/ew/cases/EWCA/Civ/2021/1374.html (quoted 5.12.2021).

[17] *Thaler v. Hirshfeld,* [2021] US District Court for the Eastern District of Virginia, 1:20-cv-903 (LMB/TCB). See more details: https://www.dwt.com/-/media/files/blogs/artificial-intelligence-law-advisor/2021/09/thaler-v-hirshfeld-decision.pdf (quoted 5.12.2021).

the decision of the Australian Federal Court,[18] which ruled that under the Australian Patent Act, AI could be listed as the inventor, but this decision has been appealed and it will be interesting to follow the result, as this may greatly affect authorship in relation to with AI.[19]

In conclusion, the EPC further noted that the inventor must be legitimate to exercise his rights, i.e., it must have a legal personality that AI systems or machines do not have.[20] AI has no personal rights. In conclusion, we only note that with regard to the other criteria that must be met in order to be a patent, as we noted above, these criteria can also be met by AI. The EPC also favors this version because the only problem that led to the rejection of patent applications was that the inventor was not a man but a machine and not the invention itself.

In view of the above, we consider the possibility of listing AI as the inventor of the invention in the light of current legislation to be more harmful than beneficial, because so far AI has not been taken into account in formulating the legal wording and the inventor has always been considered an inventor. An inventor who is registered assumes some responsibility and if we only recognized AI as an inventor, problems could arise later, especially if the solution or invention began to be used after the patent expired and cause problems or damage. We consider co-invention to be a suitably chosen form, especially so that the responsibility for the invention lies not only with AI but also with the one who took the initial step and filed the patent application and is interested in the patent being registered.

9.2 Copyright

The European Union has already begun to work actively on the harmonization of copyright, which we consider to be the right move for AI, but we do not consider the form it has chosen to be very appropriate, as the implementation of directives may lead to different national copyright regulations

[18] *Thaler v Commissioner of Patents* [2021] FCA 879, VID 108 of 2021. See more details: https://www.judgments.fedcourt.gov.au/judgments/Judgments/fca/single/2021/2021fca0879 (quoted 5.12.2021).

[19] South Africa has granted AI Dabus patent, this patent was published in July 2021 in the South African Patent Journal. Online: https://iponline.cipc.co.za/Publications/PublishedJournals/E_Journal_July%202021%20Part%202.pdf (quoted 5.12.2021).

[20] EPO: EPO publishes grounds for its decision to refuse two patent applications naming a machine as inventor. 28.1.2020. Online: https://www.epo.org/news-events/news/2020/20200128.html (quoted 9.11.2021).

which may cause problems or problems. with authorship or copyright ownership. We think that it would be more appropriate to proceed with harmonization through regulations, which would also contribute to more uniform national solutions for works that would be created with the help of AI or used AI.

With AI, the same problem arises with copyright as with the patent and the patent originator. In the Slovak Republic, copyright is regulated by Act no. 185/2015 Coll. The Copyright Act (CA), which in §3 (1) defines the subject of copyright, which is *"a work in the field of literature, art or science which is a unique result of the author's creative intellectual activity through perceptible senses, whatever its form, content, quality, purpose, form of expression or degree of completion*" and also defines the author as *"a natural person, who created the work."*[21] European Copyright Code defines the author in the chapter 2 art. 2.11: *„The author of a work is the natural person or group of natural persons who created it."* The author is thus the original subject of copyright and is always the person who created the work. The natural person is needed here mainly because the work will be created on the basis of the expression of his literary, scientific or artistic creative activity, and this work will be a personal and unrepeatable creation of his soul. Authorship of a work can only arise on the basis of creative activity and not another activity such as technical, mechanical, automatic or other activities. The primary intention of continental European copyright is not to protect investment in the creation of an intangible asset but to protect unique creative activity.[22] So if the work does not contain even a part of creative freedom in the sense of copyright, then we will not be able to be a copyright work, e.g., sports matches that have defined rules of the game. [23] Prof. Vojčík states that it is not possible to talk about an author's work if it will be the result of the activities of nature, animals, robots or other machines, etc.[24]

As for uniqueness, it is necessary to understand the originality of the work, its originality and not just statistical uniqueness.[25] In Slovakia, uniqueness

[21] §13 ods.1 zákona č. 185/2015 Z. z. Autorský zákon.

[22] TELEC, I., TUMA, P.: *Autorský zákon: Veľký komentár*, 2. vyd., Praha: C. H. Beck, 2019, p. 92, ISBN 978-80-7400-748-4.

[23] Court of Justice of the European Union (hereinafter as the "CJEU"), Judgment of 4.10.2011, *FAPL Ltd and others v. QC Leisure and others,* C-403/08.

[24] VOJČÍK, P. et. al.: *Právo duševného vlastníctva*, op. quoted, p. 261.

[25] For more details, see: CJEU, Judgment of 2.5.2012, *SAS Institute Inc. v. World Programming Ltd*, C-406/10; CJEU, Judgment of 7.3.2013, *Eva-Maria Painer v. Standard Verlags GmbH and others,* C-145/10; CJEU, Judgment of 13.11.2018, *Levola Hengelo BV gegen Smilde*

is assessed even more strictly than at the European level,[26] This is mainly due to the fact that in order to determine originality, it is necessary to assess the work in such a way that those parts that gave the author freedom of creative decision, to what extent the author used this opportunity in these parts and whether his creative decisions imprint", respectively "creative personal footprint".[27]

In view of the above, it is impossible in the current legislation for AI to become the author of a work, not only because the original author can only be a natural person, but also because it will not be a unique creative intellectual activity. However, efforts to acknowledge AI authorship are still increasing. The work of art "Edmond de Belamy"[28] Rembrandt-style, created by AI in 2018 and subsequently auctioned off at Christies's Auction House in New York for $ 400,000[29] and another curiosity that happened on October 9, 2021 is the replay of Beethoven's 10th Symphony, also known as "Unfinished," which was completed by AI. Both works have one thing in common, and that is the collection of data in large quantities, whether of musical compositions or images. In connection with copyright, it is also appropriate whether there has been an infringement of copyright. In the European Union, copyright expires 70 years after the author's death, and since Beethoven has been dead for 200 years, his copyright could not have been infringed. [30]

What about works from which AI draws, but still have copyright protection? In this case, we believe that for AI, certain rules should apply to reward authors for the use of their works. Since this is not a common inspiration but a thorough analysis of the author's work by AI, it is necessary to look at it differently than from people who are only inspired by the author's work

Foods BV, C-310/17; CJEU, Judgment of 22.12.2010, *Bezpečnostní softwarová asociace – Svaz softwarové ochrany gegen Ministerstvo kultury*, C-393/09.

[26] Nález Ústavného súdu Slovenskej republiky, sp. zn. III. ÚS 651/2016 z 28. novembra 2017.

[27] Uznesenie Ústavného súdu Slovenskej republiky, sp.zn. II. ÚS 647/2014 z 30. septembra 2014.

[28] For more details, see GUADAMUZ, A.: Artificial intelligence and copyright. WIPO MAGAZINE, October 2017. Online: https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html (quoted 5.12.2021).

[29] Meyer-Dulheuer Patent- und Rechtsanwaltskanzlei: Artificial Intelligence creates art – patentable? 30.10.2018. Online: https://legal-patent.com/international-intellectual-property/artificial-intelligence-creates-art-patentable/ (quoted 9.11.2021).

[30] Meyer-Dulheuer Patent- und Rechtsanwaltskanzlei: Beethoven's 10th AI Symphony. 10.10.2021. Online: https://legal-patent.com/international-intellectual-property/beethovens-10th-ai-symphony/ (quoted 9.11.2021).

and cannot subject it to such a detailed analysis. With such a reward, it is worth mentioning the blockchain technology, on which all works that would be provided to AI could be recorded. Any handling of data that would be bound to the author's work would be recorded on this chain. Any recording with data that would be tied to the author's work would mean for the author a small amount of reward for providing his work. [31]

We are of the opinion that, as far as copyright is concerned, it should be considered even stricter here than with patents, because copyright has been created since its inception to protect and promote creativity. Therefore, it is not appropriate for us to change the meaning of these rights and to acknowledge their standing, because even the machine itself could not exist without human creativity. Even if AI copyrights were granted, it would not only be a precedent but a comprehensive legal regulation so that there could be no damage on the part of authors, such as natural persons, e.g., if it is the duration of the copyright or co-authorship, and there are also other copyrights that have never been created in such a way that they think of the possibility that the machine may also become the author of the work.

9.3 Trademark

A trademark can be understood as a symbol, a designation that is associated with an entity that manufactures goods or provides services. [32] In addition to their typical distinctive function, trademarks have a communication, advertising and investment function, which means that the trademark may be used by its proprietor to acquire or maintain a reputation capable of attracting or committing consumers. .[33]

Artificial intelligence in relation to trademarks is changing the way consumers view products and brands, e.g., if you use AI Amazon Alexa, which will recommend three products to the consumer in order to buy the product. The consumer does not have the space to get acquainted with other products on the market and therefore indirectly leads you to choose one of the three

[31] For more information on blockchain see BARULLI, M.: IP is a journey: blockchain and encrypted storage are your best friends. WIPO MAGAZINE, February 2021. Online: https://www.wipo.int/wipo_magazine_digital/en/2021/article_0002.html (quoted 5.12.2021).

[32] LAZÁR J. et. al.: *Občianske právo hmotné 2,* Bratislava: IURIS LIBRI, 2018, p. 619, ISBN 978-80-89635-35-1.

[33] ADAMOVÁ, Z.: *Právo duševného vlastníctva*, Bratislava: TINCT, 2020, p. 73, ISBN 978-80-973544-0-4.

brands offered. AI in this case serves as a filter between brands and we can say that in a way it also acts as an "influencer". Since it offers only 3 options, one of which is a product from Amazon. The choice of only three options greatly influences the consumer's decision. If AI were to influence decisions between purchases of products, it is appropriate to consider whether, in this case, AI should follow the regulatory framework for influenza marketing, all the more so if it can benefit financially from the sale of a particular brand. AI can affect the way a consumer buys, which is related to the trademark. .[34]

What if AI recommends a product that is counterfeit or infringes a trademark and makes a decision to buy it? This situation would occur e.g., if we allow AI to choose for us with a purchase permit. Would AI be secondarily responsible for the wrong decision of the consumer? The answer is uncertain, but AI can, on the other hand, help identify and alert you to counterfeits. [35] The trademark creates a certain emotional connection with the consumer, which breaks the moment the choice remains on AI. Will the trademark have the same meaning in the future as it does today? Will the average consumer be the same as an artificial consumer? Can an artificial consumer really think like a person and realize the value of a brand? [36]

Conclusion

In order for Europe to become a world leader in AI, it is also necessary to create a favorable environment, especially in the field of intellectual property law. The guarantees provided by EU Patent Law should be important with regard to support for the protection of innovation, and we consider it wrong to recognize AI as the sole inventor of the current legislation, but we also want to point out the need for new legislation in this area, which would also add some credit to AI, as it can make a significant contribution to innovation in the future that might be unattainable for man without his help. With this is-

[34] CURTIS, L., PLATTS, R.: Trademark Law Playing Catch-up with Artificial Intelligence? WIPO MAGAZINE, June 2020. Online: https://www.wipo.int/wipo_magazine_digital/en/2020/article_0001.html (quoted 9.11.2021).

[35] CASEY, K.: Artificial Intelligence in the Trademark World. In: *Stradley Ronon,* Fall 2020. Online: https://www.stradley.com/-/media/files/publications/2020/10/ip-appeal-fall-2020.pdf (quoted 9.11.2021).

[36] REVALLA, K.: Intelligent Trademarks: Is Artificial Intelligence Colliding with Trademark Law? In: *IUP Law Review*, vol. 8, no. 4, 2018, p. 13 – 20, ISSN 22313095. Online: https://search.ebscohost.com/login.aspx?direct=true&db=lgs&an=133637760&scope=site (quoted 9.11.2021).

sue, it must be emphasized that it is also necessary to keep in mind that artificial intelligence should not serve as an option that would *de facto* provide for the avoidance of responsibility. Every AI needs to be programmed, and there is no need to forget the people who should be responsible for such programming.

With regard to copyright, we have assessed that it is necessary to evaluate them more strictly, because they were created precisely in order to protect man and his ability to be creative, while human and not machine creativity should come to the fore. We also came to the conclusion that the works that AI uses for its learning should be registered in some way, e.g., through the blockchain system so that a reasonable reward can be obtained for the author who provided his work.

Attention in AI is focused mainly on patents and copyright, and that is why at the end of the article we draw attention to others such as. trademark law, which may be significantly affected in the future by AI's presence on the market. Although the machine can recognize our personality perfectly, this does not mean that in certain situations we cannot decide otherwise than we would normally do.

In the end, we will just emphasize that even though we are tempted to look at AI in a good light, we still need to maintain critical thinking, especially given that this technology can already evolve on its own and therefore it is good to determine its certain limits before damages can occur, not only specifically in the field of intellectual property law but also in general.

Bibliography

1. ADAMOVÁ, Z.: *Právo duševného vlastníctva*. Bratislava: TINCT, 2020, ISBN 978-80-973544-0-4.
2. AGARWAL, R.: Why Low-Skilled Workers Will Win In The Robot Revolution. 16. January 2019. Online: https://www.forbes.com/sites/washingtonbytes/2019/01/16/why-low-skilled-workers-will-win-in-the-robot-revolution/?sh=7842bda37538 (quoted 9.11.2021).
3. BARR, A., FEIGENBAUM, E.: *The handbook of artificial intelligence*. 1981, ISBN 978-0-86576-089-9. Online: https://doi.org/10.1016/B978-0-86576-089-9.50013-2 (quoted 9.11.2021).
4. BARULLI, M.: IP is a journey: blockchain and encrypted storage are your best friends. WIPO MAGAZINE, February 2021. Online: https://www.wipo.int/wipo_magazine_digital/en/2021/article_0002.html (quoted 5.12.2021).

5. CASEY, K.: Artificial Intelligence in the Trademark World. In: *Stradley Ronon,* Fall 2020. Online: https://www.stradley.com/-/media/files/publications/2020/10/ip-appeal-fall-2020.pdf (quoted 9.11.2021).
6. CJEU, Judgment of 13.11.2018, *Levola Hengelo BV gegen Smilde Foods BV,* C-310/17.
7. CJEU, Judgment of 2.5.2012, *SAS Institute Inc. v. World Programming Ltd*, C-406/10.
8. CJEU, Judgment of 22.12.2010, *Bezpečnostní softwarová asociace – Svaz softwarové ochrany gegen Ministerstvo kultury*, C-393/09.
9. CJEU, Judgment of 4.10.2011, *FAPL Ltd and others v. QC Leisure and others,* C-403/08.
10. CJEU, Judgment of 7.3.2013, *Eva-Maria Painer v. Standard Verlags GmbH and others,* C-145/10.
11. Companies and Intellectual Property Commission: Patent Journal, Part II of II, vol. 54, no. 7. Online: https://iponline.cipc.co.za/Publications/PublishedJournals/E_Journal_July%202021%20Part%202.pdf (quoted 5.12.2021).
12. CURTIS, L., PLATTS, R.: Trademark Law Playing Catch-up with Artificial Intelligence?. WIPO MAGAZINE, June 2020. Online: https://www.wipo.int/wipo_magazine_digital/en/2020/article_0001.html (quoted 9.11.2021).
13. DOBREV, D.: A Definition of Artificial Intelligence. 19. January 2004. Online: https://arxiv.org/pdf/1210.1568.pdf (quoted 9.11.2021).
14. EPO: EPO publishes grounds for its decision to refuse two patent applications naming a machine as inventor. 28.1.2020. Online: https://www.epo.org/news-events/news/2020/20200128.html (quoted 9.11.2021).
15. EUIPO and EPO: Odvetvia intenzívne využívajúce PDV a ekonomická výkonnosť v Európskej únii. Správa o analýze na úrovni odvetví, september 2019. Online: https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observatory/documents/IPContributionStudy/IPR-intensive_industries_and_economicin_EU/summary/IP_Contribution_Report_092019_execsum_sk.pdf (quoted 5.12.2021).
16. European Parliament News: AI rules: what the European Parliament wants. 21.10. 2020. Online: https://www.europarl.europa.eu/news/sk/headlines/society/20201015STO89417/formovanie-europskej-legislativy-v-oblasti-umelej-inteligencie (quoted 9.11.2021).
17. European Parliament News: Parliament leads the way on first set of EU rules for Artificial Intelligence. 20. October 2020. Online: https://www.europarl.europa.eu/news/en/press-room/20201016IPR89544/parliament-leads-the-way-on-first-set-of-eu-rules-for-artificial-intelligence (quoted 9.11.2021).
18. European Patent Office. European patent applications. Section „Top tech. fields". Online: https://www.statista.com/statistics/1032627/worldwide-machine-learning-and-ai-patent-owners-trend/ (quoted 5.12.2021).

19. GESLEY, J.: Legal and Ethical Framework for AI in Europa: Summary of Remarks. In: *Proceedings of the ASIL Annual Meeting*, vol. 114, 2020, ISSN 0272-5037. Online: https://doi.org/10.1017/amp.2021.46 (quoted 9.11.2021).
20. GUADAMUZ, A.: Artificial intelligence and copyright. WIPO MAGAZINE, October 2017. Online: https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html (quoted 5.12.2021).
21. LAZÁR J. et. al.: *Občianske právo hmotné 2.* Bratislava: IURIS LIBRI, 2018, ISBN 978-80-89635-35-1.
22. Meyer-Dulheuer Patent- und Rechtsanwaltskanzlei: Artificial Intelligence creates art – patentable? 30.10.2018. Online: https://legal-patent.com/international-intellectual-property/artificial-intelligence-creates-art-patentable/ (quoted 9.11.2021).
23. Meyer-Dulheuer Patent- und Rechtsanwaltskanzlei: Beethoven's 10th AI Symphony. 10.10.2021. Online: https://legal-patent.com/international-intellectual-property/beethovens-10th-ai-symphony/ (quoted 9.11.2021).
24. Nález Ústavného súdu Slovenskej republiky, sp. zn. III. ÚS 651/2016 z 28. novembra 2017.
25. REVALLA, K.: Intelligent Trademarks: Is Artificial Intelligence Colliding with Trademark Law? In: *IUP Law Review*, vol. 8, no. 4, 2018, p. 13, ISSN 22313095. Online: https://search.ebscohost.com/login.aspx?direct=true&db=lgs&an=133637760&scope=site (quoted 9.11.2021).
26. ŠVIDROŇ, J. (ed.): *Právo duševného vlastníctva v informačnej spoločnosti a v systéme práva*. Bratislava: VEDA, 2009, ISBN 978-80-224-1033-5.
27. TELEC, I., TUMA, P.: *Autorský zákon: Veľký komentár*. 2. vydání, Praha: C. H. Beck, 2019, ISBN 978-80-7400-748-4.
28. *Thaler v Commissioner of Patents* [2021] FCA 879, VID 108 of 2021. See more details: https://www.judgments.fedcourt.gov.au/judgments/Judgments/fca/single/2021/2021fca0879 (quoted 5.12.2021).
29. *Thaler v Comptroller General of Patents Trademarks And Designs* [2021] EWCA Civ 1374 (21 September 2021). Online: https://www.bailii.org/ew/cases/EWCA/Civ/2021/1374.html (quoted 5.12.2021).
30. *Thaler v. Hirshfeld,* [2021] US District Court for the Eastern District of Virginia, 1:20-cv-903 (LMB/TCB). See more details: https://www.dwt.com/-/media/files/blogs/artificial-intelligence-law-advisor/2021/09/thaler-v-hirshfeld-decision.pdf (quoted 5.12. 2021).
31. Uznesenie Ústavného súdu Slovenskej republiky, sp.zn. II. ÚS 647/2014 z 30. septembra 2014.
32. VOJČÍK, P. et. al.: *Právo duševného vlastníctva*. 2. vydání. Plzeň: Aleš Čeněk, 2014, ISBN 978-80-7380-527-2.
33. Zákon č. 185/2015 Z. z. Autorský zákon.

34. Zákon č. 435/1990 Zb. o patentoch, dodatkových ochranných osvedčeniach a o zmene a doplnení niektorých zákonov.

Author details

Mgr. Renáta ROMANOVÁ
Pavol Jozef Šafárik University in Košice
Faculty of Law
Department of Civil Law
renata.romanova@student.upjs.sk

10 ARTIFICIAL INTELLIGENCE AS A CHALLENGE TO COPYRIGHT OF THE NEW AGE

Lukáš Macko

Abstract

The paper focuses primarily on current issues in the field of copyright with regard to the need to modify the works of artificial intelligence. It is conceived as an intersection between the practical and theoretical level in connection with the comparison of the results of the activity of artificial intelligence and the very mental activity of a person to whom copyright is granted. At the same time, it focuses on the current available regulation at the national as well as at the community level. Last but not least, the aim of the paper will be to summarize the current challenges related to the development of artificial intelligence in the field of copyright.[1]

Introduction

It should be emphasized at the outset that the age of the Internet, as well as new technological approaches to the development of artificial intelligence, have played a very important role in the development of copyright itself. The development of artificial intelligence has gained almost a hundred times the strength of the previous millennium in recent years. This is not only due to the daily use of technological advances in the professional sphere, but also to the increasingly popular use of the Internet for entertainment and education. It must be said that the present is very different from the beginnings of the first advances of technology, which counted only on the results of human intellectual creativity and is beginning to have strong competition in technological processes based on cognitive algorithms of artificial intelligence, which is itself capable of creating copyrighted work.

When processing the article, we will try to define the term artificial intelligence. Furthermore, our task will be to use the methods of analysis and observation to compare the scope of the rights granted by the Community legislation associated with the protection of works of artificial intelligence in comparison with ordinary creations of human creative intellectual activity.

[1] This paper was prepared within the framework of the research project VEGA no. 1/0643/20 "Legal and Ethical Perspective of Artificial Intelligence".

Within the concept of the article, it is necessary to deal with the concept of responsibility for the use of such works by third parties, while it is important to answer the question of who will benefit from works created by artificial intelligence.

Last but not least, the role of this work will be to solve the question of how to approach the protection of artificial intelligence in the future and to provide views on, for example, ensuring the protection of works.

10.1 The concept of artificial intelligence versus human creativity

It may be said at the outset that, nowadays, artificial intelligence no longer avoids creative activity, which is inextricably linked to copyright as a set of standards that protect this creative activity. However, how to deal with the fact that the concept of creative activity is no longer necessarily connected only with a man.

Not so long ago, the world was told that the artificial intelligence, which people imagine only as an intangible entity operating in computer programs, which is incapable of creative activity, refutes the dogma that this society has adopted in creating its first images. The artificial intelligence algorithm, based on a pattern of recognition created by French artists called Obvious, was able to paint a picture of a man that auctioned off as the first artificial creation of artificial intelligence.[2] Another group of Engineered Arts enthusiasts managed to claim that the Ai-da robot, as a work of artificial intelligence, does not create its work of art using artificial algorithmic machine code, but as an entity using individual art recognition.[3]

The concept of artificial intelligence is not unknown in the modern world. This concept appears in the literature as such in the late 1930s. Alan Turing was one of the pioneers in technology who first tried to distinguish between the creative output of artificial intelligence and man himself. In the field of copyright, his test for works of art is well known and used, which focuses on comparing the ability of artificial intelligence to behave like a human be-

[2] ŠVEC, T.: Umelá inteligencia maľuje obrazy. Jej dielo vydražili za státisíce dolárov. Pravda, 26.10.2018. Online: https://vat.pravda.sk/technologie/clanok/489525-umela-inteligencia-maluje-obrazy-jej-dielo-sa-vydrazilo-za-statisice-dolarov/ (quoted 07.11.2021).

[3] BIELIKOVÁ, V.: Umenie tvorené umelou inteligenciou: Roboti maľujú a ich obrazy sa predávajú. techbyte.sk, 13.02.2019. Online: https://www.techbyte.sk/2019/02/umenie-roboti-maluju-obrazy-predavaju/ (quoted 07.11.2021).

ing. Defining the basic purpose of this test, it can be concluded that in the last century, the general society has dealt with the question of what distinguishes the human being from artificial intelligence. Although the answer to this question seems easy, in today's complex times of information technology, it is relatively difficult to determine whether a particular piece of music or a work of art can really be the result of the creative activity of artificial intelligence.[4]

Another pioneer in this field, whose menu cannot be overlooked, was the American computer scientist and scientist John McCarthy, who made the concept of artificial intelligence as we know it today and addressed the question of its independence of artificial intelligence from the human one, which is attributed to creative thinking. It cannot be denied that the shift in understanding the concept of artificial intelligence has flourished in recent years, but it is still a relatively debated issue whether this intelligence is capable of being a separate entity to which personal and property rights to a work can be attributed as it is in the case of people.

The term artificial intelligence can be heard in the current understanding in connection with machine learning based on the generation of certain outputs according to the setting of the artificial intelligence pattern. By this model is meant a pre-programmed algorithm of behavior of a certain technology, which in the use of data can, with its functionality and aesthetic result of the activity, fall into the same category of creation as man himself.[5] However, if we look at the importance of machine learning, which is also associated with artificial intelligence, it is based on the conditioned logical operations contained in the rules determined by man. That is, it is based on the creativity of the human being. In contrast, the current wave of artificial intelligence is based on reading and encoding data perceived from the human environment, not from data. Such algorithms, pre-programmed for the perception of reality as a whole, do not rely on data, but can independently create additional experiences from their basic data, which are no longer dependent on human creativity. In connection with this topic, the question then arises as to whether the definition of artificial intelligence is sufficient with regard to the ability of technology to adapt to situations without human intervention.

[4] COECKELBERGH, M.: Can Machine Create Art? In: *Philosophy & Technology*, 2017, p. 285–288. ISSN 2210-5433. Online: https://link.springer.com/article/10.1007/s13347-016-0231-5 (quoted 16.01.2021).

[5] Cognilytica: The Seven Patterns of AI, 04.04.2019. Online: https://www.cognilytica.com/2019/04/04/the-seven-patterns-of-ai/ (quoted 31. 01. 2021).

Taking into account modern artificial intelligence as a technology, it can be understood as an attempt to imitate natural or human intelligence that can learn, perceive, process, compose, make decisions and provide outputs that can be characterized as a work or invention. Thus, artificial intelligence is a system that exhibits intelligent behavior by analyzing the environment and taking steps with a degree of autonomy to achieve specific goals.

The result of these artificial intelligence outputs is a process whose output is an output that, however, could easily be characterized as creative. It is very likely that it would not be possible to distinguish whether a musical work is the result of human creation or artificial intelligence. But can a machine create art? From the point of view of theorists, the question of creativity at the interface of artificial intelligence is thus still unresolved. Given the ability of machines to learn, it can be said that even if the initial creativity is up to the programmer, the other outputs of the machine, respectively. however, technologies will no longer emerge independently of the will of the individual.

The answer to whether a machine can also create art is then, first and foremost, the answer to the question of how a machine "creates" such a work. Subsequently, such a work must be analyzed whether the machine to create it requires creativity, which requires imagination. Opinions on whether a machine can be full of imagination vary. One side argues that this feature is characteristic of man only, others express the view that the creation of a work by machines came without prior human programming, one can speak of creativity. Furthermore, it is necessary to focus on whether the creations of artificial intelligence are not just imitations of the outside world.[6] If the machine is able to analyze its environment, evaluate them and take separate steps to achieve the goals, we can only talk about imitations. However, this includes the concept of creativity. Even man can imitate another work within his creativity, using creativity based on being itself. Unlike machines, such creative activity of man, unlike works of artificial intelligence, has the status of a work of art under copyright protection. From a philosophical point of view, the answer as to whether a machine can have its interior independent of the environment and human intervention is ambiguous.

[6] COECKELBERGH, M.: Can Machine Create Art? In: *Philosophy Technology*, 2017, p. 285–288. ISSN 2210-5433. Online: https://link.springer.com/article/10.1007/s13347-016-0231-5 (quoted 08.11.2021).

10.2 Artificial intelligence and copyright

The nature of copyright as such counts on a model based on aesthetic creativity and the mental abilities of man. Only a human can be the author of a work, as the work can only come from the human mind. Moreover, this presumption is the basis of the whole system of protection of moral rights and does not allow the possibility of the existence of an internal world of the machine, which abounds in artificial intelligence.

In addition, in copyright law, works as a result of creative intellectual activity are conceptually linked to the concept of originality. The legal concept of originality, although a very important prerequisite for the definition of a work of art in the legislation, is not defined. However, the concept of the case law of the Court of Justice as well as the creation of law at Community level focus on a human-centered approach that defines originality as the result of the author's intellectual creation.[7] Such an approach is, of course, applicable if the work is created using artificial intelligence, as long as the person himself participates in it. However, the question is the degree of involvement of a given person in a given output and whether it is enough for such a natural person to be credited with a work created by artificial intelligence if the elements of originality will show the procedure of an algorithm that created and works by evaluating and analyzing external perceptions. If a natural person only runs the algorithm, but is no longer responsible for the result, we do not think we can talk about the originality of the person, as he did not have control over the output and did not participate in its creation.

However, a major dilemma arises with works made exclusively by artificial intelligence. Despite the future, it must be said that advanced technological developments will allow such artificial works to be autonomous, independent of any human intervention. From the point of view of the adopted regulation and the conclusions of the European Parliament, it follows that there is a general consensus in society that it is not possible to provide legal protection for the creation of artificial intelligence, as only artificial intelligence cannot be granted legal personality.[8]These starting points support, in particular, the above-mentioned claims that artificial intelligence can perfectly mimic the external environment through analysis and synthesis of knowl-

[7] Court of Justice of the European Union (hereinafter as the "CJEU"), Judgment of 01.12.2011, *Eva-Maria Painer*, C-145/10, EU:C:2013:138, par. 94.

[8] European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence Technologies, (2020/2015(INI)), par. 13.

edge based on rules and principles, but it should be emphasized that it lacks other, less understood brain functions and have not yet become part of machine learning, such as inspiration, imagination, consciousness.[9]

In the opinion of the Court of Justice of the European Union, human creativity cannot be compared to random outputs of artificial intelligence, despite the predominance of its cognitive abilities in relation to humans. Even if we accept that a machine can create a work of art, the inner expression of such a work will not contain an emotional component that is undoubtedly associated with creation, and therefore artificial intelligence cannot achieve the same degree of protection. Such understanding is clearly based on moral and philosophical foundations, and a change in the position of artificial intelligence in the current copyright system would probably completely undermine the whole basis of moral rights.[10]

From the above conclusions, the output created by artificial intelligence should not be granted the same status as copyrighted works. However, as we will see, the outputs generated by artificial intelligence deserve some protection by introducing *sui generis* law. In essence, this idea is also supported by the Parliament itself, which favors the human element, in particular by stating that strengthening the creative process of generating artistic content can raise questions about the ownership of intellectual property rights. In this context, it considers that it would not be appropriate to give legal personality to artificial intelligence technologies, and points to the negative impact of such an option on stimulating human creators.[11] It also considers that technical creations generated by artificial intelligence technology must be protected under the legal framework of intellectual property rights in order to encourage investment in this form of creation and improve legal certainty for citizens, businesses and inventors, as they are still major users of artificial intelligence technologies. On the other hand, the regulation states that these works are not justified for copyright protection, in particular in order to respect the principle of originality, which applies to a natural person, and

[9] GUNKEL, D.: Special Section: Rethinking Art and Aesthetics in the Age of Creative Machines. In: *Philosophy & Technology*, 2017, p. 263–265, ISSN 2210-5433. Online: https://link.springer.com/article/10.1007/s13347-017-0281-3#article-info (quoted 09.11.2021).

[10] CJEU, Judgment of 01.12.2011, *Eva-Maria Painer*, C-145/10, EU:C:2013:138, par. 94; CJEU, Judgment of 07.08.2018, *Land Nordrhein- Westfalen κατά Dirk Renckhoff, C-161/17*, EU:C:2018:634, par. 14.

[11] European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence Technologies, (2020/2015(INI)), par. 13.

because the term "creative intellectual activity" refers to the author's personality.[12]

10.3 The need to protect "copyright" works of artificial intelligence

Although in the previous chapter we answered the question that intelligence alone cannot be granted legal personality and subsequently granted the status of an author whose artwork would be protected in the same way as human creations, we must unequivocally agree that they deserve special legal protection. Failure to provide protection could encourage not only unfair competition. If the creation of artificial intelligence as such did not provide sufficient protection, it could happen that the consumer himself could appropriate such a work by copying and incorporating an internal element without having to take a specific part in its creation. In our opinion, the need for an internal element in the form of a reflection of personality creates a problem that can be abused by people themselves, as a certain author could appropriate the work in this way, while exercising the minimum creativity required by copyright for protection.

The question of creativity, and especially of its representative in connection with the output of artificial intelligence, is of particular importance for determining the person who will enjoy economic advantages as the holder of the rights to such a work. Another specific issue will be to address the issue of liability for the unauthorized distribution and use of works of art. In the case of piracy or other illegal infringements of copyright, it will be difficult for both law and society to resolve the concept of who is actually entitled to claim liability for the unauthorized distribution of a work.

If we look at the circle of people who can participate in the creation of a work produced by artificial intelligence, we can define three categories of people who make a significant contribution to the process of operating an artificial intelligence system. First, it is the owner of the artificial intelligence system, which is a natural or legal person who invests in the functionality of the system. The second person will be a developer, a natural person, or a team of natural persons who conceptually create a system of artificial intelligence. The third article will be the user of the system, i.e., the person who enters the data/inputs and trains the system for reliable output, and this is the

[12] *Ibidem,* par. 15.

last person to intervene in the chain of final and autonomous operation of the artificial intelligence system.

If we have outlined the concept of people who can participate together or even exclusively in the output of artistic intelligence, we must once again cast our gaze on the element of creativity that includes mental creative activity. Probably the most well-known role will testify in the case of output in the form of a work of art to the user of the system himself, whose activity is closest to the creative output. It is the person who provides the training data and sets the goal. It is the person who controls the result if the artificial intelligence system is dependent on rules and training data. If artificial intelligence is dependent on training data throughout the production of the output, we can conclude that such a work of art is protected by copyright and the original copyright holder is the user, because the output was created by artificial intelligence based on machine learning. But what if the work is created by the autonomous operation of the system?

In our opinion, there is a need to re-distinguish between the user and the developer of the artificial intelligence system itself, as they can fulfill the role of both participants at the same time. The autonomous artistic output of artificial intelligence is likely to be determined by the fact that it has given artificial intelligence a figure capable of stimulating artificial intelligence to self-learning and autonomous activity that is no longer under human control. Thus, it may happen that in the concept in question we completely eliminate the requirement of system user participation, as artificial intelligence will be able to independently generate output already during its development, which may necessarily be associated with more difficult determination of the person's right to a certain output.[13]

Conclusion

Although no one denies the need for protection, most legal opinions only accept the possibility of copyright protection only on condition of the involvement of the human factor. Conversely, if the creative output is autonomous, unpredictable and does not involve human intervention, protection would not be allowed under copyright law. The introduction of a new specific sui generis law could thus provide the necessary specific legal protection for

[13] PAPADOPOULOU, A.: Creativity in crisis: are the creations of artificial intelligence worth protecting? In: JIPITEC, 2021. Online: https://www.jipitec.eu/issues/jipitec-12-3-2021/5352/#ftn.N101AA (quoted 09.11.2021).

these works, as well as strengthen investment in this area, which is essentially what the very concept of Community law envisages.

Efforts to support artificial intelligence in the field can be researched especially in the current period. Research into the potential of artificial intelligence technologies is being adopted here by adopting several legal acts. In this respect, unnecessary legal barriers need to be removed so as not to hamper the growth or innovation of the Union's emerging data economy. However, experts emphasize the crucial importance of balanced protection of intellectual property rights in relation to artificial intelligence technologies and the multidimensional nature of such protection in terms of ensuring a high level of protection of intellectual property rights.

From the perspective of not only the European Union, the Court of Justice but also society, it seems that the intellectual property rights associated with the development of artificial intelligence technologies should be different from the intellectual property rights potentially granted in the case of works generated by artificial intelligence. Indeed, it should be assumed that if artificial intelligence is used only as a tool to assist the author in the creative process, the current mental protection framework continues to apply. In the case of works created independently by artificial intelligence, they are not eligible for copyright protection, in particular in relation to the obligation of respect for the principle of originality and creativity which applies to a natural person, as the term "creative intellectual activity" applies only to personality of the author, which was presented in other parts of this article.

Last but not least, the issue of creating a new sui generis right is linked to the adoption of legislation aimed at a fully harmonized regulatory framework in the field of artificial intelligence technologies. It is therefore proposed that such a framework take the form of a regulation, not a directive, in order to avoid fragmentation of the European digital single market.[14]

In conclusion, we must evaluate the development of artificial intelligence as very beneficial. We will see what challenges the next decade of this century will bring and how the law itself will respond to the situation. It should be added, however, that if moral and ethical principles are taken into account, the creations of artistic intelligence cannot be placed on the same scale compared to man's creative activity, at least not yet.

[14] Processed according to European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence Technologies, (2020/2015(INI)).

Bibliography

1. BIELIKOVÁ, V.: Umenie tvorené umelou inteligenciou: Roboti maľujú a ich obrazy sa predávajú. techbyte.sk, 13.02.2019. Online: https://www.techbyte.sk/2019/02/umenie-roboti-maluju-obrazy-predavaju/.
2. COECKELBERGH, M.: Can Machine Create Art? In: Philosophy Technology, 2017, p. 285, ISSN 2210-5433. Online: https://link.springer.com/article/10.1007/s13347-016-0231-5.
3. Cognilytica: The Seven Patterns of AI, 04.04.2019. Online: https://www.cognilytica.com/2019/04/04/the-seven-patterns-of-ai/.
4. CJEU, Judgment of 01.12.2011, *Eva-Maria Painer*, C-145/10, ECLI:EU:C:2013:138.
5. CJEU, Judgment of 07.08.2018, *Land Nordrhein- Westfalen κατά Dirk Renckhoff*, C-161/17, EU:C:2018:634.
6. EBERS. M.: Liability For Artificial Intelligence And EU Consumer Law, In: JIPITEC, 2021. Online: https://www.jipitec.eu/issues/jipitec-12-2-2021/5289/#ftn.N100CC.
7. European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence Technologies, (2020/2015(INI)).
8. GUNKEL, D.: Special Section: Rethinking Art and Aesthetics in the Age of Creative Machines. In: *Philosophy & Technology*, p. 263, ISSN 2210-5433. Online: https://link.springer.com/article/10.1007/s13347-017-0281-3#article-info.
9. PAPADOPOULOU, A.: Creativity in crisis: are the creations of artificial intelligence worth protecting? In: JIPITEC, 2021. Online: https://www.jipitec.eu/issues/jipitec-12-3-2021/5352/#ftn.N101AA.
10. ŠVEC, T.: Umelá inteligencia maľuje obrazy. Jej dielo vydražili za státisíce dolárov. Pravda, 26.10.2018. Online: https://vat.pravda.sk/technologie/clanok/489525-umela-inteligencia-maluje-obrazy-jej-dielo-sa-vydrazilo-za-statisice-dolarov/.

Author details

Mgr. Lukáš MACKO
PhD Candidate
Pavol Jozef Šafárik University in Košice
Faculty of Law
Department of Civil Law
lukas.macko@student.upjs.sk

11 ARTIFICIAL INTELLIGENCE AND THE SLOVAK LAW OF OBLIGATIONS: NON-CONFORMING PERFORMANCE AND NON-CONTRACTUAL LIABILITY ARISING OUT OF DAMAGE CAUSED TO ANOTHER

Ľuboslav Sisák

Abstract

*The paper examines selected issues involving artificial intelligence (AI) in the Slovak law of obligations. Starting with non-conforming performance (defects liability) in a sales contract, we focus on the question of how the autonomy of AI should be treated in the assessment of whether it was defective at the time of handover. Proceeding to non-contractual (tortious, delictual) liability for damage caused by AI and presumably attributable to its principal, we first test the applicability of the special liability clauses under the Civil Code, particularly § 421a. Afterwards, we concentrate on potential problems in the assessment of illegality and fault under § 420 as the general liability clause of the Civil Code. Finally, we move to product liability law, where we answer whether an unembedded AI-software is a product under the EU Directive 85/374/EEC and the Slovak Product Liability Act.**

Introduction

The effort to mimic human behaviour via technology – so called artificial intelligence (hereinafter also as "AI") – could be one of humanity's defining undertakings of the last decade. Based on the current popularity of AI in numerous areas of life, coupled with constant technological progress one may easily expect an even swifter growth of AI than we have hitherto witnessed. In the private sphere, the proliferation of AI is particularly visible in the B2C market. Each year, devices utilizing AI (smart phones, smart homes, smart lawn mowers, smart washing machines, smart cars, etc.) are becoming more and more commonplace. Apart from this, AI is creeping into large scale busi-

* This paper was prepared within the framework of the research project VEGA no. 1/0643/20 „Legal and Ethical Perspective of Artificial Intelligence" at the Pavol Jozef Šafárik University in Košice, Slovakia.

ness and industrial production (B2B) with the goal to intensify the already advanced automatization.

The constant expansion of AI *inter alia* in the consumer area and commerce might be a challenge for contemporary continental private law, mainly for two reasons. Firstly, the foundation of private law is still inspired by chiefly three historical milestones – its first complex and sophisticated elaboration in Ancient Rome, reception of Roman law in the middle-ages, and the upsurge of rationality in the Enlightenment era – which heavily influenced the first great private law codifications in the 19th (France, Austria) and 20th (Germany, Switzerland) century. These, in turn, served as models for private law-making (not only) in the rest of Europe. Thus, one could legitimately ask whether the historical underpinning of private law hampers its application when dealing with modern inventions, e.g., artificial intelligence. Secondly, never before has private law had to deal with objects (tangible or intangible) capable of simulating human behaviour, as is the case with AI. It is, therefore, only fitting to study how does private law cope with AI as an object of various legal relations.

Against this background, from all the branches and institutes of private law, we dedicate this paper to selected topics of Slovak obligation law, namely to the breach of a sales contract in the form of non-conforming (defective) performance, and non-contractual (tortious) liability arising out of damage caused to another. Our goal is to expose potential weak points in the framework of the mentioned areas in scenarios involving AI and to present ideas as to their resolution. We stick strictly to the law as it is (*de lege lata*) with no ambition to contemplate on how it should be (*de lege ferenda*).

11.1 Artificial intelligence: a working definition

Many experts, be it from law, IT, or other discipline, have already made the effort to define AI.[1] This paper has neither the ambition to expand on these efforts nor to contribute to them. However, for the purposes of the following examinations, it is necessary to at least choose a working definition of AI from among those already existing. To this end, the definition provided by

[1] Summarized by, e.g., KAULARTZ, M., BRAEGELMANN, T.: Einführung. In: KAULARTZ, M., BRAEGELMANN, T. (hrsg.): *Rechtshandbuch Artificial Intelligence*, C. H. Beck, 2020, p. 2 et seq., ISBN 9783406746581; KONERTZ, R., SCHÖNHOF, R.: *Das technische Phänomen „Künstliche Intelligenz" im allgemeinen Zivilrecht*, Nomos, 2020, p. 30 et seq., ISBN 978-3-7489-1050-3.

the European Commission might serve us well: "Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals."[2] Based on this definition, we identify these key characteristics of AI: 1. it collects, analyses, and evaluates data from its environment, 2. it takes steps to achieve a certain goal, 3. it tries to simulate the decision-making process of a human, 4. it does all of the aforementioned with substantial autonomy.

11.2 Breach of a sales contract: non-conforming performance

Generally, non-conforming performance in contract law means that a debtor fulfilled his legal duties towards the creditor in the right time and place, but the performance has proven to be defective. Non-conforming performance is thus also called "defects liability". Translated into the terms of a sales contract, non-conforming performance occurs only on the seller's side by handing over the purchased item to the buyer with defects. In such a case, the subjective equivalence[3] of the contract is impaired and is meant to be restored by a set of remedies granted to the buyer by defects liability law.

In Slovak law, one may observe a fragmentation of defects liability law in a sales contract, largely due to the plurality of sales contract regimes, each having its own scope of application and a more or less autonomous defects liability law (2.1). Only after clarifying this issue, we can proceed to the examination of AI in the context of defects liability law applicable to individual sales contract regimes (2.2).

[2] COM (2018) 237 final, p. 1. Online: https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2018)237&lang=en (quoted 5.12.2021).

[3] The subjective equivalence principle means that each party (subjectively) perceives the value of its performance as equal to that of the other party. From recent Czech writing thereto see, e.g., KOLMAČKA, V.: Práva z vad při smíšeném darování. In: *Časopis pro právní vědu a praxi*, 2021, no. 3, 2021, p. 449 et seq., ISSN 1805-2789. Online: https://journals.muni.cz/cpvp/article/view/15322 (quoted 5.12.2021).

11.2.1 Non-conforming performance in individual sales contract regimes

The Slovak Civil Code[4] has a general regulation of liability for defective performance applicable to the whole law of obligations.[5] These rules may or may not apply for the following individual sales contract regimes. Firstly, there is a sales contract of a *lex generalis* nature,[6] meaning that it applies only if no other of the *lege specialis* regimes are applicable. The general sales contract has its own rules on defects liability.[7] The general regulation of liability for defective performance applies also to the general sales contract, but only if the law on the latter does not state otherwise. Secondly, there is a consumer sales contract,[8] which represents the implementation of the EU Directive on consumer sales and guarantees.[9] This set of rules applies to B2C sales contracts.[10] Consumer sales contract law also has special rules on defects liability.[11] The general rules on defects liability and/or those within the general sales contract law apply to a consumer sales contract only if the lastly mentioned does not state otherwise.[12] Thirdly, there is a commercial sales contract found in the Slovak Commercial Code.[13, 14] This sales contract has its own exhaustive regulation of defects liability[15] and, therefore, does not require any subsidiary application of the rules in the Civil Code. Fourthly, there is a contract on the international sale of goods found in the CISG,[16] which notoriously also has an autonomous defects liability law[17] precluding any subsidiary application of another source of law.

---

[4] Act No. 40/1964 Coll. Civil Code (hereinafter as the "Civil Code").
[5] §§ 499–510 of the Civil Code.
[6] §§ 612–627 of the Civil Code.
[7] §§ 596–600 of the Civil Code.
[8] §§ 612–627 of the Civil Code.
[9] Directive 1999/44/EC of the European Parliament and of the Council of 25 May 1999 on certain aspects of the sale of consumer goods and associated guarantees, OJ L 171/12, 7. 7. 1999.
[10] § 612 with reference to § 52 para. 1 of the Civil Code.
[11] §§ 619–627 of the Civil Code.
[12] § 612 of the Civil Code.
[13] Act No. 513/1991 Coll. Commercial Code (hereinafter as the "Commercial Code").
[14] §§ 409–475 of the Commercial Code.
[15] §§ 422–441 of the Commercial Code.
[16] United Nations Convention on Contracts for the International Sale of Goods, incorporated into the Slovak legal order by the Notification of the Federal Ministry of Foreign Affair No. 160/1991 Coll.
[17] Chapter II., Section III.; Chapter III., Section III., Chapter V. of the CISG.

11.2.2 AI and non-conforming performance in individual sales contract regimes

The pivotal feature and probably the main asset of AI is its operational autonomy. It is precisely AI's autonomy that happens to cause problems in defects liability law within such sales contract regimes where the seller's liability for defects is incurred only if the purchased item's lack of conformity existed *at the time of handover* to the buyer. Sales contract regimes embracing such a rule are the general sales contract under the Civil Code,[18] sales contract under the Commercial Code,[19] and CISG.[20] Let us assume that one of these regimes applies to the following case. A seller sells a robot with AI to a buyer. The robot stands out for its high level of autonomy, manifested by the ability to adapt to its environment and thus become more effective in fulfilling its task. After weeks of flawless performance, the robot suddenly malfunctions and thereby causes financial loss to the buyer, who wants to reach for a remedy granted by defects liability law. Is the seller liable for the defect?[21]

No particular issues arise with regard to the classification of the defect, as it is a rather straightforward case of lack of quality known to all abovementioned sales contract regimes.[22] The problematic part, however, starts while examining whether the defect existed at the time when the robot was handed over to the buyer, since this is the prerequisite for incurring liability of the seller. We identify two hypotheses.

Firstly, the robot's algorithm causing his malfunction was faulty from the very beginning (including at the time of handover), making the seller liable for defects. However, even such a constellation is not altogether a win for the buyer, for he is the one who bears the burden of proof as to the existence of the defect at the time of handover. Because of the immense technical complexity of AI, satisfying the burden of proof might be a herculean task for the buyer, even with the help of an external expert.

[18] § 499 of the Civil Code, as a rule from general defects liability law, since general sales law of the Civil Code does not have a special rule for this matter.

[19] § 425 para. 1 with reference to § 455 of the Commercial Code.

[20] Art. 36 para. 1 CISG.

[21] Abridged version of the case in IFSITS, C., MINIHOLD, A. M., ROUBIK, M.: *Haftungsfragen beim Einsatz künstlicher Intelligenz,* Linde Verlag, 2020, p. 33, ISBN 9783707342628.

[22] For the general sales contract – § 499 of the Civil Code; for the sales contract under the Commercial Code – § 422 para. 1 with reference to § 420 para. 1 of the Commercial Code; and under the CISG – Art. 35 para. 1 thereof.

Secondly, the robot's algorithm was accurate and the faulty behaviour is a result of his relatively autonomous learning, adaptation, and operation abilities, which eventually went wrong. This argumentation precludes the seller's liability for defects, which is why he would be the one pleading it to the court. The following question though stands: under the current *lege artis,* can AI fail to fulfil its task solely because of "its own fault"? While the opinions are not unanimous, the prevailing one seems to lean towards a positive answer.[23]

None of these hypotheses are advantageous to the buyer. The first, because proving the existence of the defect at the time of handover appears to be an overly arduous task for the buyer due to the advanced technology of AI. What is even worse: should the technological progress lead to the second hypothesis coming out on top more frequently, there will be a danger that AI, as an ever more common object of sales contracts, will by its nature preclude defects liability under contemporary law.

One might notice that from all the sales contracts known to Slovak law, the consumer sales contract is omitted in the list of regimes for which the aforementioned considerations are relevant. The reason is that the consumer sales contract is the only one which remarkably embraces a different rule as to the point in time when the defect must exist in order to incur liability. Namely, consumer sales law introduces a (!) statutory guarantee (meaning that defects liability is incurred if a defect emerges anytime during the guarantee period) for all things except those of swift decay (e.g., groceries with very short shelf life), and used things.[24] This means that the debate presented

[23] Thus, e.g., ZECH, H.: Künstliche Intelligenz und Haftungsfragen. In: *Zeitschrift für die gesamte Privatrechtswissenschaft,* no. 2, 2019, p. 200, ISSN 2363-4960; KIRN, S., MÜLLER-HENGSTENBERG, C. D.: *Rechtliche Risiken autonomer und vernetzter Systeme,* De Gruyter, 2016, p. 112 et seq., ISBN 978-3-11-043144-5; somewhere in between KONERTZ, R., SCHÖNHOF, R.: *Das technische Phänomen „Künstliche Intelligenz" im allgemeinen Zivilrecht,* p. 68 et seq.; in the negative REICHERTZ, J.: Algorithmen als autonome Akteure? SozBlog. Blog der Deutschen Gesellschaft für Soziologie (DGS), 24.2.2013. Online: https://blog.soziologie.de/2013/02/algorithmen-als-autonome-akteure/ (quoted 5.12.2021); HERBERGER, M.: „Künstliche Intelligenz" und Recht – Ein Orientierungsversuch. In: *Neue Juristische Wochenschrift,* 2018, p. 2827, ISSN 0341-1915.

[24] § 619 para. 2 of the Civil Code. By introducing a statutory guarantee in consumer sales contracts, the legislator went far beyond what was necessary to implement under Art. 3 and 5 of the EU Directive 1999/44/EC. Critically (not only) thereto CSACH, K.: Dezintegrácia slovenského právneho systému právom na ochranu spotrebiteľa. In: DOBROVIČOVÁ, G. (ed.): *Vplyv medzinárodného a európskeho práva na právny poriadok Slovenskej republiky. Zborník príspevkov,* Košice, 2007, p. 25, ISBN 9788089089673. Online: https://www.upjs.sk/public/media/1084/zbornik_1.pdf (quoted 5.12.2021).

in the second hypothesis becomes irrelevant, since statutory guarantee does not care whether the defect existed at the time of handover or if it emerged afterwards, as long as it happened within the guarantee period of two years[25] counted from the moment of handover.[26] Therefore, the seller in a consumer sales contract would be liable even for those defects that the AI "obtained" by its autonomous operation, if such a thing is possible at all.[27]

11.3 Non-contractual liability for damages

In this section, we point out and try to resolve selected issues in the law of non-contractual (delictual, tortious) liability for damages (tort law) caused by AI. From all the liability regimes known Slovak law, we focus on tort law found in the Civil Code, and product liability law addressed by a special act.

11.3.1 The Civil Code

Tort law in the Civil Code is divided into two basic regimes: liability under the general clause,[28] and several special clauses.[29] The general liability clause is used only if none of the special clauses are applicable (*lex specialis derogat legi generali*).

Special liability clauses

From among the special liability clauses in the Civil Code, only one deserves a closer examination as to its applicability to cases of damage caused by AI *ex delicto*, and that is § 421a of the Civil Code. This liability is incurred if damage is dealt to another by a device or another thing which the debtor used to perform his obligation. The Civil Code goes on to demonstratively specify the types of services to which the given liability applies (medical, social, veterinary, and other biological services). The liability under § 421a of the Civil Code is strict, meaning that the tortfeasor's fault is irrelevant and, in this case, with no ground for liberation.

Now, the key to answering whether damage caused by AI *ex delicto* could fall under § 421a of the Civil Code is the interpretation of the part where it

[25] As the basic guarantee period under § 620 para. 1 of the Civil Code.
[26] § 621 of the Civil Code.
[27] Thereto see footnote no. 23 and the associated main text.
[28] § 420 of the Civil Code.
[29] §§ 420a–437 of the Civil Code.

says "(…) which the debtor used to perform his *obligation*". Judging by the wording of the norm itself and the related case law or literature,[30] it is rather clear that "obligation" refers only to a contractual obligation. However, one might ask if the injured party must be the contractual creditor of the tortfeasor, or also a third person. For example: a company providing grass mowing services has a contract with the city, according to which the company mows the grass in city parks, for which the city pays money. The company uses mowers with AI that operate without human interference once they are turned on. While mowing, the mower injures a dog belonging to another person. Can this person claim damages under § 421a of the Civil Code, since the harm was caused by a thing which the company, as the city's debtor, used to perform its obligation? Although the norm leaves room for such speculation, it must be rejected. Such a reading could establish an *erga omnes* strict liability of persons fulfilling their contracts with the help of things, which would be way too harsh for the supposed tortfeasor and with no solid basis in the Civil Code.

11.3.1.1 General liability clause

Liability under the general clause is incurred after fulfilling four cumulative requirements: 1. illegality of the harmful action, 2. harm being caused, 3. causality between illegal action and harm, 4. fault of the injuring party. Involvement of AI in delictual liability for damages might bring about difficulties particularly when dealing with requirements 1. and 4.

Ad a) Illegality of the harmful action. Generally, there are mainly two approaches to illegality as a requirement for delictual liability for damages in Europe. Some countries, e.g., Germany or in principle also Switzerland, prefer the wrongfulness of the result theory, which means that every harm to an absolutely protected right is illegal, as long as no circumstance precluding liability is given.[31] Other countries favour the wrongfulness of action theory, which says that the harmful conduct must deviate from what the law expects, i. e., that one behaves as he should not, or he does not behave as he

[30] FEKETE, I.: *Občiansky zákonník – Veľký komentár (2. zväzok),* Eurokódex, 2015, commentary to § 421a, section 3., ISBN 9788081550409; ŠKÁROVÁ, M.: § 421a. In: ŠVESTKA, J., ŠPÁČIL, J., ŠKÁROVÁ, M., HULMÁK, M. et al.: *Občanský zákoník I, II*, 2. vydání, Praha: C. H. Beck, 2009, p. 1226 et seq., ISBN 978-80-7400-108-6; HUMENÍK, I.: § 421a. In: ŠTEVČEK, M., DULAK, A., BAJÁNKOVÁ, J. et al.: *Občiansky zákonník I*, 2. vydanie, C. H. Beck, 2019, p. 1494 et seq., ISBN 978-80-7400-770-5.

[31] FUCHS, M., PAUKER, W., BAUMGÄRTNER, A.: *Delikts- und Schadensersatzrecht,* 9. Auflage, Springer, 2017, p. 89–90, ISBN 978-3-662-52664-4.

should. Slovakia follows the wrongfulness of action theory,[32] alongside, e.g., Austria.[33]

A device utilizing AI and damaging another *ex delicto* typically harms either health (or, in the worst scenario, takes life) or leads to economic (property) losses. Respecting the wrongfulness of action theory, one must identify a specific rule which was violated by the harmful conduct of AI attributable to a subject of law (typically the owner or possessor, hereinafter also as the "principal").[34] Obviously, there are plenty of rules protecting *inter alia* life, health, and property in criminal law and administrative law and their violation gives rise to criminal responsibility or administrative responsibility respectfully. However, a civil court, dealing with a claim for damages and examining illegality as the first requirement for tortious liability, is forbidden to assess whether a crime or administrative offence was committed as a matter of a preliminary question.[35] Only if a final (non-appealable) decision of a criminal court or an administrative body found the offender guilty of a crime or administrative offence, must a civil court respect it in its proceedings on damages. Now, neither does every injury caused by AI amount to a crime (or administrative offence) nor is the injured person always interested in criminal or administrative repression just to get compensation. Quite the opposite, an injured person is normally completely satisfied with a civil law resolution of the dispute. Barred from using crimes or administrative offences as a foundation for the illegality test in damages proceedings and with no interest of the injured person to engage in criminal or administrative proceedings, a civil court could only conclude that the harmful conduct did not violate any rule and, as a result, rule out civil liability. Such an outcome is irrational, as it would practically strip civil proceedings on damages *ex delicto* in cases of health injury or pure economic loss from any use. Thus, the virtually undisputed case law holds that, in the absence of any oth-

[32] DULAK, A.: Záväzky zo spôsobenia škody a bezdôvodného obohatenia. In: LAZÁR, J. et al.: *Občianske právo hmotné. Záväzkové právo. Právo duševného vlastníctva,* Bratislava: Iuris Libri, 2018, p. 358–359, ISBN 9788089635351.

[33] APATHY, P., RIEDLER, A.: *Bürgerliches Recht. Band III. Schuldrecht Besonderer Teil,* 4. Auflage, SpringerWienNewYork, 2010, p. 166 et seq., ISBN 978-3-211-99426-9.

[34] We acknowledge that the attribution of AI's conduct to another subject is not a matter of course, but we do not address this issue here in more detail and presume that it is given.

[35] Via interpretation of § 193 of the Act No. 160/2015 Coll. Code on Civil Contentious Proceedings. BAJÁNKOVÁ, J., ŠTEVČEK, M.: § 194. In: ŠTEVČEK, M., FICOVÁ, S., BARICOVÁ, J. et al.: *Civilný sporový poriadok. Komentár*, Praha: C. H. Beck, 2016, p. 734 et seq., ISBN 978-80-7400-629-6.

er explicit rule, one must examine a potential violation of the general duty of prevention under § 415 of the Civil Code (catch-all clause, *Auffangklausel*).[36] In essence, this duty compels every subject of law to prevent or at least mitigate harm being done to life, health, and property, by maintaining a level of carefulness which is reasonable to expect under the circumstances of a given case. Yet, the duty of prevention does not require to anticipate every possible injury *pro futuro*.[37] Over time, a more specific case law has developed for individual case groups.[38]

Although the general prevention duty is far from unproblematic in plenty situations due to its inherent vagueness, it sparks a debate hitherto unheard of in relation to AI. The question stands what the required level of carefulness and anticipation is when using devices based on AI in order to prevent them from harming legally protected values of others. Probably the biggest challenge is to ascertain the degree of human supervision over AI required to satisfy the duty of prevention. There are two undesired interpretative extremities in this regard, one too strict and one too lenient. The strict reading of the prevention duty would set a high degree of supervision, forcing people to constantly watch over AI so that it does not cause damage to others and if it does, to avoid liability. Such an approach would discourage from using AI and consequently disturb the desired technological progress. The lenient reading would be based on the relatively high reliability of AI, allowing little to no supervision, meaning that the principal could not be blamed for the lack thereof, resulting in impunity for alleged torts. The key is to find a middle way between these two extremities. But, similarly to other case groups of the prevention duty, it is hard and probably even undesired to come up with an abstract standard of supervision over AI, as it would be naturally inflexible and thus detrimental for the assessment of some portion of cases. However, what could prove helpful is a demonstrative set of factors which should be taken into account while evaluating the compliance with the prevention duty in torts committed by AI and attributable to a principal: nature of the tasks performed by the AI and their potential to cause harm; how advanced the AI is; environment in which the AI operates; whether the principal re-

[36] Decision of the Supreme Court of the Czech Republic from 16th May 2002, no. 25 Cdo 1427/2001.

[37] Decision of the Supreme Court of the Czech Republic from 25th February 2003, no. 25 Cdo 618/2001.

[38] For more details thereto see FEKETE, I.: *Občiansky zákonník – Veľký komentár (2. zväzok),* Eurokódex, 2015, commentary to § 415, section 3.

spected the usage instructions of the manufacturer; quality of maintenance. If the injured party proves that, under these circumstances, the supervision was insufficient, the illegality requirement will be fulfilled due to a breach of the prevention duty.

Ad b) Fault of the tortfeasor. Slovak civil law does not have an autonomous definition of fault and relies on the perception of this legal category in criminal law. Descending from the most severe forms of fault to the mildest ones, Slovak criminal law distinguishes direct intent,[39] indirect intent,[40] conscious negligence,[41] and unconscious negligence.[42] The upcoming lines require to become familiar only with unconscious negligence, which, adapted for tort law purposes, has the following meaning: the tortfeasor did not know that his action could harm legally protected values of another but, taking into account the circumstances of the case and his person, he should have and could have known.[43]

In the general liability regime of the Civil Code, fault of the injuring party is presumed. Thus, it is not the duty of the injured party to prove the tortfeasor's fault but, in order to exculpate himself, the tortfeasor must prove that the damage caused was not his fault.[44] From all the forms of fault, always the mildest one – unconscious negligence – is presumed.[45]

Based on the abovementioned, if damage was caused by AI, the principal must prove that he did not act in unconscious negligence. When pondering over the question of how the principal must have acted for him not to be unconsciously negligent, we see the debate essentially the same as in the case of the prevention duty within the illegality test. Although phrased a bit differently, the duty of prevention, identically to unconscious negligence, is also about what the principal should have and could have done to avert damage, while (differently to unconscious negligence) paying no attention to *mens rea.* Though the content of both may be largely the same, the main difference can be seen in the subject carrying the burden of proof. In the illegality test, it is the injured party who must prove that the prevention duty was

[39] § 15 lit. a) of the Act No. 300/2005 Coll. Criminal Code (hereinafter as the "Criminal Code").

[40] § 15 lit. b) of the Criminal Code.

[41] § 16 lit. a) of the Criminal Code.

[42] § 16 lit. b) of the Criminal Code.

[43] *Ibidem*.

[44] § 420 para. 3 of the Civil Code.

[45] DULAK, A.: Záväzky zo spôsobenia škody a z bezdôvodného obohatenia. In: LAZÁR, J. et al.: *Občianske právo hmotné. Záväzkové právo. Právo duševného vlastníctva*, p. 371.

violated. In the fault requirement, the tortfeasor must prove that he was not even unconsciously negligent. But when it comes to legal issues caused by AI in both cases, we do not see a difference. Thus, the problems related to AI identified within the prevention duty and their suggested resolution accordingly apply to the debate on the tortfeasor's unconscious negligence, i.e., that it all comes down to the required degree of supervision over AI; that there should be no general standard; but that there are several factors influencing the assessment of an individual case. If the tortfeasor proves that his supervision was sufficient in the given circumstances, he will not be liable due to lack of fault.

11.3.2 Product liability

Slovak law on product liability is found in the Act No. 294/1999 Coll. (hereinafter as the "Product Liability Act"). As in other EU Members States, Slovak product liability law is modelled after the EU Directive 85/374/EEC.[46] The idea of product liability is to hold the producer accountable for the damage his defective product caused to its users. The producer's liability is strict (irrespective of fault) with the possibility of exoneration. The question we want to touch upon is whether AI could be understood as a "product" under the Product Liability Act. If answered in the positive, AI falls under the scope of the Product Liability Act.

Thanks to the harmonising EU Directive 85/374/EEC, product liability law is more or less the same in all EU Member States. Naturally, this applies also to the definition of a "product". For this reason, and because of sparse Slovak and Czech literature dealing with the examined issue, we take advantage of foreign scholarship which we consider applicable *mutatis mutandis* to Slovak law.

Starting with the Slovak definition of a product, the Product Liability Act opts for the following one: "A product under this Act is any movable thing which has been manufactured, mined, or otherwise obtained, irrespective of the stage of its processing, and which is intended to be put into circulation. A movable thing which is part of or an accessory to another movable or immovable thing shall also be a product under this Act. Electricity and gas which are intended for consumption shall also be deemed to be a prod-

[46] Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 210/29, 7.8.1985.

uct." Clearly, the central requirement of this definition is that a product must be a movable thing, except for electricity and gas. Now, neither the EU Directive 85/374/EEC nor the Product Liability Act define what is a "movable thing". But because the term originates in EU law, we hold that it should have an autonomous meaning, independent of national perceptions of things in private law.[47]

If a physical device is embedded with AI and the functionality of the former is dependent on the latter (e.g., service robots), it falls, for the most part, under the definition of a product as a movable thing pursuant to the Product Liability Act.[48] The same goes for AI-software sold on data carriers (e.g., CD, DVD, USB, external hard drives, etc.).[49] However, the answer is not entirely clear when AI exists only as a downloadable or online-used ("unembedded") software and the device serves only as a tool to run it (e.g., internet of things, apps for smart phones). In such instances, the AI-software is no longer tangible. The mere fact that an AI-software was downloaded to a (tangible) hard drive does not make any difference in the said observation, just like writing an idea on paper does not make the idea a tangible thing only because the paper is tangible.[50] The question thus arises whether an intangible thing in the form of an AI-software could be a "movable thing" under the EU Directive 85/374/EEC and the Product Liability Act. In our view, it should be answered in the positive. The argument is a combination of grammatical and teleological interpretation.

Grammatical interpretation. Article 2 of the EU Directive 85/374/EEC (hereinafter also as "directive") defines a product principally as a movable thing. The Product Liability Act does not go beyond that what must be im-

[47] Thus with regard to German law WAGNER, G. § 2 ProdHaftG, margin number 1. In: HABERSACK, M. (hrsg.): *Münchener Kommentar zum Bürgerlichen Gesetzbuch: BGB, Band 7: Schuldrecht Besonderer Teil IV,* 8. Auflage, C. H. Beck, 2020, ISBN 978-3-406-72607-1.

[48] OUCHSLER, J.: § 2 ProdHaftG, margin number 67. In: STAUDINGER, J. (hrsg.): *Kommentar zum Bürgerlichen Gesetzbuch, Buch 2: Recht der Schuldverhältnisse: §§ 826–829; ProdHaftG,* Sellier, De Gruyter, Stand 28.2.2020, ISBN 978-3-8059-1258-7. Problems arise when the producers of the hardware and AI-software are different subjects and the defect lies in the AI-software. As to the question of whether the producer of the AI-software could be liable under the Product Liability Act, the upcoming lines dedicated to unembedded software apply accordingly.

[49] WAGNER, G.: § 2 ProdHaftG, margin number 22. In: HABERSACK, M. (hrsg.): *Münchener Kommentar zum Bürgerlichen Gesetzbuch: BGB, Band 7: Schuldrecht Besonderer Teil IV.*

[50] Thus already ANDREEWITCH, M.: Zur Anwendbarkeit des Produkthaftungsgesetzes für Softwarefehler. In: *Medien und Recht*, 1990, p. 50 et seq.

plemented in this regard and follows the directive as to the requirement that a product must be, first of all, a movable thing. Therefore, neither the EU Directive 85/374/EEC nor the Product Liability Act declare that a product must be tangible.[51] If nothing else, this observation at least opens the door to the potential inclusion of intangible things, *inter alia* unembedded software, in the scope of product liability law.

Teleological interpretation. As per usual in secondary EU law, the objectives and purposes of the EU Directive 85/374/EEC are found in its preamble. It is here that the directive expresses one of its central beliefs that "liability without fault on the part of the producer is the sole means of adequately solving the problem, *peculiar to our age of increasing technicality*, *of a fair apportionment of the risks inherent in modern technological production*".[52] In the light of this idea, would it not be flawed if we left software out of the scope of product liability law, knowing that precisely software, above all else, are "peculiar to our age of increasing technicality" and are a result of "modern technological production"? Is it not desired to extend the strict product liability to software, which are no less potent to cause damage to consumers than tangible goods? We hold that both questions should be answered in the positive.

Taking all of the above into consideration, we count AI-software as products pursuant to the EU Directive 85/374/EEC and the Product Liability Act[53] (as the European Commission itself does without further a due[54]).

[51] Unlike, e.g., the Austrian implementation in § 4 of the *Produkthaftungsgesetz,* which, from the directive's perspective unnecessarily, defines a product as a movable *tangible* thing.

[52] EU Directive 85/374/EEC, Preamble, para. 2.

[53] With this conclusion and by using similar or even more extensive argumentation, e.g., WAGNER, G.: § 2 ProdHaftG, margin number 25 et seq. In: HABERSACK, M. (hrsg.): *Münchener Kommentar zum Bürgerlichen Gesetzbuch: BGB, Band 7: Schuldrecht Besonderer Teil IV*; KOZIOL, H.: Produkthaftung, margin number 136 et seq. In: KOZIOL, H., APATHY, P., KOCH, B. A.: *Österreichisches Haftpflichtsrecht Band III*, 3. Auflage, Jan Sramek Verlag, 2014, ISBN 978-3-7097-0022-8; DENKMAIER, CH.: 30 Jahre PHG – Software als Produkt? In: FELTE, E., KOFLER, G., MAYRHOFER, M. et al. (hrsg.): *Digitale Transformation im Wirtschafts- & Steuerrecht,* Wien: Linde Verlag, 2019, ISBN 978-3-7073-3995-6; REUSCH, P. Produkthaftung. In: KAULARTZ, M., BRAEGELMANN, T. (hrsg.): *Rechtshandbuch Artificial Intelligence*, p. 114; REBIN, I.: § 2 ProdHaftG, margin number 49 et seq. In: GSELL, B., KRÜGER, W., LORENZ, S., REYMANN, CH. (hrsg): *beck-online. GROSSKOMMENTAR zum Zivilrecht.* Stand 1.11.2020, C. H. Beck; however, in a more rejective fashion OUCHSLER, J.: § 2 ProdHaftG, margin number 64 et seq. In: STAUDINGER, J. (hrsg.): *Kommentar zum Bürgerlichen Gesetzbuch, Buch 2: Recht der Schuldverhältnisse: §§ 826-829.*

[54] Written question No 706/88 by Mr Gijs de Vries (LDR-NL) to the Commission of the European Communities and answer given by Lord Cockfield on behalf of the Commission, OJ

Conclusion

Attention was first paid to non-conforming performance (defects liability) within a contract on the sale of goods with AI. We identified issues in the assessment of whether the purchased item with AI was defective at the time of handover, which under most Slovak sales contract regimes must be answered in the affirmative in order to eventually incur the sellers' defects liability. The most important finding in this regard is that if the autonomy of AI allows it to behave largely independently and this leads to malfunction, the defect did not exist at the time of handover, thus precluding defects liability of the seller.

Secondly, we concentrated on non-contractual (delictual, tortious) liability for damages caused by AI and started with examining the Civil Code. We rejected the application of all special liability clauses, including § 421a of the Civil Code as the only *prima facie* candidate. Therefore, tortious liability for damage caused by AI attributable to the principal pursuant to the Civil Code can be incurred only by the general liability clause – § 420. Here, the involvement of AI leads to issues when dealing with the illegality and fault conditions. As to the first, the ground for illegality of the principal's conduct will be regularly the alleged breach of the prevention duty under § 415 of the Civil Code. However, due to the AI's autonomy, it is problematic to ascertain the degree of supervision required from the principal to satisfy the duty of prevention. Although we did not come up with a general standard of supervision, we put together a non-exhaustive list of factors which should be considered in the overall assessment of an individual case. Moving on to the requirement of the tortfeasor's fault, we found the issue essentially the same as in the illegality test. As a result, the conclusions made in relation to the latter accordingly apply to the question of whether the damage caused by AI was the principal's fault. Finally, we proceeded to products liability law and examined whether an unembedded AI-software is a "product" under EU Directive 85/374/EEC and the Product Liability Act. Based on grammatical and teleological interpretation of the directive, we answered this question in the affirmative.

No C 114/42, 8. 5. 1989. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:JOC_1989_114_R_0001_01&qid=1429892489522&from=EN (quoted 5.12.2021).

Bibliography

1. Act No. 300/2005 Coll. Criminal Code.
2. Act No. 40/1964 Coll. Civil Code.
3. Act No. 513/1991 Coll. Commercial Code.
4. ANDREEWITCH, M.: Zur Anwendbarkeit des Produkthaftungsgesetzes für Softwarefehler. In: *Medien und Recht*, 1990, p. 50.
5. APATHY, P., RIEDLER, A.: *Bürgerliches Recht. Band III. Schuldrecht Besonderer Teil,* 4. Auflage, SpringerWienNewYork, 2010, ISBN 978-3-211-99426-9.
6. Bundesgesetz vom 21. Jänner 1988 über die Haftung für ein fehlerhaftes Produkt (Produkthaftungsgesetz) (Austria).
7. COM (2018) 237 final. Online: https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2018)237&lang=en.
8. Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 210/29, 7.8.1985.
9. DOBROVIČOVÁ, G. (ed.): *Vplyv medzinárodného a európskeho práva na právny poriadok Slovenskej republiky. Zborník príspevkov*. Košice, 2007, ISBN 9788089089673. Online: https://www.upjs.sk/public/media/1084/zbornik_1.pdf.
10. Decision of the Supreme Court of the Czech Republic from 16th May 2002, no. 25 Cdo 1427/2001.
11. Decision of the Supreme Court of the Czech Republic from 25th February 2003, no. 25 Cdo 618/2001.
12. Directive 1999/44/EC of the European Parliament and of the Council of 25 May 1999 on certain aspects of the sale of consumer goods and associated guarantees, OJ L 171/12, 7. 7. 1999.
13. FEKETE, I.: *Občiansky zákonník – Veľký komentár (2. zväzok).* Eurokódex, 2015, ISBN 9788081550409.
14. FELTE, E., KOFLER, G., MAYRHOFER, M. et al. (hrsg.): *Digitale Transformation im Wirtschafts- & Steuerrecht,* Wien: Linde Verlag, 2019, ISBN 978-3-7073-3995-6.
15. FUCHS, M., PAUKER, W., BAUMGÄRTNER, A.: *Delikts- und Schadensersatzrecht. 9. Auflage, Springer,* 2017, ISBN 978-3-662-52664-4.
16. GSELL, B., KRÜGER, W., LORENZ, S., REYMANN, CH. (hrsg): *beck-online. GROSSKOMMENTAR zum Zivilrecht.* Stand 1.11.2020, C. H. Beck.
17. HABERSACK, M. (hrsg.): *Münchener Kommentar zum Bürgerlichen Gesetzbuch: BGB, Band 7: Schuldrecht Besonderer Teil IV*, 8. Auflage, C. H. Beck, 2020, ISBN 978-3-406-72607-1.
18. HERBERGER, M.: „Künstliche Intelligenz" und Recht – Ein Orientierungsversuch. In: *Neue Juristische Wochenschrift*, 2018, p. 2825, ISSN 0341-1915.

19. IFSITS, C., MINIHOLD, A. M., ROUBIK, M.: *Haftungsfragen beim Einsatz künstlicher Intelligenz.* Linde Verlag, 2020, ISBN 9783707342628.
20. KAULARTZ, M., BRAEGELMANN, T. (hrsg.): *Rechtshandbuch Artificial Intelligence*. C. H. Beck, 2020, ISBN 9783406746581.
21. KIRN, S., MÜLLER-HENGSTENBERG, C. D.: *Rechtliche Risiken autonomer und vernetzter Systeme*. De Gruyter, 2016, ISBN 978-3-11-043144-5;
22. KOLMAČKA, V.: Práva z vad při smíšeném darování. In: *Časopis pro právní vědu a praxi*, 2021, no. 3, 2021, p. 447, ISSN 1805-2789. Online: https://journals.muni.cz/cpvp/article/view/15322.
23. KONERTZ, R., SCHÖNHOF, R.: *Das technische Phänomen „Künstliche Intelligenz" im allgemeinen Zivilrecht*. Nomos, 2020, ISBN 978-3-7489-1050-3.
24. KOZIOL, H., APATHY, P., KOCH, B. A.: *Österreichisches Haftpflichtsrecht Band III*. 3. Auflage, Jan Sramek Verlag, 2014, ISBN 978-3-7097-0022-8.
25. LAZÁR, J. et al.: *Občianske právo hmotné. Záväzkové právo. Právo duševného vlastníctva*. Bratislava: Iuris Libri, 2018, ISBN 9788089635351.
26. REICHERTZ, J.: Algorithmen als autonome Akteure? SozBlog. Blog der Deutschen Gesellschaft für Soziologie (DGS), 24.2.2013. Online: https://blog.soziologie.de/2013/02/algorithmen-als-autonome-akteure/;
27. STAUDINGER, J. (hrsg.): *Kommentar zum Bürgerlichen Gesetzbuch, Buch 2: Recht der Schuldverhältnisse: §§ 826-829; ProdHaftG,* Sellier, De Gruyter, Stand 28.2.2020, ISBN 978-3-8059-1258-7.
28. ŠTEVČEK, M., DULAK, A., BAJÁNKOVÁ, J. et al.: *Občiansky zákonník I*. 2. vydanie, C. H. Beck, 2019, ISBN 978-80-7400-770-5.
29. ŠTEVČEK, M., FICOVÁ, S., BARICOVÁ, J. et al.: *Civilný sporový poriadok. Komentár*. Praha: C. H. Beck, 2016, ISBN 978-80-7400-629-6.
30. ŠVESTKA, J., ŠPÁČIL, J., ŠKÁROVÁ, M., HULMÁK, M. et al.: *Občanský zákoník I, II*. 2. vydání, Praha: C. H. Beck, 2009, ISBN 978-80-7400-108-6.
31. United Nations Convention on Contracts for the International Sale of Goods, incorporated into the Slovak legal order by the Notification of the Federal Ministry of Foreign Affair No. 160/1991 Coll.
32. Written question No 706/88 by Mr Gijs de Vries (LDR-NL) to the Commission of the European Communities and answer given by Lord Cockfield on behalf of the Commission, OJ No C 114/42, 8. 5. 1989. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:JOC_1989_114_R_0001_01&qid=1429892489522&from=EN.
33. ZECH, H.: Künstliche Intelligenz und Haftungsfragen. In: *Zeitschrift für die gesamte Privatrechtswissenschaft*, no. 2, 2019, ISSN 2363-4960.

Author details

JUDr. Ľuboslav SISÁK
PhD Candidate
Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of International Law and European Law
luboslav.sisak@student.upjs.sk

12 LIABILITY FOR DAMAGE CAUSED BY AUTONOMOUS VEHICLES: THE NEWS IN THE EU IN 2021

Ľubica Gregová Širicová

Abstract

*The attention of the author is focused on two legislative acts related to liability for damage caused by AVs that have recently been adopted in the EU. The first part of the article introduces the new German act on autonomous driving with its unique concept of the "technical supervisor", a natural person responsible to ensure compliance with road traffic law. This concept is going to be compared with the solutions proposed by the European Parliament in its draft Regulation on liability for the operation of artificial intelligence-systems. Several weak places of the draft regulation have to be noted: it is criticized for creating a parallel application of the regulation and the national rules that is potencially too complicated and lacks persuasive explanation on the distinction between the conventional versus autonomous vehicles. Moreover, the draft would need to be revisited to rethink its relationship with the rules on conflict of laws.****

Introduction

The benefits of autonomous vehicles (AVs) are generally acknowledged: decreasing the high number of traffic accidents caused by human factor,[1] increasing efficiency, sustainability and flexibility in transport, a tremendous potential for persons with disabilities and reduced mobility etc. The self-driving vehicles' market is expected to grow exponentially creating new jobs and developing profits of up to €620 billion by 2025 for the EU automotive industry.[2]

* This paper was prepared within the framework of the research project VEGA no. 1/0643/20 "Legal and Ethical Perspective of Artificial Intelligence".

[1] *"Automated vehicles have the potential to make a huge contribution to reducing road fatalities, given that more than 90 % of road accidents are estimated to result from some level of human error."* – recital no. 23, Regulation (EU) 2019/2144 of the European Parliament and of the Council of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users, OJ L 325, 16.12.2019, p. 1–40.

[2] European Parliament News: Self-driving cars in the EU: from science fiction to reality. 14.1.2019. Online: https://www.europarl.europa.eu/news/en/headlines/economy/2019

At the same time, there is concern over risks that the AVs may pose to the society: risks in road safety, security of personal data in the event of their massive collection, driver job losses etc. Moreover, ethical issues are likely to appear[3] (what preconditions should guide the decision that a car must make before a potentially fatal accident?).

In order to establish the AVs as accepted means of transport, it is inevitable that the risks of damage that may occur due to the operation of AVs are regulated by appropriate regimes of liability.

The current legal framework was developed to deal with the liability issues concerning the operation of the AVs. If the legislator does not duly address the above issues, the current regulatory framework will result in many uncertainties and it is likely that the cost of the risks will be borne by the injured parties and consumers will find it increasingly difficult to claim damages. This could ultimately lead to reduced consumer confidence in AVs and, consequently, to slower uptake of AVs in the market.[4]

It has to be noted, that in order to operate autonomously, the vehicles need to use artificial intelligence (AI). With self-driving, the driving tasks are shifted from the human driver to the AI. The opacity ("black box" element) of the AI makes it extremely expensive or even impossible to identify which code or input caused the harmful operation. Therefore, the existing liability regimes need to evolve and clarify who is accountable in case of accident: the driver, the manufacturer or some other person?

The following general approach to the concept of liability can be formulated: each obligation should rest on the actor who is best placed to address the risk. However, the specification of this actor is going to vary according to the level of automation incorporated in the AV. In order to address the issue of liability, a clear distinction between the various forms of AVs has to be established.

0110STO23102/self-driving-cars-in-the-eu-from-science-fiction-to-reality (quoted 8.12. 2021).

[3] About ethical guidelines for AVs see: BARTNECK, C., LÜTGE, C., WAGNER, A., WELSH, S.: *An Introduction to Ethics in Robotics and AI,* SpringerLink, 2021, p. 86–89, ISBN 978-3-030-51110-4. Online: https://link.springer.com/book/10.1007/978-3-030-51110-4 (quoted 8.12.2021).

[4] A common EU approach to liability rules and insurance for connected and autonomous vehicles. European Added Value Assessment Accompanying the European Parliament's legislative own-initiative report (Rapporteur: Mady Delvaux), p. 5.

The **SAE Levels of Driving Automation™** is the most-cited reference for the AV standard of capabilities.[5] The Society of Automotive Engineers (SAE) has created the latest version in partnership with the International Organization for Standardization (ISO). It defines six levels of driving automation, from SAE Level Zero (no automation) to SAE Level 5 (full vehicle autonomy):

Level 0 – No Driving Automation
Level 1 – Driver Assistance
Level 2 – Partial Driving Automation
Level 3 – Conditional Driving Automation
Level 4 – High Driving Automation
Level 5 – Full Driving Automation

To sum up the characteristics of the levels: When the AV operates in Level 1, the driver and the automated system share control of the vehicle and the driver must be ready to retake full control at any time. Level 2: the automated system takes full control of the vehicle: accelerating, braking, and steering, however, the driver must monitor the driving and be prepared to intervene immediately at any time if the automated system fails to respond properly. Level 3: the drivers can safely turn their attention away from the driving tasks; however, the driver must still be prepared to intervene within some limited time. Level 4: no driver attention is ever required for safety, however, self-driving is supported only in limited spatial areas (geofenced) or under special circumstances (outside of these areas or circumstances, the vehicle must be able to safely abort the trip, e.g., slow down and park the car, if the driver does not retake control). Level 5: no human intervention is required at all (all kinds of surfaces, in all weather conditions).

It will be argued below that the problem of the liability is directly related to the individual levels of automation. There is another key point concerning liability caused by the AVs operation: the human in the AV has to be fully aware of the particular level of automation deployed in the given situation (and of its consequences for his liability). Some technological names that are currently used, such as AutonoDrive, PilotAssist, Full-Self Driving or DrivePilot, may confuse a driver who may think that his action is never expected

[5] SAE International: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, J3016_202104. 30.4.2021. Online: https://www.sae.org/standards/content/j3016_202104/ (quoted 8.12.2021).

(while the driver must still be involved in driving). This may result in overconfidence and lead to accidents.[6]

This article will be focused on two legislative acts related to liability for damage caused by AVs. Currently, a battle of AVs regulation takes place among the national legal systems (e.g., Germany, Great Britain, France) to become a leader in AV industry. It appears that, at the moment, as far as the regulation is concerned, Germany has won.[7] The first part of the article will introduce the new German Act on AV. However, the German legislator has envisaged that this regulation will only be applicable in so far as the European unified regulation will be enacted. And indeed, a proposal for the EU regulation on the liability for the operation of Artificial Intelligence-systems was adopted recently. This article is going to study a question: what influence would the regulation have (if enacted) over the liablity for damage related to AVs?

12.1 The new German act on autonomous driving

> *„Das ist ein Riesenschritt Richtung Zukunft: Morgen tritt unser Gesetz zum autonomen Fahren in Kraft. Damit ist der Weg frei, um selbststeuernde Fahrzeuge ganz regulär auf die Straße zu holen – als erstes Land weltweit. Damit setzen wir internationale Standards."*
>
> Andreas Scheuer, Federal Minister of Transport and Digital Infrastructure, Germany[8]

On July 28, 2021, the Act Amending the Road Traffic Act and the Compulsory Insurance Act (Autonomous Driving Act)[9] entered into force in Ger-

[6] In one of the lawsuits against Tesla (in China), the father of the deceased asks Tesla to recognize that designating the level of automation at Level 2 SAE as an "autopilot" is misleading to customers (over-reliance on autonomous functions). FELTON, R.: Two Years On, A Father Is Still Fighting Tesla Over Autopilot And His Son's Fatal Crash. JALOPNIK, 27.2.2018. Online: https://jalopnik.com/two-years-on-a-father-is-still-fighting-tesla-over-aut-1823189786 (quoted 8.12.2021).

[7] AYAD, P., SCHUSTER, S., KOEPFERICH, K.: Germany takes a pioneering role with a new law on autonomous driving. Hogan Lovells, 2.8.2021. Online: https://www.engage.hoganlovells.com/knowledgeservices/analysis/germany-takes-a-pioneering-role-with-a-new-law-on-autonomous-driving (quoted 8.12.2021).

[8] Federal Ministry of Transport and Digital Infrastructure: Gesetz zum autonomen Fahren tritt in Kraft. 27.7.2021. Online: https://www.bmvi.de/SharedDocs/DE/Artikel/DG/gesetz-zum-autonomen-fahren.html (quoted 8.12.2021).

[9] Gesetz zur Änderung des Straßenverkehrsgesetzes und des Pflichtversicherungsgesetzes – Gesetz zum autonomen Fahren vom 12. Juli 2021.

many. The Autonomous Driving Act is one of several measures to create a legal framework to implement the German federal government's 2015 Strategy for Automated and Connected Driving.[10] Since 2015, the Federal Ministry of Transport and Digital Infrastructure has authorized testing[11] of automated and connected vehicles under real-life conditions. In June 2017, Germany amended its Road Traffic Act to allow drivers to transfer control of their vehicles to highly or fully automated driving systems and for those vehicles to be used on public roads.[12] According to this amendment, the driver is obligated to take over the driving functions from the automated driving system without undue delay if the driving system told him to do so or if the driver realized or should have realized that the conditions for using an automated driving system were no longer fulfilled (corresponding to SAE Level 3).

With the new law on autonomous driving from 2021, Germany has created the legal framework so that autonomous vehicles can drive regularly in specific areas of operation in public road traffic (SAE Level 4) – nationwide. The aim is to bring vehicles with autonomous driving functions into regular operation by 2022.[13] The Federal Ministry of Transport and Digital Infrastructure indicates examples of the application scenarios: shuttle traffic from A to B, people movers (buses that are traveling on a specified route), Hub2Hub transport (e.g., between two distribution centers), demand-oriented offers in off-peak times, the transport of people and/or goods on the first or last mile, "dual mode vehicles" such as Automated Valet Parking (AVP).[14]

The new act has introduced a new role of a "technical supervisor", a natural person responsible to ensure compliance with road traffic law. Technical supervisors are defined as natural persons who can deactivate the motor vehicle during operation and activate driving maneuvers for the vehicle. (§ 1d, para. 3.). The technical supervisor has several duties, such as as-

[10] Published by the Federal Ministry of Transport and Digital Infrastructure, September 2015. Online: https://perma.cc/TNL6-DFYN (quoted 8.12.2021).

[11] Federal Ministry of Transport and Digital Infrastructure: Digitale Testfelder. Online: https://perma.cc/5WJ8-HY7F (quoted 8.12.2021).

[12] For an analysis in English see JUHÁSZ, A.: The legal framework of autonomous driving in Germany. In: *MultiScience – XXXIII. microCAD International Multidisciplinary Scientific Conference*, University of Miskolc, 23 – 24 May 2019, ISBN 978-963-358-177-3. Online: https://www.uni-miskolc.hu/~microcad/cd2019/e1/E_Juhasz_Agnes.pdf (quoted 8.12.2021).

[13] Federal Ministry of Transport and Digital Infrastructure: Gesetz zum autonomen Fahren tritt in Kraft. 27.7.2021. Online: https://www.bmvi.de/SharedDocs/DE/Artikel/DG/gesetz-zum-autonomen-fahren.html (quoted 8.12.2021).

[14] *Ibidem*.

sessing transmitted vehicle data in order to be able to activate alternative driving manoeuvres or switch off the automated driving system as well as communicating with passengers. The role of the technical supervisor was assigned to the registered keeper of the motor vehicle (§ 1f, para. 1). However, the owner may delegate the technical supervision to another person. The owner must obtain additional liability insurance for the technical supervision. Moreover, the owner will be obligated to store and, upon request, transmit certain vehicle data (e.g., vehicle identification number, position data, environmental conditions and speed) in the case of specific safety-relevant occasions (e.g., in the event of interventions by the technical supervisor, conflict scenarios or malfunctions during operation) to the German Federal Motor Transport Authority (*Kraftfahrtbundesamt*) and the competent federal or state authorities. The liability of the vehicle owner under the current road traffic regime will generally not change. However, should the vehicle owner decide to delegate the tasks for technical supervision, he/she will be responsible for any fault of the person entrusted with those tasks.[15]

In addition to the aforementioned new role of a technical supervisor, other parties involved will need to fulfil additional obligations. The manufacturer of a vehicle with autonomous driving functions is, for example, responsible for aspects of cybersecurity (e.g., provision of proof that the electrical and electronic architecture of the vehicles is protected against attaches), risk assessment and trainings for persons involved in operation. The manufacturer must submit a certification that the vehicle complies with the technical requirements. The duty to comply with traffic regulations that do not relate to vehicle control and that a machine cannot perform, such as wearing a seat belt, rests with the humans in the vehicle. (§ 1f, para. 1.)[16] In the absence of a driver, the liability of the driver (§ 18 of the Road Traffic Act) is not applicable.

[15] AYAD, P., SCHUSTER, S., KOEPFERICH, K.: Germany takes a pioneering role with a new law on autonomous driving. Hogan Lovells, 2.8.2021. Online: https://www.engage.hoganlovells.com/knowledgeservices/analysis/germany-takes-a-pioneering-role-with-a-new-law-on-autonomous-driving (quoted 8.12.2021).

[16] GESLEY, J.: Germany: Road Traffic Act Amendment Allows Driverless Vehicles on Public Roads. Library of Congress. Online: https://www.loc.gov/item/global-legal-monitor/2021-08-09/germany-road-traffic-act-amendment-allows-driverless-vehicles-on-public-roads/ (quoted 8.12.2021).

The experts have generally praised this new German Act on Autonomous Driving.[17] However, from the point of view of Marion Jungbluth from the Federation of German Consumer Organizations (vzbv), the law is "sensible and correct for commercial applications" – but not for private vehicle owners. The new tasks imposed by the law would be overwhelming. Jungbluth sees a need for change in the question of liability. She criticized that instead of extending liability to the manufacturer according to the degree of automation, the law even extends liability of the owner.[18] In this concern, it is questionable whether in practice there would be any "private" vehicle owners falling under the scope of the new provision on autonomous driving. The relevant level of automation (SAE Level 4) is restricted to specific areas and conditions and the examples given by the Ministry relate to business purposes. The only possibly imaginable case of private use of AVs (for now) would probably be the Automated Valet Parking system where the owner of the car would need to fulfil the obligations of the "technical supervisor".

12.2 Proposal for a Regulation on liability for the operation of artificial intelligence-systems

On 20 October 2020 the European Parliament adopted a Resolution with recommendations to the Commission on a civil liability regime for artificial intelligence[19] (it was published in the Official Journal on 6 October 2021). Interestingly, despite not having the legislative initiative, the Parliament included a Proposal for a Regulation on liability for the operation of artificial intelligence-systems. It has yet to be seen how the Commission is going to react towards this resolution.

The Parliament explains in the introduction that it considers that the challenge related to the introduction of AI-systems into society, the workplace and the economy is one of the most important questions on the current political agenda. The Parliament believes that in order to efficiently exploit the advantages and prevent potential misuses of AI-systems and to avoid regula-

[17] Deutscher Bundestag: *Experten: Gesetz zum autonomen Fahren geht in die richtige Richtung*. Online: https://www.bundestag.de/dokumente/textarchiv/2021/kw18-pa-verkehr-autonomes-fahren-835640 (quoted 8.12.2021).

[18] *Ibidem*.

[19] European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)), P9 TA(2020)0276.

tory fragmentation in the Union, uniform, principle-based and future-proof legislation across the Union for all AI-systems is crucial. Therefore, it believes that the new common rules for AI-systems should only take the form of a regulation and that the question of liability in cases of harm or damage caused by an AI-system is one of the key aspects to address within this framework. The Parliament considers that the Product Liability Directive (PLD) has, for over 30 years, proven to be an effective means of getting compensation for harm triggered by a defective product, but should nevertheless be revised to adapt it to the digital world. It urges the Commission to assess whether the PLD should be transformed into a regulation, to clarify the definition of 'products' by determining whether digital content and digital services fall under its scope and to consider adapting concepts such as 'damage', 'defect' and 'producer'). The Parliament then states that the existing fault-based tort law of the Member States offers in most cases a sufficient level of protection for persons that suffer harm caused by an interfering third party like a hacker or for persons whose property is damaged by such a third party, as the interference regularly constitutes a fault-based action (only for specific cases, including those where the third party is untraceable or impecunious, does the addition of liability rules to complement existing national tort law seem necessary). Therefore, the Parliament considers it appropriate for this report to focus on civil liability claims against the operator of an AI-system. Due to the AI-system's complexity and connectivity, the operator will be in many cases the first visible contact point for the affected person.

12.2.1 The operator of AV as the target of liability

Under this Regulation 'operator' should be understood to cover both the frontend and the backend operator.

"Frontend operator" means any natural or legal person who exercises a degree of control over a risk connected with the operation and functioning of the AI-system and benefits from its operation.

"Backend operator" means any natural or legal person who, on a continuous basis, defines the features of the technology and provides data and an essential backend support service and therefore also exercises a degree of control over the risk connected with the operation and functioning of the AI-system.

If there is more than one operator of an AI-system, they shall be jointly and severally liable. The proposed regulation sets out conflict rules dealing

with the different combinations of actors (frontend operator/backend operator/producer/user):

- If a frontend operator is also the producer of the AI-system, this regulation shall prevail over the Product Liability Directive.
- If the backend operator also qualifies as a producer, that Directive should apply to him or her.
- If there is only one operator and that operator is also the producer of the AI-system, this Regulation should prevail over the Product Liability Directive.
- If a user, namely the person that utilises the AI-system (e.g., a passenger in an AV), is involved in the harmful event, he or she should only be liable under this Regulation if the user also qualifies as an operator. If not, the extent of the user's grossly negligent or intentional contribution to the risk might lead to the user's fault-based liability to the claimant. An example could be a passenger who spills water over technical device in an AV thereby causing its malfunction (and the situation results in a traffic accident).

According to *Wagner,* the fundamental question remains why the operators are identified as the main targets of the proposed liability scheme.[20] He argues that the user largely forfeits his influence over the "behavior" of the appliance, while control by the manufacturer increases; automobiles are named as a classic example. He underscores the need to shift liability away from the user of the car towards the manufacturer: the reason being that the user no longer controls the speed and direction of movement.[21] However, the concept of the liability of the operator is understandable from the point of view of those member states that distinguish between two types of liability: the fault-based liability of the driver and the strict liability of the owner, possessor and/or keeper of conventional motor vehicle.[22] For example, in Slovakia, the latter liability is named "liability for damage caused by operation of vehicles",[23] i.e., the liability of the operator. The driver and the operator can be the same person (e.g., the driver drives his own car) or they can be two different persons (e.g., when the driver is not the owner of the car). The liability of the operator is not rooted in his control over driving; the idea is that the

[20] WAGNER, G.: Liability for Artificial Intelligence: A Proposal of the European Parliament. 16.7.2021, p. 8. Online: https://ssrn.com/abstract=3886294 (quoted 8.12.2021).

[21] *Ibidem*, p. 9.

[22] E.g.: Germany, the Netherlands, Belgium, France, Slovakia, the Czech Republic.

[23] §§ 427–431 of the Slovak Civil Code.

operator benefits from the risk posed by the operation of the car[24] and exercises control over the technical state of the car (maintenance of the car).[25] The operator is liable for the damage caused by the car even if another person was driving. Therefore, from this perspective, the concept of operator as the target of the liability for the AVs appears logical.[26] This interpretation of the term "operator" can be supported by the wording of point 10 in the Introduction of the Resolution: *"the operator's liability is justified by the fact that he or she is controlling a risk associated with the AI-system, comparable to an owner of a car"*.

The qualification of the frontend operator is based on two elements: control (over the AI-system) and benefit (from the operation of the AI-system). The qualification of the backend operator is based solely on the control (over the AI-system). How do these definitions translate to the operation of the AVs? At first, it looks only fair that the frontend operator shares liability with some other actor. Once the AI-system is involved, the frontend operator is no longer in a position to exercise control over the technical conditions of the car. The control has shifted to the backend operator who on a continuous basis, defines the features of the technology and provides data and an essential backend support service (unless the front- and backend operator are the same person). The difficulty could be with the fully autonomous vehicles (Level 4 and 5 SAE): it is hard to imagine that there is any level of control on the side if the frontend operator. In order to impose liability on the frontend operator, it would be necessary to argue that there is enough control if the operator is able to turn the vehicle on and off.

[24] NOVOTNÁ, M., JURČOVÁ, M.: Zodpovednosť za škodu spôsobenú autonómne a poloautonómne riadenými vozidlami podľa slovenského právneho poriadku. In: SUCHOŽA, J., HUSÁR, J., HUČKOVÁ, R. (eds.): *Právo, obchod, ekonomika VII*, Košice: Univerzita P. J. Šafárika v Košiciach, 2017, p. 260–272, p. 265, ISBN 978-80-8152-528-5. Online: https://poe.pravo.upjs.sk/wp-content/documents/POE_2017_zbornik.pdf (quoted 8.12.2021).

[25] KRIŽAN, M.: Autonómne vozidlá: otázka zodpovednosti. In: *ACTA FACULTATIS IURIDICAE UNIVERSITATIS COMENIANAE*, Tomus XXXIX, 2/2020, p. 136–154, p. 143, ISSN 2729-8027. Online: https://afi.flaw.uniba.sk/index.php/AFI/article/download/63/56/111 (quoted 8.12.2021).

[26] This differentiation may sometimes not be obvious, especially when the author uses the term "operator" synonymously with "driver" (e.g., a common EU approach to liability rules and insurance for connected and autonomous vehicles. European Added Value Assessment Accompanying the European Parliament's legislative own-initiative report (Rapporteur: Mady Delvaux), p. 82.

12.2.2 Two regimes of liability

The Draft Regulation creates two sets of rules for claims against operators depending on the level of risk connected with the operation of the AI-system:

(a) Liability for high-risk AI systems and
(b) Liability for other AI systems.

12.2.2.1 High-risk AI systems

All AI-systems with a high risk should be exhaustively listed in an Annex to this Regulation and frequently reviewed by the Commission so that the list remaind up to date. The Annex that was originally attached to the draft Parliament resolution listed AVs SAE Level 4 and 5 in the high-risk AI-systems (however, there is no such list in the draft regulation itself).

Operators of high-risk systems would have strict liability for harm or damage caused, subject only to force majeure. The Draft Regulation also sets out maximum compensation amounts,[27] methods for calculating the extent of compensation, and limitation periods, while subjecting operators of high-risk systems to a mandatory insurance regime. This liability regime should take precedence over the national legal rules for liability of the vehicle operator – e.g., the rules in § 427 and following of the Slovak Civil Code. The national rules on the liability of the driver should not be applicable at all – simply because there is no person in the position of the driver in the SAE Levels 4 and 5 (in Slovakia there is a fault-based liability regime in § 420 of the Civil Code).

12.2.2.2 Other AI systems

The second set of rules applies to AI systems not perceived as high-risk. *A contrario* to the list of AVs included in the high-risk category, the SAE Levels 1–3 would fit in the other AI systems. Article 8 establishes a regime where fault-based liability is presumed, unless it is a case of force majeure or one of the grounds in Art. 8(2) applies:

[27] According to Article 5: up to a maximum amount of 2 million EUR in the event of the death of, or in the event of harm caused to the health or physical integrity of, an affected person; up to a maximum amount of 1 million EUR in the event of significant immaterial harm that results in a verifiable economic loss or of damage caused to property; no compensation shall be paid, if the total amount of the damage to property or the significant immaterial harm is of a value that falls below 500 EUR.

a) the AI-system was activated without his or her knowledge while all reasonable and necessary measures to avoid such activation outside of the operator's control were taken, or
b) due diligence was observed by performing all the following actions: selecting a suitable AI-system for the right task and skills, putting the AI-system duly into operation, monitoring the activities and maintaining the operational reliability by regularly installing all available updates.

The operator shall not be able to escape liability by arguing that the harm or damage was caused by an autonomous activity, device or process driven by his or her AI-system. The operator would also be liable for harm or damage caused by untraceable or impecunious interfering third parties (e.g., a hacker interfered with the AI-system). The Draft Regulation also imposes a duty on the producer to cooperate with and provide information to requesting operators and affected persons in order to facilitate the identification of liabilities. This Parliament has been criticized that by referring to the causes for excluding fault-based liability it recommended a heterodox subjective civil liability, with clear elements of objective apportionment.[28]

Importantly, the limitation periods as well as the amounts and the extent of compensation would be governed by the laws of the Member State where the harm or damage occurred (Article 9). The draft regulation thus creates a mosaic approach where the rules of the regulation have to be combined with the national fault-based liability regimes.[29] Galbois-Lehalle suggests that as a consequence, in cases of harm caused by AI-systems not considered high-risk, where the applicable national law is particularly favourable to the victims (e.g., France or Belgium), the victim will be able to obtain compensation for losses that could not be repaired had the damage been caused by a high-risk AI-system (such as purely moral damage, psychologic distress).[30] This observation depends on the terminology: is "the amounts and the ex-

[28] SOUSA ANTUNES, H.: Civil Liability Applicable to Artificial Intelligence: A Preliminary Critique of the European Parliament Resolution of 2020. 5.12.2020. Online: https://ssrn.com/abstract=3743242 (quoted 8.12.2021).

[29] In Slovakia, the rules for compensation are to be found in §§ 442–450 of the Slovak Civil Code.

[30] GALBOIS-LEHALLE, D.: AI and Civil Liability: Welcomed but Perfectible Recommendations of the European Parliament, Chair Legal and Regulatory Implications of Artificial Intelligence. MIAI Grenoble Alpes, 26.1.2021. Online: https://ai-regulation.com/civil-liability-regime-for-artificial-intelligence-welcomed-but-perfectible-recommendations-of-the-european-parliament/ (quoted 8.12.2021).

tent of compensation" the same issue as "the type of loss"? Not likely: the draft regulation specifies that "harm or damage" means an adverse impact affecting the life, health, physical integrity of a natural person, the property of a natural or legal person or causing significant immaterial harm that results in a verifiable economic loss (the definition is in Article 3 which is common to high-risk and other-risk AI-systems).

It has to be noted, that the liability for the damage caused by a conventional vehicle (where there is no involvement of the AI-system) would remain not harmonized among the EU member States (with considerable differences in the liability regimes). As the result there will be member states, where the operator of the conventional car would be strictly liable for the damage caused by the operation of the car (e.g., the rules in § 427 and following of the Slovak Civil Code) while the operator of the non-high-risk AV would bear the fault-based liability. More clarification from the legislator on this aspect would be helpful in order to establish that this is indeed a desirable discrimination.

12.2.3 Private International Law perspective

The relationship between the draft regulation and the rules for the conflict of laws is potentially problematic. In case of a traffic accident with foreign element, either the Convention of 4 May 1971 on the Law Applicable to the Traffic Accidents or the Rome II Regulation[31] are applicable. However, the wording of the scope of the draft regulation in Article 2(1) could cause confusion as to the applicability of these rules: "*This Regulation applies on the territory of the Union where a physical or virtual activity, device or process driven by an AI-system has caused harm or damage to the life, health, physical integrity of a natural person, to the property of a natural or legal person or has caused significant immaterial harm resulting in a verifiable economic loss.*"

It has been argued by *van Hein* that this wording can be deemed as a unilateral conflicts rule (very unusual in modern private international law). As such, it would prevail over the Rome II Regulation on the law applicable to non-contractual relations pursuant to Art. 27 Rome II, which states that the Rome II Regulation shall not prejudice the application of provisions of EU law which, in relation to particular matters, lay down conflict-of-law rules re-

[31] Regulation (EC) No 864/2007 of the European Parliament and of the Council of 11 July 2007 on the law applicable to non-contractual obligations (Rome II), OJ L 199, 31.7.2007, p. 40–49.

lating to non-contractual obligations. Insofar, it must be noted that Art. 2(1) DR deviates considerably from the choice-of-law framework of Rome II. It does not provide any exception to the lex loci damni rule (in Rome II: lex loci domicili communis), there is no escape clause or facility for agreement.[32] There is another conflict rule formulated in Article 9, whereby claims brought under against non-high-risk operators would have limitation periods and compensation amounts determined according to 'the laws of the Member State in which the harm or damage occurred.'

There is another possible interpretation of Art. 2(1) in a way that allows for the operation of Rome II. The rule in Art. 2(1) may be conceptualised as a 'substantive corrective', as is arguably the case with Art. 3(2) of the e-Commerce Directive, although in such a form it would have an indirect influence on conflict of laws.[33]

Furthermore, no relationship with the 1971 Convention on the law applicable to traffic accidents is established.[34] Generally, one would expect that this Convention should prevail over the draft regulation, because it was not concluded exclusively between two or more EU member states (there are also non-EU member states parties to this convention). Therefore, the obligations set under the EU regulation should not prevail over the international obligations of the EU members states towards the third states.

Conclusion

In order to create an appropriate liability regime for AVs a clear legal distinction has to be established between the particular levels of automation in vehicles. This approach has been followed by in Germany where the new Act on autonomous driving is specifically designed for SAE Level 4. At the same time, it is inevitable that the drivers are properly trained and always fully aware of a vehicle's level of automation and the corresponding liability. As far

[32] von HEIN, J.: Forward to the Past: A Critical Note on the European Parliament's Approach to Artificial Intelligence in Private International Law. Conflictoflaws.net, 22.10.2020. Online: https://conflictoflaws.net/2020/forward-to-the-past-a-critical-note-on-the-european-parliaments-approach-to-artificial-intelligence-in-private-international-law/ (quoted 8.12. 2021).

[33] Study on the Rome II Regulation (EC) 864/2007 on the law applicable to non-contractual obligations, JUST/2019/JCOO_FW_CIVI_0167, p. 72–73. Online: https://op.europa.eu/en/publication-detail/-/publication/11043f63-200c-11ec-bd8e-01aa75ed71a1 (quoted 8.12. 2021).

[34] The convention is binding upon 14 EU member states (including Slovakia).

as the level of the fully autonomous vehicles is concerned (SAE Level 4 and 5) there is a tendency for a shift towards the strict liability of the operator for the AV. This is visible in the new German national liability rules as well as in the draft regulation on liability for the operation of AI-systems proposed by the European Parliament. The German model is criticised that the tasks imposed by the law would be overwhelming for the private AVs owners. In comparison, the draft regulation further distinguishes between the frontend and backend operator and could stand up to this objection. However, the rules that the regulation proposed for fault-based liability of the operator could be blamed for creating a parallel application of the regulation and the national rules that is potentially too complicated and lacks persuasive explanation on the distinction between the conventional versus autonomous vehicles. Moreover, the draft would need to be revisited to rethink its relationship with the rules on conflict of laws (Rome II Regulation and the 1971 Convention on the law applicable to traffic accidents).

Bibliography

1. Act No. 40/1964 Coll. Civil Code (Slovakia).
2. AYAD, P., SCHUSTER, S., KOEPFERICH, K.: Germany takes a pioneering role with a new law on autonomous driving. Hogan Lovells, 2.8.2021. Online: https://www.engage.hoganlovells.com/knowledgeservices/analysis/germany-takes-a-pioneering-role-with-a-new-law-on-autonomous-driving (quoted 8.12.2021).
3. BARTNECK, C., LÜTGE, C., WAGNER, A., WELSH, S.: *An Introduction to Ethics in Robotics and AI*. SpringerLink, 2021, ISBN 978-3-030-51110-4. Online: https://link.springer.com/book/10.1007/978-3-030-51110-4 (quoted 8.12.2021).
4. Convention of 4 May 1971 on the Law Applicable to Traffic Accidents. Online: https://www.hcch.net/en/instruments/conventions/full-text/?cid=81 (quoted 8.12.2021).
5. Deutscher Bundestag: *Experten: Gesetz zum autonomen Fahren geht in die richtige Richtung*. Online: https://www.bundestag.de/dokumente/textarchiv/2021/kw18-pa-verkehr-autonomes-fahren-835640 (quoted 8.12.2021).
6. European Added Value Assessment Accompanying the European Parliament's legislative own-initiative report (Rapporteur: Mady Delvaux).
7. European Parliament News: Self-driving cars in the EU: from science fiction to reality. 14.1.2019. Online: https://www.europarl.europa.eu/news/en/headlines/economy/20190110STO23102/self-driving-cars-in-the-eu-from-science-fiction-to-reality (quoted 8.12.2021).

8. European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)), P9 TA(2020)0276.
9. Federal Ministry of Transport and Digital Infrastructure: 2015 Strategy for Automated and Connected Driving. September 2015. Online: https://perma.cc/TNL6-DFYN (quoted 8.12.2021).
10. Federal Ministry of Transport and Digital Infrastructure: Digitale Testfelder. Online: https://perma.cc/5WJ8-HY7F (quoted 8.12.2021).
11. Federal Ministry of Transport and Digital Infrastructure: Gesetz zum autonomen Fahren tritt in Kraft. 27.7.2021. Online: https://www.bmvi.de/SharedDocs/DE/Artikel/DG/gesetz-zum-autonomen-fahren.html (quoted 8.12.2021).
12. FELTON, R.: Two Years On, A Father Is Still Fighting Tesla Over Autopilot And His Son's Fatal Crash. JALOPNIK, 27.2.2018. Online: https://jalopnik.com/two-years-on-a-father-is-still-fighting-tesla-over-aut-1823189786 (quoted 8.12.2021).
13. GALBOIS-LEHALLE, D.: AI and Civil Liability: Welcomed but Perfectible Recommendations of the European Parliament, Chair Legal and Regulatory Implications of Artificial Intelligence. MIAI Grenoble Alpes, 26.1.2021. Online: https://ai-regulation.com/civil-liability-regime-for-artificial-intelligence-welcomed-but-perfectible-recommendations-of-the-european-parliament/ (quoted 8.12.2021).
14. Gesetz zur Änderung des Straßenverkehrsgesetzes und des Pflichtversicherungsgesetzes – Gesetz zum autonomen Fahren vom 12. Juli 2021 (Germany).
15. GESLEY, J.: Germany: Road Traffic Act Amendment Allows Driverless Vehicles on Public Roads. Library of Congress. Online: https://www.loc.gov/item/global-legal-monitor/2021-08-09/germany-road-traffic-act-amendment-allows-driverless-vehicles-on-public-roads/ (quoted 8.12.2021).
16. JUHÁSZ, A.: The legal framework of autonomous driving in Germany. In: *Multi-Science – XXXIII. microCAD International Multidisciplinary Scientific Conference*, University of Miskolc, 23-24 May 2019, ISBN 978-963-358-177-3. Online: https://www.uni-miskolc.hu/~microcad/cd2019/e1/E_Juhasz_Agnes.pdf (quoted 8.12.2021).
17. KRIŽAN, M.: Autonómne vozidlá: otázka zodpovednosti. In: *ACTA FACULTATIS IURIDICAE UNIVERSITATIS COMENIANAE*. Tomus XXXIX, 2/2020, p. 136, p. 143, ISSN 2729-8027. Online: https://afi.flaw.uniba.sk/index.php/AFI/article/download/63/56/111 (quoted 8.12.2021).
18. Regulation (EC) No 864/2007 of the European Parliament and of the Council of 11 July 2007 on the law applicable to non-contractual obligations (Rome II), OJ L 199, 31.7.2007.
19. Regulation (EU) 2019/2144 of the European Parliament and of the Council of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users, OJ L 325, 16.12.2019.

20. SAE International: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, J3016_202104. 30.4.2021. Online: https://www.sae.org/standards/content/j3016_202104/ (quoted 8.12.2021).
21. SOUSA ANTUNES, H.: Civil Liability Applicable to Artificial Intelligence: A Preliminary Critique of the European Parliament Resolution of 2020. 5.12.2020. Online: https://ssrn.com/abstract=3743242 (quoted 8.12.2021).
22. Study on the Rome II Regulation (EC) 864/2007 on the law applicable to non-contractual obligations, JUST/2019/JCOO_FW_CIVI_0167. Online: https://op.europa.eu/en/publication-detail/-/publication/11043f63-200c-11ec-bd8e-01aa75ed71a1 (quoted 8.12.2021).
23. SUCHOŽA, J., HUSÁR, J., HUČKOVÁ, R. (eds.): *Právo, obchod, ekonomika VII*. Košice: Univerzita P. J. Šafárika v Košiciach, 2017, ISBN 978-80-8152-528-5. Online: https://poe.pravo.upjs.sk/wp-content/documents/POE_2017_zbornik.pdf (quoted 8.12.2021).
24. von HEIN, J.: Forward to the Past: A Critical Note on the European Parliament's Approach to Artificial Intelligence in Private International Law. Conflictoflaws.net, 22.10.2020. Online: https://conflictoflaws.net/2020/forward-to-the-past-a-critical-note-on-the-european-parliaments-approach-to-artificial-intelligence-in-private-international-law/ (quoted 8.12.2021).
25. WAGNER, G.: Liability for Artificial Intelligence: A Proposal of the European Parliament. 16.7.2021. Online: https://ssrn.com/abstract=3886294 (quoted 8.12.2021).

Author details

Mgr. Ľubica GREGOVÁ ŠIRICOVÁ, PhD.
Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of International and European Law
lubica.siricova@upjs.sk

13 AUTONOMOUS WEAPON SYSTEMS IN THE LIGHT OF THE PRINCIPLES OF DISTINCTION AND PROPORTIONALITY

Juraj Panigaj

Abstract

It is a known fact that the technological progress around artificial intelligence is evolving constantly, not excluding weapon systems. Progress of unmanned (or "robotic") systems is growing not only in the meaning of their quantity but also regarding abilities to perform more and more functions and actions without a direct human supervisor or intervention of human factor. The article sets a goal to analyze these weapon systems regarding possible use while respecting the fundamental principles of international humanitarian law nowadays and in the near future. Such assessment is quite a challenge due to criticism of many authors and organizations who claim that autonomous weapons could not comply with international requirements and rules or guarantee the security of civilians. This group calls for an absolute ban on autonomous weapon systems. However, due to its huge potential and interest in the development of some of the world-powers (some of the members of the UN´s Security council included), the result in the form of an absolute ban is unlikely. That is why the article is oriented more to the issue of their potential use and necessity of compliance with the most important principles of international humanitarian law associated with it. By analyzing and researching the current technological state, relevant international rules, and attitude of the international community, we want to summarize a certain counterpart to the opinions leading to the ban of autonomous weapons in general without further discussions based on its "potential danger." In the event of being compliant with the principles of distinction and proportionality and in the case of strict international regulation, the use of these weapons could be preferable and safer than the deployment of its human counterparts.

Introduction

Over the years, artificial intelligence has gained solid ground in many aspects of our lives. Without any doubt, it is well deserved, because it can make things easier, simpler, or perform human activities faster, better, and safer. However, such positivity has no place regarding its militarization. In this field, artificial intelligence raises legitimate concerns in a form of autonomous weapon systems.

It is the militarization of artificial intelligence in the form of autonomous weapon systems that is a core issue of this paper in the context of its compli-

ance with the principles of distinction and proportionality. The author draws on the views of both proponents and opponents of autonomous weapon systems. The core argument, the author is working with is the argument of the opponents regarding the "inhuman" nature of the autonomous systems and its impact on the fundamental principles of international humanitarian law. The article tries to compare these opinions in the light of current technological development and international humanitarian legislation as well. It attempts to answer the possible capability of the autonomous weapon systems to be compliant with the principles of the international humanitarian law by thoroughly analyzing interdisciplinary professional literature and the international humanitarian rules accordingly. To this end, it also presents an illustrative concept of an international legal instrument with a demonstrative calculation of the issues it should contain to ensure compliance with these principles.

13.1 Autonomous weapon systems and compliance with the core principles of the international humanitarian law

Firstly, for the purposes of this article, it is necessary to define autonomous weapon systems. Although international law or especially the international humanitarian law does not contain a legal definition of such weapons, we can define it as a weapon system with autonomy in almost every aspect of its operation, thus it independently chooses (seeks out, identifies, follows…) a target, which it subsequently eliminates or hit in a different manner depending on the goal of the mission. All of that is executed without human intervention. This is also the most important issue that it can operate without direct human supervision.[1]

The nature of such a weapon raises considerable concerns and brings risks related to the difficulty to predict and reduce its effects alternatively. It is the loss of human control that multiplies these concerns. A number of states and organizations call for an absolute ban of autonomous weapon systems without any relevant results as to whether they will be able to act under the principles of international humanitarian law or not.

[1] DAVISON, N.: A legal perspective: Autonomous weapon systems under International humanitarian law. International committee of the red cross, 31.1.2018. Online: https://www.icrc.org/en/document/autonomous-weapon-systems-under-international-humanitarian-law (quoted 9.10.2021).

The International Committee of the Red Cross states in its position on autonomous weapon systems that the process by which autonomous weapon systems function:

- brings risks of harm for those affected by armed conflict, both civilians and combatants, as well as dangers of conflict escalation,
- raises challenges for compliance with international law, including international humanitarian law, notably, the rules on the conduct of hostilities for the protection of civilians,
- raises fundamental ethical concerns for humanity, in effect substituting human decisions about life and death with sensor, software, and machine processes.[2]

Given the scope of this issue, the paper is limited to analyzing the challenges of compliance with international humanitarian law, in particular with the principles of proportionality and distinction. Therefore, the paper does not address the ethical or moral aspects of the use of the autonomous weapons.

As mentioned before, the essence of this article is to assess the issue of the capability of the autonomous weapon systems to be compliant with the principles of international humanitarian law. Principles or to be more accurate, rules of the international humanitarian law apply to victims of armed conflicts or those hors de combat.[3]

Fundamental rules of international humanitarian law can be found in the Geneva Conventions of 1949 and their Additional Protocols. The issue of autonomous weapons, in general, is regulated in the Article no. 36 of the Additional protocol I. which states:

"In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party."[4]

The most dominant principles of international humanitarian law include the principles of discrimination and proportionality.

[2] International committee of the red cross: ICRC's position on autonomous weapon systems. 12.5.2021. Online: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems (quoted 9.10.2021).

[3] KLUČKA, J.: *Medzinárodné právo verejné (všeobecná a osobitná časť)*, Bratislava: Wolters Kluwer, 2017, p. 469, ISBN 978-80-8168-744-0.

[4] Notice from the Federal Ministry of Foreign Affairs no. 168/1991 Coll. on the Additional Protocols I and II to the Geneva Conventions of 12.8.1949.

13.1.1 Compliance with the principle of distinction

An expression of the principle of distinction is also contained in the Additional Protocol I. to the Geneva Conventions. The principle is defined in Article no. 48 as follows:

"In order to ensure respect for and protection of the civilian population and civilian objects, the Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives."

The principle can be found in several other provisions, for example in the Article no. 57 regarding precautions in the attack:

"With respect to attacks, the following precautions shall be taken do everything feasible to verify that the objectives to be attacked are neither civilians nor civilian objects and are not subject to special protection but are military objectives within the meaning of paragraph 2 of Article 52 and that it is not prohibited by the provisions of this Protocol to attack them."[5]

To simplify, the essence of this principle is to distinguish between civilians, those hors de combat and civilian objects on the one side, and combatants and military objects on the other side. Concerning the autonomous weapon systems, compliance with the principle of distinction raises questions mostly regarding the urbanization of the armed conflicts and diminishing differences between combatants and civilians.

A machine would probably be able to comply with the principle of distinction if deployed outside populated areas, while the target would be for example a military base. In such a scenario, specification of a location and target itself should ensure international humanitarian law would not be violated and we would not have to presume the object (or a person) is a civil one because in case of doubts international humanitarian law dictates to presume, it is civilian (according to the Article no. 52, sec. 3 and 50 sec. 1 of the Additional Protocol I.).[6]

On the other hand, it is not the same case with armed conflicts which take place in a populated area, let´s say in a city. In this scenario, a robot would most likely fail to comply with the principle. The reason is many entities in armed conflicts nowadays have a non-state character. Their combatants do not wear uniforms, they look alike civilians or make civilians or even chil-

[5] *Ibidem.*
[6] *Ibidem.*

dren perform various acts instead of them. As a result, the ability to distinguish combatants from civilians or those hors de combat usually requires considering an individual´s intentions based on other characteristics, for example, body language, gestures, or tone of voice. Opponents of Autonomous weapons are convinced that humans who can relate to other people, can better interpret those cues than inanimate robots.[7] We need to point out these situations often cause problems even for human soldiers to distinguish between lawful and unlawful targets.

However, the development of artificial intelligence shows us future machines could be able to operate even in populated areas despite these objections. It looks like one day we should be able to teach artificial intelligence to "read" body language, gestures, or to evaluate a behavior based on a tone, pitch, or even rhythm of speaking. In 2017 researchers at Carnegie Mellon University have enabled a computer to understand body poses and locomotion of several people from video in real-time and for the first time even the pose of an individual´s hands and fingers.[8] In the same year, researchers from MIT's Computer Science and Artificial Intelligence Laboratory and Institute of Medical Engineering and Science have reached a milestone in the development of artificial intelligence capable of predicting whether a conversation is pleasant, sad, or neutral based on various speech patterns and vitals. In the research, the system analyzed a voice (tone, pitch, energy of voice, or vocabulary), transcripts, and physiological signals (movements, heartbeat, blood pressure, etc.) to determine the emotionality of a story told by a person. The system was successful at 83%. It was one of the first research focused on collection both physiological and speech data, with the results showing that it is possible to classify the emotional tone of a conversation in real-time.[9]

The research proofs that dynamic, speed, and capacity concerning the ability to learn are much higher with machines than humans. However, what machines are not better in yet, in addition to the emotional concept (despite

[7] DOCHERTY, B.: The Need for and Elements of a New Treaty on Fully Autonomous Weapons. Human rights watch, 20.2.2020. Online: https://www.hrw.org/news/2020/06/01/need-and-elements-new-treaty-fully-autonomous-weapons (quoted 10.10.2021).

[8] SPICE, B.: Computer Reads Body Language. Carnegie Mellon University. 6.7.2017. Online: https://www.cmu.edu/news/stories/archives/2017/july/computer-reads-body-language.html (quoted 10.10.2021).

[9] CONNER-SIMONS, A., GORDON, R.: Wearable AI system can detect a conversation´s tone. MIT news, 2.1.2017. Online: https://news.mit.edu/2017/wearable-ai-can-detect-tone-conversation-0201 (quoted 10.10.2021).

the abovementioned research), is the ability to predict and link seemingly unrelated information.[10] To illustrate this matter we could compare an autonomous weapon with an autonomous vehicle. To be fully autonomous it is not sufficient enough we "teach" the car the traffic rules. Such autonomous vehicles must be able to react even in unpredictable situations. The very recognition of the surroundings is a complicated matter, and the complications would just multiply when talking about a busy environment of a modern urban battlefield. Although the well-trained system can recognize various objects, in accordance with the abovementioned it must be able to relate these objects to other factors, facts.[11]

Regarding the autonomous weapon systems, we could picture it in the following scenario. Should an enemy soldier drop the weapon, run away, and hide in a building, a human soldier must presume, such maneuver could be a trap and in this building are more hidden enemies or even hostages. Because of that, he must carefully evaluate a situation before taking another step. A soldier cannot be satisfied by the mere fact the enemy dropped the weapon, and presume it is safe to pursue him. To simplify, we can´t focus only on completing the mission. Setting a human goal to a machine could potentially cause a problem.

Based on British computer scientist Stuart Russell´s opinion, there is a particular risk that the current model of artificial intelligence understands the assigned task as the ultimate goal that needs to be achieved. Thus, a complex machine can pursue a simply formulated goal, basically "through dead bodies," not to look beyond anything else. Human preferences on other hand are not so fatal and they tend to change due to circumstances. At the same time, Russell brings attention to the possibility that one of the first steps of artificial intelligence pursuing the goal could be simply a deactivation of a "kill switch" button.[12] In other words, having a robot instead of a human soldier in the abovementioned scenario might cause that a machine would pursuit the enemy regardless of the consequences and circumstances not related to the situation at the first sight. At present, autonomous weapon systems should not be used under such circumstances with a respect to the current state of development.

[10] KOLAŘÍKOVÁ, L., HORÁK, F.: *Umělá inteligence & právo,* Praha: Wolters Kluwer, 2020, p. 25, ISBN 9788075987839.
[11] *Ibidem*.
[12] *Ibidem,* p. 54–55.

Even if the autonomous weapon systems would not be able to comply with the principle of distinction in a populated area in the near future, it still does not have to be a reason for the absolute, general ban. There would be still a possibility to think about deploying a machine in isolated areas, where the requirements for distinguishing would be within its capabilities. The machine could be for example programmed with certain geographic limitations preventing the machine from entering the inhabited area.[13] It would be unnecessarily risky to deploy such a weapon in an area inhabited by civilians, however, it is not out of the question in the near future. In such cases, for example, the machine would be obliged to evaluate such results as justified to interrupt the mission based on data questioning the attack. Nevertheless, the mere existence of doubt does not bring the presumption into operation. The degree of doubt must cause a reasonable attacker in the same or similar circumstances would hesitate before attacking. On the other hand, the doubt threshold is framed in terms of human reasonableness, so in the context of autonomous weapons, it is another complication.[14] In the field of autonomy this problem of artificial intelligence, where the machine prefers to "freeze because it fears" its further steps are dangerous is referred to as the "freezing robot problem."[15] Needless to say, such a process should be applied by the machine regardless of the isolation of its operation, or target. Therefore, the machine should attack only after a sufficient number of pre-programmed characteristics consistent with the objectives that are legal under the principle of distinction. This type of analysis would be largely based on quantitative data (e.g., shape or size of a target).[16] Sensors recognizing specific categories of people come into account to minimize doubts in connection with the size or shape of the target. Assuming the system is capable of recognizing a child regarding its anatomical-morphological characteristics distinguishing it from adults, whether a person holds a weapon, or wears a uniform,

[13] SCHMITT, M. N.: Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics. In: *Harvard National Security Journal Feature*, 2013, p. 13–14, ISSN 2153-1358. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2184826 (quoted 10.10.2021).

[14] *Ibidem,* p. 15–17.

[15] KOLAŘÍKOVÁ, L., HORÁK, F.: *Umělá inteligence & právo*, op. quoted, p. 25.

[16] WAGNER, M.: The Dehumanization of International Humanitarian Law: Legal, Ethical and Political Implications of Autonomous Weapon Systems. In: *Vanderbilt Journal of Transnational Law*, vol. 47, 2014, p. 1391–1392, ISSN 0090-2594. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2541628 (quoted 10.10.2021).

such an exclusion method would increase the likelihood that it is a legal target in accordance with the principle of distinction.[17]

Only after the machine meets the requirements of the principle of distinction, we can subsequently deal with the issue of complying with the principle of proportionality. If the autonomous weapon system is not capable of complying with the principle of distinction, it is almost certain, it will fail in relation to the principle of proportionality as well.

13.1.2 Compliance with the principle of proportionality

The principle is expressed in several provisions of the Additional Protocol I. For example, it is codified in Article no. 51, sec. 5, b) and then repeated in Article no. 57. The principle demands *"to cancel or stop any attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated."* At the same time, the Protocol states that *"it is prohibited to employ weapons, projectiles and material and methods of warfare of a nature to cause superfluous injury or unnecessary suffering."* (Article no. 35, sec. 2)[18]

The principle itself compares the incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof with the anticipated result of the attack in the time when the attack was initiated. The assessment does not take place retrospectively. Therefore, if the attack resulted in more excessive collateral damage than initially expected, the principle of proportionality was not violated if the original assumptions of collateral damage could not be considered excessive in light of the expected military advantage.[19]

To be compliant with the principle in the context of artificial intelligence brings even greater challenges as was the case with the principle of distinction. Proportionality is a highly individual matter, and it can be adequately specified only in relation to the concrete attack, time, by using specific weapons or strategies of the attack.[20] Therefore, individual situations almost

[17] SCHMITT, M. N.: *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics,* op. quoted, p. 17.

[18] Notice from the Federal Ministry of Foreign Affairs no. 168/1991 Coll. on the Additional Protocols I and II to the Geneva Conventions of 12.8.1949.

[19] WAGNER, M.: *The Dehumanization of International Humanitarian Law: Legal, Ethical and Political Implications of Autonomous Weapon Systems*, op. quoted, p. 1396.

[20] *Ibidem,* p. 1393.

always bring different variables, within which we must balance on the one hand anticipated military advantage and incidental damages on the other hand. Such analysis cannot be inserted into some universal equation as these factors are not quantifiable, measurable quantities, they always need to be addressed individually[21] (by extreme simplification, incidental loss of 4 civilian lives cannot be always balanced with eliminating 15 enemy combatants). It would be a mistake to grant the machine full autonomy after this evaluation.

Even though the issue of comparison between anticipated military advantage and incidental loss cannot be universally quantified, math is still an integral part of evaluating compliance with the principle of proportionality. A robot should be faster and more precise in calculating for example radius of a blast or other effects of a weapon, which could cause incidental damage, and on top of that, in real-time. A human soldier would not be able to make such calculations during an attack in real-time, since these calculations would be too complex to calculate in a short period of time under a lot of pressure. The autonomous system would be able to perform a tremendous number of calculations in real-time leading to increase military advantage and decrease the probability of incidental damages at the same time.[22] A hand in hand with this ability goes what weapon would be robot equipped with. As with other weapons, it should be noted that legal use also depends on what conditions the machine would operate in or what weapons it would be equipped with. It is not a factor justifying a ban on autonomous weapon systems, but it is one of the determining factors in specific situations to determine the legality of a particular attack. Moreover, it can affect whether the machine meets the requirements of the principles of proportionality and distinction. Therefore, certain weapon limitations should be set in the international legal instrument regulating the use of autonomous weapon systems.

Regarding the principle of proportionality, it is necessary to mention the so-called Collateral Damage Estimation Methodology (hereinafter also as "CDEM"). As European External Action Service defines it, CDEM is *"a*

[21] ANDERSON, K., REISNER, D., WAXMAN, M.: Adapting the Law of Armed Conflict to Autonomous Weapon Systems. In: *International Law Studies*, vol. 90, 2014, p. 405. Online: https://apps.dtic.mil/dtic/tr/fulltext/u2/a613290.pdf (quoted 16.10.2021).

[22] EASON, P. M.: *Lethal Autonomous Weapon Systems: Reconciling the Myth of Killer Robots and the Reality of the Modern Battlefield*, thesis for the degree of Master of Arts in Humanities, Graduate School of Duke University, 2021, p. 42. Online: https://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/23212/Eason_duke_0066N_16305.pdf?sequence=1 (quoted 16.10.2021).

body of joint standards, methods, techniques, and processes to conduct collateral damage analysis and produce Collateral Damage estimates." CDEM serves as a supportive instrument to perform the proportionality test. To pass the test, CDEM must predict with a high accuracy anticipated collateral damages in the light of a type of attack. Such a process contains a form of "tiering" methodology. The higher risk of collateral damage is present, the higher tier must be used (e.g., with a higher tier there is a need for the consent of hierarchically higher authority). To describe the level of risk there will be used a form measurable by defined CDEM tools. Such tools will be used by designated and qualified operators and also used to mitigate risks by designing the variables in an attack, in particular regarding lethal attacks.[23] "It includes variables such as:

- parameters (e.g., direction) of the attack,
- time of the attack,
- the weapons to be used,
- the environment and geography of the target,
- the target itself,
- the proximity of any other structures or civilians within a given distance from the target."[24]

As mentioned above, the CDEM process contains a form of "tiering" methodology, so in other words, the higher risk of collateral damages comes into account, the higher tier must be applied. For example, European External Action Service states 5 possible levels or tiers. In the event, all reasonable and known techniques to avoid or at least reduce incidental damages fail, it is necessary to apply the highest tier possible (level 5). Level 5 application occurs if there are concerns of biological, radiological, or environmental nature present, or if the attack may cause civilian casualties or damage civilian objects. Casualty assessment needs available intelligence (Pattern of life, Demographic data) to provide an estimate of population density in a specific collateral concern area. Commanders must be informed that they suspect a serious risk of collateral damage when engaging a target assessed under level 5.[25]

[23] European External Action Service (EEAS): Avoiding and Minimizing Collateral Damage in EU-led Military Operations Concept. 3.2.2016, p. 12. Online: https://data.consilium.europa.eu/doc/document/ST-5785-2016-INIT/en/pdf (quoted 16.10.2021).

[24] *Ibidem.*

[25] *Ibidem,* p. 20.

Even though the CDEM can provide essential information needed to successfully pass the proportionality test, the Methodology does not answer, whether the specific attack violated the principle of proportionality or not.[26] Autonomous weapon systems could operate on a similar approach based on CDEM, as they can process an enormously large amount of data in a short period of time, and the Methodology itself requires many calculations consisting of a lot of variables. The issue regarding the calculation of anticipated military advantage concerning the complexity of modern battlefields would be more challenging[27] as well as assessment of proportionality on the basis of calculated incidental damages.

As we mentioned before, evaluation and calculation of quantitative data are not sufficient enough to assess the proportionality. Although such data and its processing is an integral part of the process, psychological processes and personal beliefs are equally important at the same time. However, artificial intelligence is not capable of ensuring its application. Aforesaid processes are exclusively linked to humans. It is not enough to rely only on objective data to initiate an attack. International Committee of the Red Cross commentary to the Additional Protocol I. states that the proportionality test is subjective, allows for a "fairly broad margin of judgment," and "must above all be a question of common sense and good faith for military commanders."[28] Simply put, while at present a machine may be able to evaluate the extent of collateral damage in a concrete attack, it cannot distinguish whether these collateral damages are proportionate to the anticipated military advantage, it cannot apply a "human, emotional approach" to the situation.

On the other hand, situations occur sometimes, where emotions are the reason, why is the principle of proportionality or distinction violated. Anger, fear, and urge to avenge someone are emotions only human is capable of. Such emotions lead to a violation of the abovementioned principles.[29] To sum it up, machines, nor humans are perfect, and the principle of propor-

[26] SCHMITT, M. N.: *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics,* op. quoted, p. 19.

[27] PANIGAJ, J.: *Medzinárodné humanitárne právo a autonómne zbraňové systémy*, Master thesis, Pavol Jozef Šafárik University in Košice, 2019, p. 28.

[28] Human Rights Watch: *Losing humanity: The case against Killer Robots*, 2012, p. 33, ISBN 1-56432-964-X. Online: https://reliefweb.int/sites/reliefweb.int/files/resources/Losing%20Humanity%20The%20Case%20Against%20Killer%20Robots.pdf (quoted 8.11.2021).

[29] PANIGAJ, J.: *Medzinárodné humanitárne právo a autonómne zbraňové systémy*, op. quoted, p. 25.

tionality such perfection does not even require.[30] However, ultimately, machines cannot currently fully comply with these principles.

There is no doubt, it is necessary to find answers to many questions regarding autonomous weapon systems. The core of these questions often has an ethical nature, and also affects future development and use of these weapons. However, finding answers to all questions does not have to be necessarily the right solution. Regarding this issue, we could analogically apply the opinion of several academics concerning autonomous vehicles. Experts from Stanford University for example call attention to the fact that sticking too much to finding the right answers to the questions outlined can lead to a dead-end and unnecessarily delay technological development, which is supposed to save the lives of accident victims. Results of statistics also support questioning the importance of theoretical moral dilemmas. According to them the number of car accidents with ethical issues is negligible.[31]

13.2 Regulation of autonomous weapon systems

Current technological development shows us first generations of autonomous weapon systems will not be able to fully comply with the fundamental principles of international humanitarian law, especially in armed conflicts increasingly taking place in populated areas.

On the other hand, it is questionable whether the absolute and preemptive ban on autonomous weapons could be a solution since the development of artificial intelligence in connection to military industry occupies several world powers and they are investing billions in the development.[32] The potential of artificial intelligence´s militarization is huge. Not surprisingly, regulation of autonomous weapons is hampered by a group of countries such as China, the United States, and the Russian federation.[33] The statement of Russian president V. Putin is eloquent enough in this context: "*Whoever becomes the leader in this sphere will become the ruler of the world.*"[34] However, this

[30] SCHMITT, M. N.: *Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics,* op. quoted, p. 21.

[31] KOLAŘÍKOVÁ, L., HORÁK, F.: *Umělá inteligence & právo*, op. quoted, p. 150.

[32] PICCONE, T.: How can international law regulate autonomous weapons? BROOKINGS, 10.4.2018. Online: https://www.brookings.edu/blog/order-from-chaos/2018/04/10/how-can-international-law-regulate-autonomous-weapons/ (quoted 16.10.2021).

[33] BARTLETT, M.: The AI Arms Race in 2020. towards data science, 16.6.2020. Online: https://towardsdatascience.com/the-ai-arms-race-in-2020-e7f049cb69ac (quoted 1.11.2021).

[34] PICCONE, T.: How can international law regulate autonomous weapons?, op. quoted

should not serve as an argument to make autonomous weapon systems illegal per se along with the fact, that the robots should be compliant with the principles of international humanitarian law for example in remote areas.

The growing need for regulation of autonomous weapon systems stems, inter alia, from a recent recommendation of the International Committee of the Red Cross. Recommendations of this nature are rare.[35] It is also important to mention the efforts of the Group of Governmental Experts and the Convention on Certain Weapons in Geneva that is the focal point of the endeavor.[36] Some of the UN representatives were quite optimistic a year ago, that a framework of a regulatory mechanism should be accomplished within 2 years.[37] Though, the last discussions from august 2021 show regulation in this field is still resented by several countries. Such an approach is slowing the whole process down and preventing the consensus body from making headway.[38] But still, it is not contrary to the opinions of some authors. According to them, the global trend is to ban the development, production, and use of autonomous weapon systems capable of attacking without direct human intervention[39] (the European Parliament takes a similar view[40]), however this "trend" is hampered by several countries.

13.2.1 The fundamental concept in relation to the principles of distinction and proportionality

Although a general legal framework for this type of weapons exists in the form of the law of armed conflicts, its basic principles, and the issue of legal review of weapons[41] a new and specific international legal instrument is necessary. Several experts say we should take a similar approach as it was with the development of nuclear weapons. The progress cannot be stopped, but it

[35] SAUER, F.: Autonomy in weapons systems: playing catch up with technology. Humanitarian Law & policy, 29.9.2021. Online: https://blogs.icrc.org/law-and-policy/2021/09/29/autonomous-weapons-systems-technology/ (quoted 1.11.2021).

[36] *Ibidem.*

[37] BARTLETT, M.: The AI Arms Race in 2020, op. quoted

[38] SAUER, F.: Autonomy in weapons systems: playing catch up with technology, op. quoted

[39] ŠTEDROŇ, B. et al.: *Právo a umělá intelligence*, Plzeň: Aleš Čeněk, 2020, p. 126, ISBN 978-80-7380-803-7.

[40] European Parliament resolution of 12 September 2018 on autonomous weapon systems (2018/2752(RSP)), OJ EU C 433/86, 23.12.2019. Online: https://eur-lex.europa.eu/legal-content/SK/TXT/HTML/?uri=CELEX:52018IP0341&from=EN (quoted 16.10.2021).

[41] ANDERSON, K., REISNER, D., WAXMAN, M.: *Adapting the Law of Armed Conflict to Autonomous Weapon Systems*, op. quoted, p. 411.

is necessary to maintain its control from the beginning.[42] Whether it takes the form of an international treaty or a protocol to the Convention on Certain Conventional Weapons, considering the current development of artificial intelligence in conjunction with the principles of distinction and proportionality, in the light of the above it should regulate the following issues:

1. Geographic limitations – Current development shows urban battlefields or populated areas, in general, should be inaccessible for autonomous weapon systems regarding the number of factors, which would make it difficult for the machine to comply with the principles. The machine should be programmed only to operate in a specific area or even to be static as a defensive instrument similar to some of the current automated defensive mechanisms.

   Within this type of limitation, it would also be useful to talk about a pre-programmed basic framework for the mission to which the machine would be sent. However, in this case, we are slipping from autonomy to automation.

   When talking about limitations, the system could be set to react only to certain kinds of objects regarding the shape, size, etc. Furthermore, the systems could be programmed with determined limits of reasonable or disproportionate collateral damage for specific categories of objectives or situations. Such programming would always have to be done by the person before the actual activation and deployment of the machine in a particular area.[43]
2. Option of remote intervention – An option of a human operator to intervene if needed and even shut the machine down remotely or cancel the targeting (or/and initiating the attack) in real-time could be also a way of ensuring compliance with the principles of proportionality and distinction.[44] At the same time, the machine could be programmed to interrupt the attack in case it would reach a high percentage of quantified doubts based on real-time calculations.
3. Thorough testing of individual scenarios before deploying the machine with respect to the specifications of a mission. However, in the case of the ability of a machine to learn would be previous tests or results

[42] KOLAŘÍKOVÁ, L., HORÁK, F.: *Umělá inteligence & právo*, op. quoted, p. 132.

[43] PANIGAJ, J.: *Medzinárodné humanitárne právo a autonómne zbraňové systémy*, op. quoted, p. 29.

[44] ANDERSON, K., REISNER, D., WAXMAN, M.: *Adapting the Law of Armed Conflict to Autonomous Weapon Systems*, op. quoted, p. 407.

would in principle lose their meaning. Furthermore, similar issues arise concerning the so-called "deep learning" of the machine, which eventually leads to its unpredictability.[45]

4. Equipment limitations – To be compliant with the international humanitarian law would also rely on what weapons would be robot equipped with. The basic framework should be defined which would answer what weapon may be a specific autonomous system equipped with. To elaborate, a lot would depend on whether the system is static or movable, whether it is aerial, marine, or operates on the ground.
5. Similarly, to airplanes, an obligation to equip the machine with some kind of a black box would be in place. It could help in finding the answer "why" and "based on what" a malfunction occurred, or a law was violated.[46]

This is, of course, just a demonstrative calculation of essentials a new international legal instrument should contain. Strong ground rules for the use and development of autonomous systems could ensure their development is under control also in terms of preventing potential arms races or possible abuse. Last but not least, such a regulatory mechanism should help to ensure that their use is in line with international humanitarian law.

Conclusion

It is necessary to point out in respect to the current state of development of the autonomous weapon systems, that the absolute ban is not an answer, although many countries, organizations, and other entities call for it. Technological state shows autonomous weapon systems should be able to operate under international humanitarian law in isolated areas even nowadays. Such systems would have probably just defensive character, not that different from automated defensive systems. The absolute ban would be also difficult to enforce due to the massive potential of artificial intelligence´s militarization and efforts of several countries to be a leader in this field. Such an approach

[45] International Committee of the Red Cross: Ethics and autonomous weapon systems: An ethical basis for human control? 3.4.2018, p. 15. Online: https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control (quoted 7.11.2021).

[46] LEWIS, J.: The Case For Regulating Fully Autonomous Weapons. In: *Yale Law Journal*, vol. 124, no. 4, 2015, ISSN 0044-0094. Online: https://www.yalelawjournal.org/comment/the-case-for-regulating-fully-autonomous-weapons (quoted 16.10.2021).

hampers the efforts of the rest of the international society to agree to the regulation of autonomous weapons.

On the other hand, the technological development shows, that even if not now, autonomous weapon systems could be used even in "attack mode" without geographic limitations in the near future. Passages of the paper dealing with their potential future ability to be more powerful and paradoxically more "humane" means of leading an armed conflict, despite the absence of a human factor, emotions, and psychological aspect as such supports it.

To sum it up, the issue of the use of autonomous weapon systems should be properly adjusted and regulated on the international level. The author also dared to outline some basic elements that should be contained in a new international legal instrument in terms of meeting the requirements of the principles of distinction and proportionality. Nevertheless, discussions at the UN suggest that regardless of initial optimism, the path to reach the consensus on the regulatory mechanism will be largely tortuous.

From the content of this article is certain that part of the international community is slowly but still deviating from the popular concept in the form of the famous laws of robotics by I. Asimov, according to which the robot should not under any circumstances harm humans. However, whether this deviation is in favor of humanity as such is not that certain.

Bibliography

1. ANDERSON, K., REISNER, D., WAXMAN, M.: Adapting the Law of Armed Conflict to Autonomous Weapon Systems. In: *International Law Studies*, vol. 90, 2014, p. 386. Online: https://apps.dtic.mil/dtic/tr/fulltext/u2/a613290.pdf (quoted 16.10.2021).
2. BARTLETT, M.: The AI Arms Race in 2020. towards data science, 16.6.2020. Online: https://towardsdatascience.com/the-ai-arms-race-in-2020-e7f049cb69ac (quoted 1.11.2021).
3. CONNER-SIMONS, A., GORDON, R.: Wearable AI system can detect a conversation´s tone. MIT news, 2.1.2017. Online: https://news.mit.edu/2017/wearable-ai-can-detect-tone-conversation-0201 (quoted 10.10.2021).
4. DAVISON, N.: A legal perspective: Autonomous weapon systems under international humanitarian law. International committee of the red cross, 31.1.2018. Online: https://www.icrc.org/en/document/autonomous-weapon-systems-under-international-humanitarian-law (quoted 9.10.2021).
5. DOCHERTY, B.: The Need for and Elements of a New Treaty on Fully Autonomous Weapons. Human rights watch, 20.2.2020. Online: https://www.hrw.org/

news/2020/06/01/need-and-elements-new-treaty-fully-autonomous-weapons (quoted 10.10.2021).

6. EASON, P. M.: *Lethal Autonomous Weapon Systems: Reconciling the Myth of Killer Robots and the Reality of the Modern Battlefield*, thesis for the degree of Master of Arts in Humanities, Graduate School of Duke University, 2021. Online: https://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/23212/Eason_duke_0066N_16305.pdf?sequence=1 (quoted 16.10.2021).
7. European External Action Service (EEAS): Avoiding and Minimizing Collateral Damage in EU-led Military Operations Concept. 3.2.2016. Online: https://data.consilium.europa.eu/doc/document/ST-5785-2016-INIT/en/pdf (quoted 16.10.2021).
8. European Parliament resolution of 12 September 2018 on autonomous weapon systems (2018/2752(RSP)), OJ EU C 433/86, 23.12.2019. Online: https://eur-lex.europa.eu/legal-content/SK/TXT/HTML/?uri=CELEX:52018IP0341&from=EN (quoted 16.10.2021).
9. Human Rights Watch: *Losing humanity: The case against Killer Robots*, 2012, ISBN 1-56432-964-X. Online: https://reliefweb.int/sites/reliefweb.int/files/resources/Losing%20Humanity%20The%20Case%20Against%20Killer%20Robots.pdf (quoted 8.11.2021).
10. International Committee of the Red Cross: Ethics and autonomous weapon systems: An ethical basis for human control?. 3.4.2018, p. 15. Online: https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control (quoted 7.11.2021).
11. International committee of the red cross: ICRC's position on autonomous weapon systems. 12.5.2021. Online: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems (quoted 9.10.2021).
12. KLUČKA, J.: *Medzinárodné právo verejné (všeobecná a osobitná časť).* Bratislava: Wolters Kluwer, 2017, ISBN 978-80-8168-744-0.
13. KOLAŘÍKOVÁ, L., HORÁK, F.: *Umělá inteligence & právo.* Praha: Wolters Kluwer ČR, 2020, ISBN 9788075987839.
14. LEWIS, J.: The Case For Regulating Fully Autonomous Weapons. In: *Yale Law Journal*, vol. 124, no. 4, 2015, ISSN 0044-0094. Online: https://www.yalelawjournal.org/comment/the-case-for-regulating-fully-autonomous-weapons (quoted 16.10.2021).
15. Notice from the Federal Ministry of Foreign Affairs no. 168/1991 Coll. on the Additional Protocols I and II to the Geneva Conventions of 12.8.1949.
16. PANIGAJ, J.: *Medzinárodné humanitárne právo a autonómne zbraňové systémy*. Master thesis, Pavol Jozef Šafárik University in Košice, 2019.
17. PICCONE, T.: How can international law regulate autonomous weapons? BROOKINGS, 10.4.2018. Online: https://www.brookings.edu/blog/order-from-chaos/2018/04/10/how-can-international-law-regulate-autonomous-weapons/ (quoted 16.10.2021).

18. SAUER, F.: Autonomy in weapons systems: playing catch up with technology. Humanitarian Law & policy, 29.9.2021. Online: https://blogs.icrc.org/law-and-policy/2021/09/29/autonomous-weapons-systems-technology/ (quoted 1.11.2021).
19. SCHMITT, M. N.: Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics. In: *Harvard National Security Journal Feature*, 2013, p. 1, ISSN 2153-1358. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2184826 (quoted 10.10.2021).
20. SPICE, B.: Computer Reads Body Language. Carnegie Mellon University. 6.7.2017. Online: https://www.cmu.edu/news/stories/archives/2017/july/computer-reads-body-language.html (quoted 10.10.2021).
21. ŠTEDROŇ, B. et al.: *Právo a umělá intelligence*. Plzeň: Aleš Čeněk, 2020, p. 126, ISBN 978-80-7380-803-7.
22. WAGNER, M.: The Dehumanization of International Humanitarian Law: Legal, Ethical and Political Implications of Autonomous Weapon Systems. In: *Vanderbilt Journal of Transnational Law*, vol. 47, 2014, p. 1371, ISSN 0090-2594. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2541628 (quoted 10.10.2021).

Author details

Mgr. Juraj PANIGAJ
PhD. student
Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of International law and European law
juraj.panigaj@student.upjs.sk

14 DRONES IN ARMED CONFLICT: ANGEL OF SALVATION OR ANGEL OF DEATH?

Ľudmila Elbert

Abstract

The production of the drones grows rapidly in quality as well as in quantity, but how about legal response of their usage? Are the current rules of international law applicable and sufficient, or do we need new legal regulation of the employment of the drones during the armed conflict? As drones play also effective role in private sector, e.g., for the deliveries, their usage can be very controversial. They can be used in positive connotation, as they could save life during the armed conflict, but also in negative sense, as they also could be the "messenger of death". This contribution demonstrates these two sides of drone. Firstly, the positive side of drone as lifesaver will be illustrated on example of its employment during the spread of coronavirus. Secondly, negative side of drone will be demonstrated on example of its employment during the armed attack. But the main goal is to show that during the armed attack drones could be more than just killer robots.[1]

Introduction

The expansion of the drones or unmanned aerial systems, their development and employment, are growing rapidly in last 20 years. Drones are able to carry on the high-definition cameras, infrared cameras, electronic security systems, etc. They could be constructed of light materials as plastic or carbon fibers, what allows the construction of big but also very light weight drones able to carry on GPS systems and control remote systems.[2] The production of the drones grows rapidly in quality as well as in quantity, but how about legal response of their usage? Are the current rules of international law applicable and sufficient, or do we need new legal regulation of the employment of the drones during the armed conflict? As drones play also effective role in private sector, e.g., for the deliveries, their usage can be very controversial. They can be used in positive connotation, as they could save life during the armed conflict, but also in negative sense, as they also could be the "mes-

[1] This paper was prepared within the framework of the research project VEGA no. 1/0643/20 "Legal and Ethical Perspective of Artificial Intelligence."

[2] ESTRADA, M. A. R.: *The Uses of Drones in Case of Massive Epidemics Contagious Diseases Relief Humanitarian Aid: Wuhan-COVID-19 Crisis*. University of Malaya, Social Security Research Centre, 29.2.2020, p. 1. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3546547 (quoted 10. 11. 2021).

senger of death". This contribution demonstrates these two sides of drone. Firstly, the positive side of drone as lifesaver will be illustrated on example of its employment during the spread of coronavirus. Secondly, negative side of drone will be demonstrated on example of its employment during the armed attack. But the main goal is to show that during the armed attack drones could be more than just killer robots. Legal employment of the drone in this case needs to comply with the rules of international law, especially international humanitarian law and international human rights law.

14.1 The Drone as an Angel of Salvation

The employment of the drones does not need to serve just for targeting a military object, but the drones are very effectively used in private sector. The employment of the drones during ongoing spread of COVID-19 can be very good example of how drones can be a lifesaver. It can inspire its employment as the "good drones" during the armed conflict for the protection of or saving the life of soldiers or civilians, as an instrument of the first aid. Their effective usage has been recognized during the COVID-19 pandemic, when the drones help people in quarantine.[3] In time of spread of the highly contagious disease it is necessary to expect the restrictions on mobility and consequently restrictions on the delivery of goods and services.[4] The COVID-19 pandemic may be identified as a biological disaster,[5] and as a response to this disaster,

[3] SAGAR, M.: How Drones are Assisting Government in China Fight COVID-19. opengovasia.com, 11.3.2020. Online: https://opengovasia.com/how-drones-are-assisting-government-in-china-fight-covid-19/ (quoted 10. 11. 2021).

[4] International Transport Forum: COVID-19 Transport Brief- Drones in the Era of Coronavirus. 19.6.2020. Online: https://www.itf-oecd.org/sites/default/files/drones-covid-19.pdf (quoted 10. 11. 2021); SHARMA, M.: How drones are being used to combat COVID-19. Geospatial World, 20.4.2020. Online: https://www.geospatialworld.net/blogs/how-drones-are-being-used-to-combat-covid-19/ (quoted 10. 11. 2021); UNSW Sydney: 'Pandemic drones': useful for enforcing social distancing, or for creating a police state?. 1.4.2020. Online: https://www.unsw.edu.au/news/2020/04/-pandemic-drones---useful-for-enforcing-social-distancing--or-fo (quoted 10. 11. 2021).

[5] BÓDIŠOVÁ, L.: Uplatňovanie suverenity štátov v procese odstraňovania následkov prírodných a priemyselných katastrof. In: BECKOVÁ, D., GIERTL, A. (eds.): *Miesto, úloha a význam vnútroštátneho práva pri zabezpečovaní plnenia záväzkov vyplývajúcich z medzinárodného práva a európskeho práva*. Košice: University of Pavol Jozef Šafárik in Košice, 2018, p. 238, ISBN 978-80-8152-595-7. Online: https://unibook.upjs.sk/img/cms/2018/pravf/miesto-uloha-a-vyznam-naweb.pdf?fbclid=IwAR3Y8sJHmfFH_mD7RZAeO7RXdMAA6iBjc6PXEfMv9qPmBm7MlFedrf1R75k (quoted 10. 11. 2021).

the drones can be used for the aerial monitoring of the spread of the disease, delivery of goods, or evaluation of the post-pandemic situation.[6]

One of the famous images of the drone used during the COVID-19 pandemic is one showing people on the balcony during the lockdown in China.[7] The drone was equipped with the infrared camera to measure the temperature of the people at their home for the purposes of quarantine. This mean of the usage of drone can be also effective in the need of the emergency call or for the medical staff to identify new potential cases of the disease. But the temperature check by the drone was also the subject of critique from the human rights activists (civil liberties groups) in Florida,[8] where the Daytona Police Department used the drones for monitoring of social distancing and breaking the curfew in relation to the homeless people. Activists pointed out that monitoring by "coronavirus detecting drones" can harm the individual human rights not only of homeless people, but also people who try to reach out to the homeless in order to help them, especially motorists who stop to feed them.

But drones can do much more than just temperature checking. For example, couple of states[9] employ the drones for sanitizing over the villages, as they can by filled with disinfectants and can cover more ground in less time and also faster than doing so with traditional methods. When equipped with loudspeakers, they can warn about quarantine curfew, make public announcements to keep people indoors, take necessary precautions, keep social distancing and wear a mask if stepping outside from home. They can also carry out banners to educate people about virus precautions. Also, the lighting drones were employed during the construction of two temporary hospitals in China, as they can operate for 24 hours a day and allowed the construction work to continue also during the night.[10] The quick construction was managed also with the help of the satellite-based technology provided high-precision positioning and accelerated the construction immensely. The drones can help with the delivery of personal protective equipment, medical sam-

[6] See: ESTRADA, *supra* no. 2.

[7] See: SHARMA, *supra* no. 4.

[8] Renegade Tribune: Police Are Using Drones To Monitor The Homeless And Check People's Temperatures. 10.5.2020. Online: http://www.renegadetribune.com/police-are-using-drones-to-monitor-the-homeless-and-check-peoples-temperatures/ (quoted 10. 11. 2021).

[9] E.g., China, India, Colombia, Indonesia, Philippines, Chile, UAE, Spain. See: SHARMA *supra* no. 4.

[10] See: SAGAR, *supra* no. 3.

ples (from hospital to laboratories) and other equipment to medical campuses and providers.[11] They can very effectively help to avoid face-to-face contact to cut the risk of infection.

The indisputable utility of drones in the current health crisis may well accelerate their deployment and may lead to increased social acceptance of its use on daily basis.[12] Outside the pandemic, in private sector[13] the drones even now continually help to deliver medical supplies to rural hospitals in Rwanda and Ghana. Drones equipped with large engines with high power and complex hardware and software can fly for long distances.[14] They can supply the territories affected by war or disaster with food, water, medicine, or just basic internet connectivity systems. We have evidences that drones are employed to help firefighters, rescue squads and during search operations, they can help to allocate missing people in the snow, lost hikers, or they can bring the emergency supplies or life-saving devices like life vests, ropes, etc.[15] Also unarmed drones were employed during the peacekeeping operations of the UN in the eastern part of the Democratic Republic of Congo, in Mali or Central African Republic for the surveillance purposes as for the prior warnings that an ambush or an attack is about to happen.[16] Moreover, during or after war, the drones with the infrared cameras can easily detect small butterfly landmines.[17]

These cases of the employment of drones are not subject of the rules of international law, but subject of the regulation of the municipal law, as every state has a complete and exclusive sovereignty over the airspace above its ter-

[11] SHAPIRO, E.: How COVID-19 is Accelerating Robot and Drone Technology for use in Everyday Activities. Robotics Tomorrow, 21.7.2020. Online: https://www.roboticstomorrow.com/article/2020/07/how-covid-19-is-accelerating-robot-and-drone-technology-for-use-in-everyday-activities/15482(quoted 10. 11. 2021).

[12] International Transport Forum: COVID-19 Transport Brief- Drones in the Era of Coronavirus. 19.6.2020. Online: https://www.itf-oecd.org/sites/default/files/drones-covid-19.pdf (quoted 10. 11. 2021).

[13] For example, the drones produced by the California drone start-up Zipline. Online: https://flyzipline.com/global-healthcare/ (quoted 10. 11. 2021).

[14] See: ESTRADA, *supra* no. 2, p. 2.

[15] Royal Aeronautical Society: Life – saving drones. 21.3.2017. Online: https://www.aerosociety.com/news/life-saving-drones/ (quoted 10. 11. 2021).

[16] TAFIRENYIKA, M.: Drones are effective in protecting civilians. April 2016. Available online: https://www.un.org/africarenewal/magazine/april-2016/drones-are-effective-protecting-civilians (quoted 10. 11. 2021).

[17] EPATKO, L.: These drone projects are saving lives and protecting nature. PBS NEWS HOUR, 10.7.2018. Online: https://www.pbs.org/newshour/world/these-drone-projects-are-saving-lives-and-protecting-nature (quoted 10. 11. 2021).

ritory.[18] For example, according to the Slovak legal system, drones are governed by the act no. 143/1998 about civil aviation (aviation act).[19] Following the para. 7 of this act, the drones as unmanned aerial vehicles need to operate in the airspace under the conditions which ensure the flight safety.[20] In case we want to use the drone for the video-recording or photos from the above, the act no. 215/2004 about the protection of classified (confidential) information[21] need to be applied. This act considers this activity as aerial sensing (letecké snímkovanie) which need to be approved by the Ministry of Defence of Slovak Republic.[22] Moreover, if the aerial sensing falls within the scope of "aerial work" defined in the para. 44 of the above-mentioned Aviation act, Transport Authority of Slovak Republic needs to grant a permission. Transport Authority[23] governs the conditions of unmanned aerial vehicles by its decision no. 2/2019.[24] This decision does not govern the drones of HALE category[25] as well as kites and free balloon without the crew. Also, within

[18] Art. 1 Convention on International Civil Aviation (1944, Chicago). Online: https://www.icao.int/publications/Documents/7300_cons.pdf (quoted 10. 11. 2021); Oznámenie Ministerstva zahraničných vecí Slovenskej republiky č. 196/1995 Z. z. o uskutočnení notifikácie sukcesie Slovenskej republiky do Dohovoru o medzinárodnom civilnom letectve (vyhláška ministra zahraničných vecí č. 147/1947 Zb.).

[19] Zákon č. 143/1998 Z. z. o civilnom letectve (letecký zákon) a o zmene a doplnení niektorých zákonov.

[20] VOJČÍK, P.: Drony a súkromné právo. In: SUCHOŽA, J., HUSÁR, J., HUČKOVÁ, R. (eds.): *Právo, obchod, ekonomika IX*. Košice: Univerzita P. J. Šafárika v Košiciach, 2019, p. 555, ISBN 978-80-8152-776-0. Online: https://poe.pravo.upjs.sk/wp-content/documents/POE_2019_zbornik.pdf (quoted 10. 11. 2021).

[21] § 63, 64 zákona č. 215/2004 Z. z. o ochrane utajovaných skutočností a o zmene a doplnení niektorých zákonov.

[22] See: Právna regulácia dronov v Slovenskej republike. 26.1.2018. Online: https://www.epravo.sk/top/clanky/pravna-regulacia-dronov-v-slovenskej-republike-3941.html?mail (quoted 10. 11. 2021).

[23] Lietadlá spôsobilé lietať bez pilota. Online: http://letectvo.nsat.sk/letova-prevadzka/lietadla-sposobile-lietat-bez-pilota/ (quoted 10. 11. 2021).

[24] Rozhodnutie Dopravného úradu SR č. 2/2019 zo 14.11.2019, ktorým sa určujú podmienky vykonania letu lietadlom spôsobilým lietať bez pilota a vyhlasuje zákaz vykonania letu určitých kategórií lietadiel vo vzdušnom priestore Slovenskej republiky. Online: http://nsat.sk/wp-content/uploads/2019/11/R2-2019.pdf (quoted 10.11.2021). This decision distinguishes between: 1. Autonomous aircraft – as unmanned aircraft with the independence system of control, which does not permit the human intervention to control the aircraft during a flight, and 2. Remotely piloted aircraft – as unmanned aircraft with the remote control of the human pilot from the base station not on board of aircraft.

[25] Types of long endurance drones include HALE (high altitude long endurance) and MALE (medium altitude long endurance). MALE UAVs fly at altitudes of 10,000 to 30,000 feet, while HALE drones can operate even higher and are often capable of faster speed. See:

the European region, the European Union plays a huge legislative role when it comes to the use of drones,[26] even for the commercial or public usage, as there is need for the harmonization of the common conditions of the use of drones under the municipal law of the member states.[27]

The cases stated above proved that the drones employed in private sector are able to make a huge impact in emergency cases. But they also raise the reasonable doubts about the privacy and security of individuals and the protection of basic human rights. As the employment of drones in private sector is governed by municipal law, there is a responsibility of every state to regulate the usage of drones by clear and up to date rules in accordance with the technological development.

14.2 The Drone as an Angel of Death

Drones as unmanned aerial vehicles are employed also in time of armed conflict. Drone strikes mostly operated by the United States (but not only) represent a big challenge for the international law. Since the first usage of drones after 9/11, the armed drones rise a question of its compliance with the generally accepted meaning of the core legal concepts, as self-defence, armed attack, necessity, proportionality, combatant, civilian, armed conflict, and hostilities, etc.[28] The armed drones represent technological development designed to enable the delivery of force from the distance. Drones can be equipped with the infrared camera or color optical sensor, they can be armed

Long Endurance Drones & UAVs. Online: https://www.unmannedsystemstechnology.com/expo/long-endurance-drones/ (quoted 10. 11. 2021).

[26] See: Regulation (EU) 2018/1139 of the European Parliament and of the Council of 4 July 2018 on common rules in the field of civil aviation and establishing a European Union Aviation Safety Agency, and amending Regulations (EC) No 2111/2005, (EC) No 1008/2008, (EU) No 996/2010, (EU) No 376/2014 and Directives 2014/30/EU and 2014/53/EU of the European Parliament and of the Council, and repealing Regulations (EC) No 552/2004 and (EC) No 216/2008 of the European Parliament and of the Council and Council Regulation (EEC) No 3922/91 [2018] OJ EU L 212, 22.8.2018.

[27] European union distinguishes between Remotely Piloted Aviation Systems (RSAP) – drone where the aircraft is controlled by a human pilot from a distant location, and Unmanned drones – which are automatically programmed without being piloted, even remotely. See: Remotely Piloted Aviation Systems (RPAS) – Frequently Asked Questions. 8.4.2014. Online: https://www.easa.europa.eu/sites/default/files/dfu/Q&A_Commission_Drones.pdf (quoted 10. 11. 2021).

[28] BROOKS, R.: Drones and International Rule of Law. In: *Ethics & International Affairs,* no. 1, 2014, p. 83, ISSN 1747-7093. Online: <https://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?article=2296&context=facpub (quoted 10. 11. 2021).

with lasser guided missiles or bomb. They can be controlled manually or navigate autonomously along preplanned route.[29] As for an advantage of the drones, they are cheaper to produce, with low risk to live as they are designed to reduce the probability of killing anyone other than target. As the drones can provide high quality information about the target and its environment, they also allow for more refined assessment of the likely collateral damage to civilians and civilian objects.[30]

Especially in the war with terrorists, the use of conventional military force is not so effective (e.g., as wars in Afghanistan and Iraq show). The first U.S. drone strike is believed to have occurred in 2002, when drone killed four al-Qaeda members in Yemen. Since this year, the employment of drone strikes is more frequent every year, especially in the context of terrorism where everything is kept in secrecy. As Brooks[31] stated, U.S. officials appear to have suggested that the self-defence framework supplements the armed conflict framework, but nowadays they shifted entirely to a self-defence framework. With regard to the sovereignty, U.S. officials have repeatedly stated that they only use force inside the borders of sovereign state when that state either consents to the use of force or is "unwilling and unable" to take appropriate action to address the threat itself.

The dark side of the drones is its ability to carry on arms, or be by itself the autonomous weapon. The nature and the way of use of drones in the armed conflict varies, so it is very hard to make a general conclusion about the compliance of drones with the rules of the international law. This task is much harder as we take into account lack of information due to its confidentiality and also security reasons. For a long time, we had information only about the employment of the drones operated by the human pilot on the ground. But nowadays we are also aware of the employment of the drones as fully autonomous lethal weapons systems as programmed to attack targets without

[29] SCHMITT, M.N.: Drone attacks under the jus ad bellum and jus in bello: Clearing the "Fog of Law". In: SCHMITT, M. N., ARMINATSU, L., MCCORMACK, T. (eds.): *Yearbook of International Humanitarian Law. Volume 13, 2010,* T.M.C. A SSER PRESS, The Hague, The Netherlands www.asserpress.nl, 2011, p. 513, ISBN 978-90-6704-811-8. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1801179 (quoted 10. 11. 2021).

[30] *Ibidem*.

[31] See: BROOKS, *supra* no. 28, p. 90.

requiring data connectivity between the operator and the munition: in effect, a true "fire, forget and find" capability.[32]

14.2.1 Drones as Autonomous Weapon Systems

Drones are used to be categorized as autonomous weapon systems, but is it right presumption? According to Davison,[33] an autonomous weapon system can be defined as any weapon system with autonomy in its critical functions, as a weapon system that can select (search for, detect, identify, track or select) and attack (use force against, neutralize, damage or destroy) targets without human intervention. Anderson and Waksman[34] define autonomous weapon system as a weapon system that, once activated, can select and engage targets without further intervention by a human operator. Concerns about the autonomous weapons, especially fully autonomous, were outlined also by Human Rights Watch in its report Losing Humanity: The Case against Killer Robots.[35] This report refers to the killer robots as fully autonomous weapons that could select and engage targets without human intervention. Autonomous weapons systems differ depending on the level of autonomy, which means the ability of systems to operate without human supervision. We can categorized them as: 1) human-in-the-loop weapons, where system can select targets and deliver force only with the human command, which are under control of human operator, 2) human-on-the-loop weapons, where system can select targets and deliver force under the oversight of a human operator who can override the systems actions, and 3) Human-out-of-the-loop weapons, where systems are capable of selecting targets and delivering force without any human input or interaction.[36] According to this categorization of the report of Human Rights Watch, fully autonomous

[32] See for example: Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council (S/2021/229).

[33] DAVISON, N.: *Autonomous weapon systems under international humanitarian law*. Online: https://www.icrc.org/en/download/file/65762/autonomous_weapon_systems_under_international_humanitarian_law.pdf (quoted 10. 11. 2021).

[34] ANDERSON, K., WAXMAN, M.C.: *Debating Autonomous Weapon Systems, Their Ethics, and Their Regulation under International Law*, Washington College of Law Research Paper No. 2017-21, written on 28.2.2017. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2978359 (quoted 10. 11. 2021).

[35] Losing Humanity: The Case against Killer Robots. 19.11.2012. Online: https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots (quoted 10. 11. 2021).

[36] Compare ANDERSON, WAKSMAN, *supra* no. 34, p. 1100 and Losing Humanity, *supra* no. 35.

weapons fall under category human-out-of-the-loop weapons, as well as human-on-the-loop weapons in case that they are in practice out-of-the-loop due to very limited supervision.[37] Wagner[38] differs between autonomous weapon system and remotely-operated or automated systems, as term autonomy refers to two characteristics. Firstly, the ability to operate independently and engage targets without being programmed to specifically target an individual object or person, secondly, the capability to make discretionary decisions. The autonomous weapons systems due to these characteristics are able to react independently to changing set of circumstances. According to Taddeo and Blanchard, four key aspects need to be considered as factors to define autonomous weapon systems, autonomy, adapting capabilities of autonomous weapon systems, human control and purpose of use.[39]

The drone can be defined by each of these categories according to the manner of its use, from the weapon system under the control of human operator, for example when the drone serves just as a launch vehicle for delivering bombs and missiles, to the fully autonomous weapon system independent in selection of the target and use of force. The assessment needs to be done on case-by-case basis with regard to the manner of its employment.

14.2.2 Drone and its Compliance with the International Law

Besides the general assessment of drones as autonomous weapon systems, every employment of the drone during the armed conflict has to be in compliance with the rules of international law, in particular *ius ad bellum* and *ius in bello*. The first one refers to the regime governing resort to military force and second one refers to the regime governing the conduct of armed force within the rules of international humanitarian law as well as international human rights law.

[37] See: ANDERSON, WAXMAN, *supra* no. 34.

[38] WAGNER, M.: Autonomous Weapon Systems. January 1, 2016. In: WOLFRUM, R. (ed.): *Max Planck Encyclopedia of Public International Law*, Oxford: Oxford University Press. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2786136 (quoted 10. 11. 2021).

[39] TADDEO, M., BLANCHARD, A.: *A Comparative Analysis of the Definitions of Autonomous Weapons Systems.* Online: https://documents.unoda.org/wp-content/uploads/2021/10/20210721-Autonomous-Weapon-Systems-Definitions-TO-SHARE.pdf (quoted 10. 11. 2021); See also: UNIDIR publications. Online: https://documents.unoda.org/wp-content/uploads/2021/08/UNIDIR-Publications-on-Lethal-Autonomous-Weapons-and-Military-Artificial-Intelligence_2021-07-28.pdf (quoted 10. 11. 2021).

Drone and ius ad bellum

In current international law, states are obliged by the principles of the Charter of United Nations (UN), in particular, to settle their international disputes by peaceful means (art. 2 para. 3) and to refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state (art. 2 para. 4). The UN Charter defines two exceptions from the prohibition of the use of force, self-defence under the art. 51 of the UN Charter and collective actions with the use of force under the Chapter VII of the UN Charter.

According to some authors,[40] the prohibition of the use of force is not applicable if state consents to the use of force on its territory by another state. State consent is important for the compliance with the prohibition of the use of force against non-state actors during the war on terrorism with the drones. Territorial state has to give consent to another state targeting non-state actors on the territorial state´s territory with drones. Following the Draft articles on responsibility of states for wrongful acts,[41] the consent of the state concerned falls within the circumstances precluding wrongfulness. Following the art. 20 of Draft articles, the consent has to be valid, has to be given freely and by the authorized entity, in advance or even at the time the conduct is occurring, and the conduct should remain within the limits of the consent given. As the commentary to Draft articles states, art. 26 makes it clear that none of the circumstances precluding wrongfulness can be relied on if to do so would conflict with a peremptory norm of general international law. And the use of force on the territory of other state with drone can be in contrast with the prohibition of aggression which International Law Commission refers to as having status of the peremptory norm.[42] But as the Commentary on art. 26

[40] HEYNS, CH., AKANDE, D., HILL-CAWTHORNE, L., CHENGETA, T.: The International Law Framework Regulating the Use of Armed Drones. In: *International & Comparative Law Quarterly*, vol. 65, no. 4, p. 791 et seq., ISSN 1471-6895. Online: https://www.cambridge.org/core/services/aop-cambridge-core/content/view/E92C0FCA200F667633B0C3686A9EDE3C/S0020589316000385a.pdf/div-class-title-the-international-law-framework-regulating-the-use-of-armed-drones-a-href-fns01-ref-type-fn-a-div.pdf (quoted 10. 11. 2021).

[41] Draft articles on Responsibility of States for Internationally Wrongful Acts, with commentaries 2001. Online: https://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf (quoted 10. 11. 2021).

[42] Peremptory norms of general international law (*ius cogens*), Text of the draft conclusions and draft annex provisionally adopted by the Drafting Committee on first reading (A/CN.4/L.936). Online: https://undocs.org/en/A/CN.4/L.936 (quoted 10. 11. 2021).

defines, the consent of a particular state may be relevant in applying some peremptory norms and in particular, a state may validly consent to a foreign military presence on its territory for a lawful purpose.

When the consent of state is missing, the use of force on the territory of other state is in compliance with the international law if engaged within the self-defence or collective measures under the Chapter VII of UN Charter. When it comes to drones, state can relate to instrument of self-defence defined by international customary law and art. 51 of the UN Charter. The self-defence is the most common instrument by which states justify their employment of the drones during the armed attack. Art. 51 defines the existence of the inherent right of individual or collective self-defence if an armed attack[43] occurs against a member of the UN. State exercising the right to self-defence has to immediately report the measures within the self-defence to the UN Security Council. But the self-defence has to meet further basic conditions, like principle of proportionality and necessity, and it has to be exercised with the purpose to halt and repel the armed attack.[44] Self-defence is applicable also in the emergency case of the immediate or imminent risk for the state, but once these threats of attack are solved, state needs to find other solution for the maintenance of its security and peace.

In case of employment of the drones, there is a need for consideration, if the group against the drones are employed is sufficiently disrupted so that it no longer poses an immediate or imminent threat to exact moment.[45] Justification of self-defence prior to an actual armed attack is very questionable.

[43] The term "armed attack" was defined by ICJ:
"it may be considered to be agreed that an armed attack must be understood as including not merely action by regular armed forces across an international border, but also 'the sending by or on behalf of a State of armed bands, groups, irregulars or mercenaries, which carry out acts of armed force against another State of such gravity as to amount to' (inter alia) an actual armed attack conducted by regular forces, 'or its substantial involvement therein'. This description, contained in Article 3, paragraph (g), of the Definition of Aggression annexed to General Assembly resolution 3314 (XXIX), may be taken to reflect customary international law... the concept of "armed attack" includes not only acts by armed bands where such acts occur on a significant scale but also assistance to rebels in the form of the provision of weapons or logistical or other support. Such assistance may be regarded as a threat or use of force, or amount to intervention in the internal or external affairs of other States. It is also clear that it is the State that is the victim of an armed attack which must form and declare the view that it has been so attacked." Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States of America). Merits, Judgment, I.C.J. Reports 1986, p. 14, para. 195.

[44] KLUČKA, J.: *Medzinárodné právo verejné (všeobecná a osobitná časť),* Košice: Wolters Kluwer, 2017, p. 231–234, ISBN 978-80-8168-744-0.

[45] See: HEYNS, AKANDE, HILL-CAWTHORNE, CHENGETA, *supra* no. 40, p. 801.

In this case, self-defence is acceptable as long as the threatened attack is imminent, no other means would deflect it and the action is proportionate.[46] There is also question of legality of the self-defence exercised against the non-state actors. They do not act on behalf of the territorial state, but they are the main target of drone strikes[47] on the territory of state not necessarily responsible for the acts of these actors.[48] Since the 9/11 events, state practice relating to the use of force abroad against non-state actors is now extensive and according to the broad acceptance and support for the actions of the United States and other countries in response to these events, it is suggested that state may use the force in self-defence on other state´s territory, where the state of self-defence has been a victim of the armed attack by non-state actors operating on the other state´s territory, even without attribution of the attack to this state.

Second exception of the prohibition of the use of force due to the UN Charter are coercive actions allowed by the UN Security Council of the United Nations under the art. 42 of the UN Charter. In case of a state posing a threat to other states or international order more generally, the Security Council may decide about the approval of any coercive action at all, including military action, against a state when it deems this necessary to maintain or restore international peace and security. This decision can be made whether the threat is occurring now or in future, whether it involves the state´s own actions or those of non-state actors it harbours or supports, whether it takes the form of an act or omission, an actual or potential act of violence or simply a challenge to the Council´s authority.[49] During such engagement of force within the Chapter VII of the UN Charter, together with the conventional weapons, drones are also being employed.

[46] A more secure world: our shared responsibility (Report of the High-level Panel on Threats, Challenges and Change, (A/59/565), para. 188. Online: https://undocs.org/A/59/565 (quoted 10. 11. 2021).

[47] For example: UAE implicated in lethal drone strike in Libya. 28.8.2020. Online: https://www.bbc.com/news/world-africa-53917791(quoted 10. 11. 2021); A Military Drone With A Mind Of Its Own Was Used In Combat, U.N. Says. 1.6.2021. Online: https://www.npr.org/2021/06/01/1002196245/a-u-n-report-suggests-libya-saw-the-first-battlefield-killing-by-an-autonomous-d (quoted 10. 11. 2021).

[48] Compare with the case *Military and Paramilitary Activities in and against Nicaragua* (Nicaragua v. United States of America). Merits, Judgment, I.C.J. Reports 1986, p. 14, para. 195, where the ICJ rejected this kind of responsibility of territorial state.

[49] See: A more secure world, *supra* no. 46, para. 193.

For the legal employment of the drone, not just the rules of the resort to use of force should be respected, but also rules governing the conduct of armed force.

Drone under the International Humanitarian Law

Drones can be employed in the context of international armed conflict, as well as non-international armed conflict. The distinction is based on the question who the parties of the conflict are. When the parties of the conflict are states, the conflict has a nature of international armed conflict in accordance with the common art. 2 of Geneva Conventions. But when the conflict exists between state and non-state armed group or two or more such groups, these conflicts are recognized as non-international armed conflicts covered by the common art. 3 of Geneva Conventions. In this type of conflict, each party to the conflict (not just state party of the Geneva Conventions, but every party of conflict, like non-state actors) shall be bound to apply the provisions of this article as minimum requirement. Persons taking no active part in the hostilities shall in all circumstances be treated humanely, without any adverse distinction. The common art. 3 of the Geneva Conventions defines the acts which shall be prohibited at any time and in any place. In particular, two of these acts can be violated by the employment of drone, either under the operation of human or as autonomous weapon. These include violence to life and person, in particular murder of all kinds, mutilation, cruel treatment and torture, and passing of sentences and carrying out of executions without previous judgment pronounced by a regularly constituted court, affording all the judicial guarantees which are recognized as indispensable by civilized peoples. Legal employment of the drone during the armed attack depends on its compliance with these rules. As the drone strikes are targeted mainly on the members of terrorist group, we need to consider conditions upon which they can be defined as a party of the conflict of non-international nature as necessity for the application of the rules of international humanitarian law. These conditions were defined by case-law as an intensity of the conflict and organization of the parties of the conflict. The armed group (e.g., terrorist group) should constitute the party of non-international conflict only if it is sufficiently organized and the intensity of conflict is higher than just internal disturbances and tensions.[50] These conditions need to be considered on

[50] See: *Prosecutor v Tadić*, Case No. IT-94-1-T, 7 May 1997, para. 562. Online: https://www.icty.org/x/cases/tadic/tjug/en/tad-tsj70507JT2-e.pdf (quoted 10. 11. 2021); Art. 1 (2) of the

case-by-case basis and if they are not met, the rules of the international humanitarian law are not applicable. In case that there is no non-international armed conflict and there is no consent of territorial state for the use of force at its territory, targeted individuals will be classified as civilians as they are not considered combatants in the international armed conflict between states.[51] Only exemption for the lawfulness of the drone strike against the civilians would be a case, when the civilians would participate directly in hostilities.[52]

Either during the international armed conflict, or non-international armed conflict, the parties of the conflicts need to comply with the principles of distinction and proportionality, as well as precautions in attack. Principle of distinction[53] focuses on the protection of civilian population with for example the prohibition of indiscriminate attacks and differentiation between civilian population and combatants and between civilian objects and military objectives. In relation to drones it could be stated that they can employ precise guided munitions, e.g., laser-guided, so the drone should not be assumed as indiscriminate means of warfare. They are, on the contrary, able of much more efficient targeting than many other commonly employed weapon systems.[54] In addition, within the principle of proportionality[55] the states shall refrain from any attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combina-

Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), 8 June 1977.

[51] Art. 50 (1) of the Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977.

[52] Art. 51 (3) of the Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977; and Art. 13 Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), 8 June 1977.

[53] Compare: Art. 48, 51, 52 of the Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977; and Art. 13 Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), 8 June 1977.

[54] See: SCHMITT, *supra* no. 29, p.10.

[55] Compare: Art. 57 (2) (a) (iii) and Art. 51 (5)(b) of the Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977.

tion thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.

Consideration of these two principles within the employment of the drones leads us to the partial conclusion that we need to distinguish between the drone operated by a human and drone as autonomous weapon system able to choose target and decide about the engagement of use of force, because the first one will be treated as any other common weapon system operated by human, but the second one can be considered as much more suspected of the breach of the principles of distinction and prohibition. The compliance with these principles will depend on an absolutely precise adjustment of the system of automation of drone as autonomous weapon system.

Drone under the International Human Rights Law

In times of armed conflict, not just international humanitarian law, but also international human rights law is applicable.[56] State has a primary responsibility to refrain from the violating the rights guaranteed by applicable human rights treaties as long as state has a jurisdiction to do so.[57] The basic rule of international human rights law which is in contrary to the consequences of armed conflict is the right to life defined in art. 6 of International Covenant on Civil and Political Rights (within the European region also art. 2 of the European Convention of Human Rights[58]). Previous article 4 in para. 2 expressively prohibits a derogation of the right to life. But during the war lives are ended. According to the mentioned art. 4, measures derogating from the Covenant are allowed only if and to the extent that the armed conflict constitutes a threat to the life of the nation. As art. 15 of the European Convention of Human Rights states, in time of war or other public emergency threatening the life of the nation, any High Contracting Party may take measures derogating from its obligations under the Convention. No derogation

[56] *"The Court observes that the protection of the International Covenant of Civil and Political Rights does not cease in times of war, except by operation of Article 4 of the Covenant whereby certain provisions may be derogated from in a time of national emergency. Respect for the right to life is not, however, such a provision. In principle, the right not arbitrarily to be deprived of one's life applies also in hostilities."* See: *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 1. C.J. Reports 1996, p. 226, para. 25.

[57] FRAU, R.: Unmanned Military Systems and Extraterritorial Application of Human Rights Law. In: *Groningen Journal of International Law,* vol. 1, no. 1, p. 3. 1.1.2013. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2264241 (quoted 10. 11. 2021).

[58] Convention for the Protection of Human Rights and Fundamental Freedoms, 1950 (European Convention on Human rights).

from the obligation of the right to life shall be made, except of deaths resulting from lawful acts of war. As such, the right to life under human rights law form a standard legal norm applicable to the protection of people affected by drone strikes.[59]

Within the employment of the drone, there is also discussion about the question of the responsibility of state operating the drone strikes according to the violation of the human rights, in particular right to life, on the territory of other state, as for its extraterritorial actions. Following the case law of the International Court of Justice,[60] the jurisdiction of States is primarily territorial, it may sometimes be exercised outside the national territory. Considering the object and purpose of the International Covenant on Civil and Political Rights, it would seem natural that even if state exercises its jurisdiction outside its national territory, States Parties to the Covenant should be bound to comply with its provisions. Even when the state operating the drone strikes has no territorial control and jurisdiction over territory of other state where the drones are employed, for the purposes of extraterritorial application of human rights treaties this state exercises authority or control over specific individuals abroad.[61] Frau[62] suggests that the use of drones leads to effective control and jurisdiction if state exercises jurisdiction in waiting, meaning an extraterritorial situation in which a state may exercise all or some of the public power normally exercised by the government within an instant. He also points out that the employment of the drone may pose the violation of the prohibition of torture or inhuman treatment,[63] in case of mental suffering when individual finds out that he is subject of monitoring by drone and he must fear an attack on his life anytime, or violation of the right to liberty and security,[64] if the monitoring by the drone results in confinement in particular restricted area. Due to these facts, we need to conclude in part once again that the manner of the drone´s employment is crucial also for the compliance with the rules of international human rights law.

[59] See: HEYNS, AKANDE, HILL-CAWTHORNE, CHENGETA, *supra* no. 40, p. 819.

[60] *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*, Advisory Opinion, I.C.J. Reports 2004, p. 136, para. 109.

[61] See: HEYNS, AKANDE, HILL-CAWTHORNE, CHENGETA, *supra* no. 40, p. 823–824.

[62] See: FRAU, *supra* no. 57, p. 10–11.

[63] Art. 7 of the International Covenant on Civil and Political Rights, Art. 3 of the European Convention on Human rights.

[64] Art. 9 of the International Covenant on Civil and Political Rights, Art. 5 of the European Convention on Human rights.

The Discussion about the Ban of Drone as Lethal Autonomous Weapon

According to the International Committee of the Red Cross (ICRC), the overall guiding principle is that parties to an armed conflict are limited in their choice of weapons and how they use them. These limits include prohibitions on weapons that cause superfluous injury or unnecessary suffering or that are incapable of distinguishing between civilians or civilian objects and military targets. These are fundamental rules of international humanitarian law,[65] and it´s a traditional expression of the principles of proportion and distinction.

In the light of art. 36 of the Additional protocol I to Geneva Conventions of 1949,[66] each State Party has to ensure that the use of any new weapons, means or methods of warfare comply with the rules of international humanitarian law. To make this assessment easier, the ICRC prepared the Guide to the Legal Review of New Weapons, Means and Methods of Warfare.[67] However, for the compliance with art. 36, every state should have its own so called review mechanism. This is also the case of the drones as new kind of weapons as they are constructed as fully autonomous weapon systems.

Major role within the consideration of specific problems or weapon play the High Contracting Parties to the CCW,[68] as they affirmed in the Convention´s preamble the need to continue the codification and progressive development of the rules of international law applicable in armed conflict. During the meeting of the Group of Governmental Experts on lethal Autonomous Weapon Systems in August 2021, ICRC[69] expressed its view that an urgent and effective international response is needed to address the

[65] Review of new weapons. November 30, 2011. Online: https://www.icrc.org/en/document/review-new-weapons (quoted 10. 11. 2021).

[66] Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977.

[67] A Guide to the Legal Review of New Weapons, Means and Methods of Warfare. 10.6.2020. Online: https://www.icrc.org/en/publication/0902-guide-legal-review-new-weapons-means-and-methods-warfare-measures-implement-article (quoted 10. 11. 2021); for more see: Legal review of new weapons. 21.5.2021. Online: https://www.icrc.org/en/document/new-weapons-factsheet (quoted 10. 11. 2021).

[68] Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which may be deemed to be Excessively Injurious or to have Indiscriminate Effects (with Protocols I, II, II as amended, III, IV and V). Geneva, 10 October 1980. Online: https://legal.un.org/avl/ha/cprccc/cprccc.html (quoted 10. 11. 2021).

[69] Statement of the International Committee of the Red Cross delivered at the Convention on Certain Conventional Weapons (CCW) before the Group of Governmental Experts on Lethal Autonomous Weapons Systems – 3–13 August 2021, Geneva. Online: https://www.

serious risks posed by autonomous weapon systems, which stem from the process by which autonomous weapon systems function. The concern of the ICRC is based on understanding that the autonomous weapon systems select and apply force to targets without human intervention, so the owner of these systems does not choose the specific target, precise time or place of the application of that force.[70] ICRC calls for the responsibility of the High Contracting Parties to the CCW as they have an opportunity to make a progress in the clarifying, considering and developing the normative and operational framework for autonomous weapon systems.[71] State parties, especially from the group of states actively developing and employing the drones as autonomous weapons, are not willing to agree with the ban of autonomous weapons systems with the lethal force as they are pointing out the missing definition of the autonomous weapons systems and also different opinions of states on the appropriate level of human control under the autonomous weapons systems.[72] From the beginning of the discussions about fully autonomous weapons employed during the armed attack, the ICRC calls for the human centered approach[73] to any use of the technologies with the autonomy in armed conflict.

---

icrc.org/en/document/autonomous-weapons-icrc-recommends-new-rules (quoted 10. 11. 2021).

[70] In the relation to the (fully) autonomous weapons systems, the ICRC stated: "Worryingly, the use of artificial intelligence and machine learning software to control the critical functions of selecting and applying force is being increasingly explored, which would exacerbate the already difficult task that users have in anticipating and limiting the effects of an autonomous weapon system."

[71] ICRS offered recommendations for all states for the new legally binding rules to regulate autonomous weapon systems as they: a) can have indiscriminate effects, b) can target human beings, c) can be employed without any limitation. In: ICRC position on autonomous weapon systems, ICRC, Geneva, 12 May 2021. Online: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems (quoted 10. 11. 2021); for the upcoming meeting of the Group of Governmental Experts on Lethal Autonomous Weapons in December 2021, see: Views and recommendations of the ICRC for the Sixth Review Conference of the Convention on Certain Conventional Weapons. November 8, 2021. Online: https://www.icrc.org/en/document/icrc-sixth-review-conference-convention-certain-conventional (quoted 10. 11. 2021).

[72] For more see: ELBERT, Ľ.: Vnútroštátne snahy o reguláciu autonómnych zbraňových systémov v súlade s medzinárodným právom. In: SUCHOŽA, J., HUSÁR, J., HUČKOVÁ, R. (eds.): *Právo, obchod, ekonomika*. Košice: Univerzita P. J. Šafárika v Košiciach, 2020, p. 87–89, ISBN 978-80-8152-931-3. Online: https://unibook.upjs.sk/sk/pravo/1440-pravo-obchod-ekonomika (quoted 10. 11. 2021).

[73] Artificial intelligence and machine learning in armed conflict: A human-centered approach. 6.6.2019. Available online: https://www.icrc.org/en/document/artificial-intelli-

Conclusion

During an armed attack, drones could be employed as an instrument of salvation or death. It´s up to the decision of human operator if the drone will be employed as lifesaver or messenger of death. Drones could be a lifesaver as they are able to detect missing person or landmines for its destruction. Main rules for the employment of the drones in private sector fall within the municipal law. In case of the employment of the drone as a weapon system or weapon, it has to be in accordance with the rules of the international law. Drones employed as just vehicle for delivering bombs or missiles[74] should be relatively uncontroversial from the perspective of the international humanitarian law because they are ultimately under the full control of human operators.[75] More controversial and problematic from the international law perspective are the autonomous weapon systems, as they are independent from human operator and their behavior can be very unpredictable. Thus, it is very problematic to ensure their compliance with the peremptory norms of international law, as well as principles of international humanitarian law or human rights law, or even international criminal law. The education of the operator who operate the drone during the armed attack or who pre-programmed the drone as autonomous weapon system plays a crucial role, as he has to be aware and be governed by the rules of international law applicable in time of armed conflict. During the consideration, there is a need to distinguish between the drone operated by human and drone as autonomous weapon system able to choose the target and decide about the engagement of use of force, because the first one will be treated like any other common weapon system operated by the human, but the second one can be considered as much more suspected of the breach of the principles of distinction and prohibition. The compliance with these principles will depend on an absolutely precise adjustment of the system of automation of drone as autonomous weapon system. Drones can employ lethal force across the borders

gence-and-machine-learning-armed-conflict-human-centred-approach (quoted 10. 11. 2021).

[74] O´CONNELL, M. E.: *Drones under International law*. International Debate Series. 8.10.2010. Online: https://law.wustl.edu/wp-content/uploads/2018/10/OConnellFullRemarksNov23.pdf (quoted 10. 11. 2021).

[75] LIU, H.-Y.: *Categorization and legality of autonomous and remote weapons systems.* In: International Review of the Red Cross, vol 94, no. 886, 2012, p. 5, ISSN 1607-5889. Online: https://www.icrc.org/en/doc/assets/files/review/2012/irrc-886-liu.pdf (quoted 10. 11. 2021).

very effectively so they can cause significant risks to the protection of life. Following international treaties of human rights and the institute of extraterritorial jurisdiction of state, every state should very carefully consider its obligations. Moreover, according to the international humanitarian law, if the person participates directly in hostilities, during international or non-international armed conflict, and he is targeted and killed by the drone strike, it does not violate his right to life. Very limited number of types of weapons or weapon systems are governed by their specific legal regulation and therefore it should be no surprise that we need to consider applicability of existed rules of the international law in order to regulate drones as just vehicle for delivering bombs or missiles or as fully autonomous weapon systems. We can conclude that the legal regulation is sufficient if the drone is employed under the human control. But in case of employment of the drone as fully autonomous weapon which is able to select the target and employ the use of force, there is a need to achieve the agreement about the ban for autonomous weapon systems or to develop new regulation in relation to the responsibility and other consequences of the employment of the drone as fully autonomous weapon. This just reaffirms the conclusion that the manner of use of drone is crucial for the assessment of its compliance with the existing rules of international law and states should maintain the human centered approach within the development and use of drones as fully autonomous weapons during the armed conflict.

Bibliography

1. A Guide to the Legal Review of New Weapons, Means and Methods of Warfare. 10.6.2020. Online: https://www.icrc.org/en/publication/0902-guide-legal-review-new-weapons-means-and-methods-warfare-measures-implement-article (quoted 10. 11. 2021); for more see: Legal review of new weapons. 21.5.2021. Online: https://www.icrc.org/en/document/new-weapons-factsheet (quoted 10. 11. 2021).
2. A Military Drone With A Mind Of Its Own Was Used In Combat, U.N. Says. 1.6.2021. Online: https://www.npr.org/2021/06/01/1002196245/a-u-n-report-suggests-libya-saw-the-first-battlefield-killing-by-an-autonomous-d (quoted 10. 11. 2021).
3. A more secure world: our shared responsibility (Report of the High-level Panel on Threats, Challenges and Change, (A/59/565), para. 188. Online: https://undocs.org/A/59/565 (quoted 10. 11. 2021).
4. ANDERSON, K., WAXMAN, M.C.: *Debating Autonomous Weapon Systems, Their Ethics, and Their Regulation under International Law*, Washington College of Law

Research Paper No. 2017-21, written on 28.2.2017. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2978359 (quoted 10. 11. 2021).

5. Art. 1 Convention on International Civil Aviation (1944, Chicago). Online: https://www.icao.int/publications/Documents/7300_cons.pdf (quoted 10. 11. 2021).
6. Artificial intelligence and machine learning in armed conflict: A human-centered approach. 6.6.2019. Available online: https://www.icrc.org/en/document/artificial-intelligence-and-machine-learning-armed-conflict-human-centred-approach (quoted 10. 11. 2021).
7. BECKOVÁ, D., GIERTL, A. (eds.): *Miesto, úloha a význam vnútroštátneho práva pri zabezpečovaní plnenia záväzkov vyplývajúcich z medzinárodného práva a európskeho práva*. Košice: University of Pavol Jozef Šafárik in Košice, 2018, ISBN 978-80-8152-595-7. Online: https://unibook.upjs.sk/img/cms/2018/pravf/miesto-uloha-a-vyznam-naweb.pdf?fbclid=IwAR3Y8sJHmfFH_mD7RZAeO7RXdMAA6iBjc6PXEfMv9qP-mBm7MlFedrf1R75k (quoted 10. 11. 2021).
8. BROOKS, R.: Drones and International Rule of Law. In: *Ethics & International Affairs,* no. 1, 2014, p. 83, ISSN 1747-7093. Online: <https://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?article=2296&context=facpub (quoted 10. 11. 2021).
9. Convention for the Protection of Human Rights and Fundamental Freedoms, 1950 (European Convention on Human rights).
10. Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which may be deemed to be Excessively Injurious or to have Indiscriminate Effects (with Protocols I, II, II as amended, III, IV and V). Geneva, 10 October 1980. Online: https://legal.un.org/avl/ha/cprccc/cprccc.html (quoted 10. 11. 2021).
11. DAVISON, N.: *Autonomous weapon systems under international humanitarian law*. Online: https://www.icrc.org/en/download/file/65762/autonomous_weapon_systems_under_international_humanitarian_law.pdf (quoted 10. 11. 2021).
12. Draft articles on Responsibility of States for Internationally Wrongful Acts, with commentaries 2001. Online: https://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf (quoted 10. 11. 2021).
13. EPATKO, L.: These drone projects are saving lives and protecting nature. PBS NEWS HOUR, 10.7.2018. Online: https://www.pbs.org/newshour/world/these-drone-projects-are-saving-lives-and-protecting-nature (quoted 10. 11. 2021).
14. ESTRADA, M. A. R.: *The Uses of Drones in Case of Massive Epidemics Contagious Diseases Relief Humanitarian Aid: Wuhan-COVID-19 Crisis*. University of Malaya, Social Security Research Centre, 29.2.2020. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3546547 (quoted 10. 11. 2021).
15. FRAU, R.: Unmanned Military Systems and Extraterritorial Application of Human Rights Law. In: *Groningen Journal of International Law,* vol. 1, no. 1, p. 3. 1.1.2013. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2264241 (quoted 10. 11. 2021).

16. HEYNS, CH., AKANDE, D., HILL-CAWTHORNE, L., CHENGETA, T.: The International Law Framework Regulating the Use of Armed Drones. In: *International & Comparative Law Quarterly*, vol. 65, no. 4, p. 791 et seq., ISSN 1471-6895. Online: https://www.cambridge.org/core/services/aop-cambridge-core/content/view/E92C0FCA200F667633B0C3686A9EDE3C/S0020589316000385a.pdf/div-class-title-the-international-law-framework-regulating-the-use-of-armed-drones-a-href-fns01-ref-type-fn-a-div.pdf (quoted 10. 11. 2021).
17. ICRC position on autonomous weapon systems, ICRC, Geneva, 12 May 2021. Online: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems (quoted 10. 11. 2021).
18. International Covenant on Civil and Political Rights.
19. International Transport Forum: COVID-19 Transport Brief – Drones in the Era of Coronavirus. 19.6.2020. Online: https://www.itf-oecd.org/sites/default/files/drones-covid-19.pdf (quoted 10. 11. 2021).
20. International Transport Forum: COVID-19 Transport Brief- Drones in the Era of Coronavirus. 19.6.2020. Online: https://www.itf-oecd.org/sites/default/files/drones-covid-19.pdf (quoted 10. 11. 2021).
21. KLUČKA, J.: *Medzinárodné právo verejné (všeobecná a osobitná časť),* Košice: Wolters Kluwer, 2017, ISBN 978-80-8168-744-0.
22. *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*, Advisory Opinion, I.C.J. Reports 2004.
23. *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 1. C.J. Reports 1996.
24. Lietadlá spôsobilé lietať bez pilota. Online: http://letectvo.nsat.sk/letova-prevadzka/lietadla-sposobile-lietat-bez-pilota/ (quoted 10. 11. 2021).
25. LIU, H.-Y.: *Categorization and legality of autonomous and remote weapons systems*. In: International Review of the Red Cross, vol 94, no. 886, 2012, p. 5, ISSN 1607-5889. Online: https://www.icrc.org/en/doc/assets/files/review/2012/irrc-886-liu.pdf (quoted 10. 11. 2021).
26. Long Endurance Drones & UAVs. Online: https://www.unmannedsystemstechnology.com/expo/long-endurance-drones/ (quoted 10. 11. 2021).
27. Losing Humanity: The Case against Killer Robots. 19.11.2012. Online: https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots (quoted 10. 11. 2021).
28. Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States of America). Merits, Judgment, I.C.J. Reports 1986.
29. O´CONNELL, M. E.: *Drones under International law*. International Debate Series. 8.10.2010. Online: https://law.wustl.edu/wp-content/uploads/2018/10/OConnell-FullRemarksNov23.pdf (quoted 10. 11. 2021).

30. Oznámenie Ministerstva zahraničných vecí Slovenskej republiky č. 196/1995 Z. z. o uskutočnení notifikácie sukcesie Slovenskej republiky do Dohovoru o medzinárodnom civilnom letectve (vyhláška ministra zahraničných vecí č. 147/1947 Zb.).
31. Právna regulácia dronov v Slovenskej republike. 26.1.2018. Online: https://www.epravo.sk/top/clanky/pravna-regulacia-dronov-v-slovenskej-republike-3941.html?mail (quoted 10. 11. 2021).
32. *Prosecutor v Tadić*, Case No. IT-94-1-T, 7 May 1997, para. 562. Online: https://www.icty.org/x/cases/tadic/tjug/en/tad-tsj70507JT2-e.pdf (quoted 10. 11. 2021).
33. Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), 8 June 1977.
34. Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977.
35. Regulation (EU) 2018/1139 of the European Parliament and of the Council of 4 July 2018 on common rules in the field of civil aviation and establishing a European Union Aviation Safety Agency, and amending Regulations (EC) No 2111/2005, (EC) No 1008/2008, (EU) No 996/2010, (EU) No 376/2014 and Directives 2014/30/EU and 2014/53/EU of the European Parliament and of the Council, and repealing Regulations (EC) No 552/2004 and (EC) No 216/2008 of the European Parliament and of the Council and Council Regulation (EEC) No 3922/91 [2018] OJ EU L 212, 22.8.2018.
36. Remotely Piloted Aviation Systems (RPAS) – Frequently Asked Questions. 8.4.2014. Online: https://www.easa.europa.eu/sites/default/files/dfu/Q&A_Commission_Drones.pdf (quoted 10. 11. 2021).
37. Renegade Tribune: Police Are Using Drones To Monitor The Homeless And Check People's Temperatures. 10.5.2020. Online: http://www.renegadetribune.com/police-are-using-drones-to-monitor-the-homeless-and-check-peoples-temperatures/ (quoted 10. 11. 2021).
38. Review of new weapons. November 30, 2011. Online: https://www.icrc.org/en/document/review-new-weapons (quoted 10. 11. 2021).
39. Royal Aeronautical Society: Life - saving drones. 21.3.2017. Online: https://www.aerosociety.com/news/life-saving-drones/ (quoted 10. 11. 2021).
40. Rozhodnutie Dopravného úradu SR č. 2/2019 zo 14.11.2019, ktorým sa určujú podmienky vykonania letu lietadlom spôsobilým lietať bez pilota a vyhlasuje zákaz vykonania letu určitých kategórií lietadiel vo vzdušnom priestore Slovenskej republiky. Online: http://nsat.sk/wp-content/uploads/2019/11/R2-2019.pdf (quoted 10.11.2021).
41. SAGAR, M.: How Drones are Assisting Government in China Fight COVID-19. opengovasia.com, 11.3.2020. Online: https://opengovasia.com/how-drones-are-assisting-government-in-china-fight-covid-19/ (quoted 10. 11. 2021).
42. SCHMITT, M. N., ARMINATSU, L., MCCORMACK, T. (eds.): *Yearbook of International Humanitarian Law. Volume 13, 2010,* T.M.C. A SSER PRESS, The Hague, The

Netherlands www.asserpress.nl, 2011, ISBN 978-90-6704-811-8. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1801179 (quoted 10. 11. 2021).

43. SHAPIRO, E.: How COVID-19 is Accelerating Robot and Drone Technology for use in Everyday Activities. Robotics Tomorrow, 21.7.2020. Online: https://www.roboticstomorrow.com/article/2020/07/how-covid-19-is-accelerating-robot-and-drone-technology-for-use-in-everyday-activities/15482(quoted 10. 11. 2021).
44. SHARMA, M.: How drones are being used to combat COVID-19. Geospatial World, 20.4.2020. Online: https://www.geospatialworld.net/blogs/how-drones-are-being-used-to-combat-covid-19/ (quoted 10. 11. 2021).
45. Statement of the International Committee of the Red Cross delivered at the Convention on Certain Conventional Weapons (CCW) before the Group of Governmental Experts on Lethal Autonomous Weapons Systems – 3–13 August 2021, Geneva. Online: https://www.icrc.org/en/document/autonomous-weapons-icrc-recommends-new-rules (quoted 10. 11. 2021).
46. SUCHOŽA, J., HUSÁR, J., HUČKOVÁ, R. (eds.): *Právo, obchod, ekonomika IX*. Košice: Univerzita P. J. Šafárika v Košiciach, 2019, p. 555, ISBN 978-80-8152-776-0. Online: https://poe.pravo.upjs.sk/wp-content/documents/POE_2019_zbornik.pdf (quoted 10. 11. 2021).
47. SUCHOŽA, J., HUSÁR, J., HUČKOVÁ, R. (eds.): *Právo, obchod, ekonomika*. Košice: Univerzita P. J. Šafárika v Košiciach, 2020, p. 87–89, ISBN 978-80-8152-931-3. Online: https://unibook.upjs.sk/sk/pravo/1440-pravo-obchod-ekonomika (quoted 10. 11. 2021).
48. TADDEO, M., BLANCHARD, A.: *A Comparative Analysis of the Definitions of Autonomous Weapons Systems*. Online: https://documents.unoda.org/wp-content/uploads/2021/10/20210721-Autonomous-Weapon-Systems-Definitions-TO-SHARE.pdf (quoted 10. 11. 2021).
49. UNIDIR publications. Online: https://documents.unoda.org/wp-content/uploads/2021/08/UNIDIR-Publications-on-Lethal-Autonomous-Weapons-and-Military-Artificial-Intelligence_2021-07-28.pdf (quoted 10. 11. 2021).
50. TAFIRENYIKA, M.: Drones are effective in protecting civilians. April 2016. Available online: https://www.un.org/africarenewal/magazine/april-2016/drones-are-effective-protecting-civilians (quoted 10. 11. 2021).
51. Text of the draft conclusions and draft annex provisionally adopted by the Drafting Committee on first reading (A/CN.4/L.936). Online: https://undocs.org/en/A/CN.4/L.936 (quoted 10. 11. 2021).
52. UAE implicated in lethal drone strike in Libya. 28.8.2020. Online: https://www.bbc.com/news/world-africa-53917791(quoted 10. 11. 2021).
53. UNSW Sydney: ‚Pandemic drones': useful for enforcing social distancing, or for creating a police state? 1.4.2020. Online: https://www.unsw.edu.au/news/2020/04/-pandemic-drones---useful-for-enforcing-social-distancing--or-fo (quoted 10. 11. 2021).

54. Views and recommendations of the ICRC for the Sixth Review Conference of the Convention on Certain Conventional Weapons. November 8, 2021. Online: https://www.icrc.org/en/document/icrc-sixth-review-conference-convention-certain-conventional (quoted 10. 11. 2021).
55. WOLFRUM, R. (ed.): *Max Planck Encyclopedia of Public International Law*, Oxford: Oxford University Press. Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2786136 (quoted 10. 11. 2021).
56. Zákon č. 143/1998 Z. z. o civilnom letectve (letecký zákon) a o zmene a doplnení niektorých zákonov.
57. Zákon č. 215/2004 Z.z. o ochrane utajovaných skutočností a o zmene a doplnení niektorých zákonov.
58. Zipline. Online: https://flyzipline.com/global-healthcare/ (quoted 10. 11. 2021).

Author details

JUDr. Ľudmila ELBERT, PhD.
Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of the International Law and European Law
ludmila.elbert@upjs.sk

15 CONTEMPORARY INTERNATIONAL LEGAL REGULATION OF ARTIFICIAL INTELLIGENCE IN HEALTH AND NURSING

Lucia Bakošová

Abstract

Artificial Intelligence (AI) is starting to play an important role in the current and future health and nursing. Shortages of medical professionals and ageing population will even accelerate the implantation of AI in this area. Diagnostics, clinical care, development of new drugs or provision of nursing care are the most notable areas where the AI is proving to be useful. However, the development and application of AI is affected by existing ethical and legal regulation. The aim of the paper is to analyse the contemporary international legal regulation of artificial intelligence used in health and nursing. Special attention is paid to the legal proposal and regulations on AI in health and nursing on the international universal, specialized level, as well as on regional level which reflect the specific legal aspects of this area. Though most of the legal aspects are addressed in soft-law instruments, over time, they are able to take on a legally binding form.[1]

Introduction

It has been 65 years, since the notion "artificial intelligence" was introduced at the 1956 Dartmouth conference.[2] The aim of the study which was conducted at the conference was to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. Furthermore, the participants at the conference attempted to find out how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.[3] Despite the fact that the conference was not successful, it initiated further interest and research in this area. Back then, the general population probably did not envision that systems, or robots powered by AI will assist humans in various

[1] This paper was prepared within the framework of the research project VEGA no. 1/0643/20 "Legal and Ethical Perspective of Artificial Intelligence".

[2] Officially known as the Dartmouth Summer Research Project on Artificial Intelligence.

[3] McCARTHY, J., MINSKY, M., ROCHESTER, N., SHANNON, C.: A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. In: *AI Magazine*, Vol. 27, No. 4, 2006, ISSN: 0738-4602, p. 12–13.

areas of our lives. For the past years, the international community is eagerly observing the implementation of AI in various areas of health and nursing. One of the reasons is the increase of ageing population which requires greater amount of health and nursing care,[4] as well as deficiency of medical staff around the world.[5] On a positive note, a handful of studies conducted around world confirm the great advantages of AI in health and nursing, especially in diagnostics, clinical care or drug development. However, everything has its advantages and disadvantages. The same applies to AI. Although technological companies are accelerating their research on new AI systems, the international community, as well as individual states are confronted with the absence of unified ethical standards and legal framework. It may seem that legal regulation in this area should be left to individual states, however, we can not ignore the fact that the development and use of AI is not solely done by one entity in the territory of one state. This is true, for instance, in data collection and sharing on which the existence of AI depends.

Currently, there are numerous international initiatives, whether on universal, specialized or regional level, that may eliminate the current state of uncertainty and legal gaps. Although some progress can be seen, there is still an urging question: Are we a long way from an international legal framework on artificial intelligence in health and nursing? The aim of the paper is to analyse the contemporary international legal regulation of artificial intelligence used in health and nursing care. Special attention is paid to the legal proposal and regulations on AI in health and nursing care on the international universal, specialized, as well as on regional level which reflect the specific legal aspects of this area. Though most of the legal aspects are addressed in soft-law instruments, over time, they are able to take on a legally binding form.

15.1 Artificial intelligence in health and nursing

The term "artificial intelligence" or "AI" has numerous definitions that were drafted by States, academics, international organizations or groups of experts

[4] United Nations: World Population Ageing 1950 – 2050, New York, 2017, p. 2. Online: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA-2017_Report.pdf (quoted 11.11.2021).

[5] WHO: Addressing the 18 million health worker shortfall – 35 concrete actions and 6 key messages, 2019. Online: https://www.who.int/hrh/news/2019/addressing-18million-hw-shortfall-6-key-messages/en/ (quoted 11.11.2021).

in this area. For the purposes of this paper, we refer to the definition drafted by the High-Level Expert Group on AI, set out by the European Commission, in the *Ethics Guidelines for Trustworthy AI*. According to this guidelines, artificial intelligence systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal.[6] AI was first used in medicine in the 1970s when medical expert systems—based on Bayesian statistics and decision theory—diagnosed and recommended treatments for glaucoma and infectious disease.[7] Nowadays, the primary purpose of using AI in health and nursing is to improve the quality of diagnosis[8] and treatment,[9] increase the independence and social inclusion of vulnerable people. This may be done by robotic replacement of reduced or lost human organ functions, robots used for home rehabilitation or care for the elderly

[6] High-Level Expert Group on Artificial Intelligence: Ethics Guidelines for Trustworthy AI, Independent High-Level Expert Group on Artificial Intelligence Set Up by the European Commission, Brussels, 8 April 2019, p. 36. Online: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_EN.pdf (quoted 11.11.2021).

[7] SAMARGHITEAN, C., VIHINEN, M.: Medical expert systems. In: *Current Bioinformatics*, Vol. 3, No. 1, 2008, p. 56, ISSN 1574-8936. Online: https://www.eurekaselect.com/82021/article (quoted 11.11.2021).

[8] See for instance: HUANG, B., TIAN, S., ZHAN, N., MA, J., HUANG, Z., ZHANG, CH., et al.: Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: A retrospective multicentre study. In: *EBioMedicine*, Vol. 73, November 1, 2021, ISSN: 103631. Online: https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(21)00424-2/fulltext#seccesectitle0001 (quoted 11.11.2021); XIANG, Y., DU, J., FUJIMOTO, K., LI, F., SCHNEIDER, J., TAO, C.: Application of artificial intelligence and machine learning for HIV prevention interventions. In: *The Lancet HIV*, Vol. 8, Issue 11, 8 November 2021, ISSN 2352-3018. Online: https://www.thelancet.com/journals/lanhiv/article/PIIS2352-3018(21)00247-2/fulltext#%20 (quoted 11.11.2021).

[9] See for instance: RAYNAUD, M., AUBERT, O., DIVARD, G., REESE, P., KAMAR, N., YOO, D., et al.: Dynamic prediction of renal survival among deeply phenotyped kidney transplant recipients using artificial intelligence: an observational, international, multicohort study. In: *The Lancet Digital Health*, 27 October 2021. Online: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(21)00209-0/fulltext (quoted 11.11.2021); PREETHA, CH., J.,MEREDIG, H., BRUGNARA, G., MAHMUTOGLU, M., FOLTYN, M., ISENSEE, F.: Deep-learning-based synthesis of post-contrast T1-weighted MRI for tumour response assessment in neuro-oncology: a multicentre, retrospective cohort study. In: *The Lancet Digital Health*, 20 October 2021. Online: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(21)00205-3/fulltext (quoted 11.11.2021).

or for the people with disabilities, or use of applications in the form of chatbots.[10] Amid the COVID-19 pandemic we have witnessed several examples of new drugs developed with AI.[11]

With the use of AI in health and nursing care, several issues may arise. Although the main aim of the paper is to analyse and highlight the international legal regulation, it is worth mentioning some ethical issues associated with AI in health and nursing. Current policy and ethical guidelines for AI technology are lagging behind the progress AI has made in the health care field.[12] Till this date, there are several ethical codes and proposals applicable in this area. Their aim is to provide guidance to stakeholders about how basic moral requirements should direct or constrain their decisions and actions in the specific context of developing, deploying and assessing the performance of AI technologies for health.[13] The most frequently mentioned are *Asilomar Artificial Intelligence Principles*,[14] the *Montreal Declaration for a Responsible Artificial Intelligence*[15] (Montreal Declaration), the *Top Ten Principles for Ethical Artificial Intelligence*,[16] the World Commission's *Report on Ethics, Scientific Knowledge and Technology on Ethics robotics*, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent*

[10] See for instance: KARL, J.: "Are You There, Chatbot?": Automated Care Grows Up. In: *Health Tech Magazine*, 14 August 2020. Online: https://healthtechmagazine.net/article/2020/08/are-you-there-chatbot-automated-care-grows (quoted 11.11.2021).

[11] See for instance: PAUL, D., SANAP, G., SHENOY, S., KALYANE, D., KALIA, K., TEKADE, R. K.: Artificial intelligence in drug discovery and development, In: *Drug Discovery Today*, Volume 26, Issue 1, January 2021, pp. 80–93; ARSHADI, A. K., WEBB, J., SALEM, M., et al.: Artificial Intelligence for COVID-19 Drug Discovery and Vaccine Development, Frontiers in Artificial Intelligence, 18 August 2021. Online: https://www.frontiersin.org/articles/10.3389/frai.2020.00065/full (quoted 11.11.2021).

[12] RIGBY, M.: Ethical Dimensions of Using Artificial Intelligence in Health Care. In: *AMA Journal of Ethics*, Volume 21, Number 2, February 2019, p. 121. Online: https://journalofethics.ama-assn.org/sites/journalofethics.ama-assn.org/files/2019-01/fred1-1902_1.pdf (quoted 11.11.2021).

[13] WHO: Ethics and Governance of Artificial Intelligence in Health: WHO Guidance, Geneva, 2021, ISBN 978-92-4-002920-0, p. 23.

[14] Future of Life Institute: *Asilomar AI Principles*. Online: https://futureoflife.org/ai-principles/ (quoted 11.11.2021).

[15] University of Montreal´s Technosocial Innovation Centre: *Montreal Declaration for a Responsible AI*. Online: https://www.montrealdeclaration-responsibleai.com/ (quoted 11.11.2021).

[16] UNI Global Union: *Top 10 Principles for Ethical Artificial Intelligence*. Online: http://www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf (quoted 11.11.2021).

Systems,[17] the *Ethics Guidelines for Trustworthy Artificial Intelligence*[18] or the newly published World Health Organization´s *Ethics and Governance of Artificial Intelligence for Health*.[19] The following ethical principles are common to these documents: (a) human dignity, (b) respect for person's autonomy, (c) transparency and explainability, (d) beneficence and nonmaleficence, (e) justice and fairness, (f) responsibility and accountability.

Regarding the legal aspects of AI in health and nursing, there are several of them that need to be addressed in the future international legal frameworks. Among the most frequently mentioned are the protection of human rights, protection of personal data, safety and liability, and legal personality of AI. The above-mentioned aspects are currently addressed by a handful of States, international intergovernmental and non-state actors with different outcomes, especially when it comes to the adoption of legal framework. In the following chapters we look closely on these aspects through international initiatives of the United Nations (UN), the World Health Organization (WHO), Organisation for Economic Co-Operation and Development (OECD), and the regional organizations, such as the European Union (EU) and the African Union (AU).

15.2 The Role of the United Nations in Regulation of AI in Health and Nursing

The UN, as a universal international intergovernmental organization, provides an appropriate forum for establishing a common approach to the adoption of adequate ethical and legal standards for AI. The UN's three foundational pillars – peace and security, human rights and development – position it well to help spotlight issues emerging in the digital age and advocate on

[17] The Institute of Electrical and Electronics Engineers: *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*. Online: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf (quoted 11.11.2021).

[18] EU: Ethics Guidelines for Trustworthy AI, Independent High-Level Expert Group on Artificial Intelligence Set Up by the European Commission, 8 April 2019, Brussels. Online: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_EN.pdf (quoted 11.11.2021).

[19] WHO: Ethics and Governance of Artificial Intelligence in Health: WHO Guidance, Geneva, 2021, ISBN 978-92-4-002920-0.

behalf of humanity's best interests.[20] The main organ in the UN dealing with the AI, although only partially, is the High-Level Panel on Digital Cooperation, which was convened by the UN Secretary-General to provide recommendations on how the international community could work together to optimise the use of digital technologies and mitigate the risks. In 2019, the Panel published the report *"The Age of Digital Interdependence"*[21] and with it a series of recommendations to improve digital cooperation. In the report, the Panel included five sets of recommendations on how the international community could work together to optimize the use of digital technologies and mitigate the risks.[22] The relevant recommendation regarding the AI is Recommendation 3C, which states that "autonomous intelligent systems should be designed in ways that enable their decisions to be explained and humans to be accountable for their use. Audits and certification schemes should monitor compliance of AI systems with engineering and ethical standards, which should be developed using multi-stakeholder and multilateral approaches. Life and death decisions should not be delegated to machines. We call for enhanced digital cooperation with multiple stakeholders to think through the design and application of these standards and principles such as transparency and non-bias in autonomous intelligent systems in different social settings."[23] In May 2020, the UN Secretary-General introduced the *Roadmap for Digital Cooperation*. The action-oriented Roadmap presents the Secretary-General's recommendations for action by diverse stakeholders that would enhance global digital cooperation in several areas.[24] Regarding

[20] UN Secretary-General's High-level Panel on Digital Cooperation: The Age of Digital Interdependence, 2019, p. 27. Online: https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf (quoted 11.11.2021).

[21] *Ibidem*.

[22] (a) Build an inclusive digital economy and society; (b) Develop human and institutional capacity; (c) Protect human rights and human agency; (d) Promote digital trust, security and stability; (e) Foster global digital cooperation.

[23] *Ibidem*, p. 30.

[24] (a) Achieving universal connectivity by 2030—everyone should have safe and affordable access to the internet; (b) Promoting digital public goods to unlock a more equitable world—the internet's open source, public origins should be embraced and supported; (c) Ensuring digital inclusion for all, including the most vulnerable—under-served groups need equal access to digital tools to accelerate development; (d) Strengthening digital capacity building—skills development and training are needed around the world; (e) Ensuring the protection of human rights in the digital era—human rights apply both online and offline; (f) Supporting global cooperation on artificial intelligence that is trustworthy, human-rights based, safe and sustainable and promotes peace; (g) Promoting digital trust and security— calling for a global dialogue to advance the Sustainable Development Goals;

the issues raised around inclusion, coordination, and capacity-building for member states on AI, the UN Secretary-General intends to establish a multi-stakeholder advisory body on global artificial cooperation to provide guidance to the UN Secretary-General and the international community on AI that is trustworthy, human-rights based, safe and sustainable and promotes peace. Such advisory body will comprise UN member states, relevant UN entities, interested companies, academic institutions and civil society groups.[25] The UN Secretary-General in *Our Common Agenda* report expressed that one of the key proposals in achieving the Sustainable Development Goals[26] is to promote regulation of AI.[27] Although the UN is gradually involved in AI, there is no specific initiative that focuses on AI in health and nursing, rather the attention is focused on a general AI. We may conclude that the UN as such did not prepare a legally binding or non-binding instrument that would address the ethical and legal aspects of AI in health and nursing.

15.3 Initiatives of Specialized International Organizations

The development and use of AI in health and nursing is affected by initiatives of several specialized international organizations. For the purposes of this paper, the author focuses its attention to the initiatives and outcomes of the World Health Organization and the Organization for Economic Co-operation and Development. Each of the mentioned organizations adopted important international reports and principles, which are, at this point, legally non-binding instruments. However, we believe that their outcomes will inevitably shape the future international and national legal regulation in this area.

(h) Building a more effective architecture for digital cooperation—make digital governance a priority and focus the United Nation's approach.

[25] UN General Assembly: Road map for digital cooperation: implementation of the recommendations of the High-level Panel on Digital Cooperation: Report of the Secretary General, 29 May 2020, A/74/821, para. 88. Online: https://undocs.org/en/A/74/821 (quoted 11.11.2021).

[26] For more on Sustainable Development Goals see UN General Assembly resolution 70/1: Transforming our world: the 2030 Agenda for Sustainable Development, A/RES/70/1, 25 September 2015. Online: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E (quoted 11.11.2021).

[27] United Nations: Our Common Agenda: Report of the Secretary-General, United Nations, New York, 2021, p. 7, ISBN 978-92-1001012-2. Online: https://www.un.org/en/content/common-agenda-report/assets/pdf/Common_Agenda_Report_English.pdf(quoted11.11.2021).

15.3.1 World Health Organization

Since 1948, the WHO plays and important part in regulating international aspects of health care. Its initial priorities were malaria, tuberculosis, venereal disease and other communicable diseases, plus women and children's health, nutrition and sanitation. From the start, it worked with member states to identify and address public health issues, support health research and issue guidelines.[28] The WHO in the last years adopted several resolutions and documents that urge States to promote the use of digital technologies and AI in healthcare. With regard to health and nursing care, in 2018 the World Health Assembly adopted the *WHO Digital Health resolution (WHA71.7)* which urges States to promote the use of digital technologies, including improving access to quality data and monitoring, and to develop data protection legislation and policies on, for example, access to data sharing, informed consent, security, privacy, interoperability and inclusiveness in line with international human rights obligations.[29] In July 2018, the WHO and International Telecommunication Union (ITU) has set up an expert group on Ethics of AI for health[30] which is in the process of creating a framework on Regulations of AI for health that would be published in November 2021. The Group works on the premise that a standardized and transparent evaluation of AI methods would benefit from the widespread adoption of AI in the field of health. It should be noted that the group does not intend to specify AI for the health algorithms themselves as an ITU recommendation, nor to standardize medical data formats, nor to set performance criteria for the hardware on which the AI algorithms are based.[31]

The use AI is inevitably connected with the use of personal data, which the WHO addressed in the 2018 Astana Declaration. The Declaration calls for promotion of rational, safe use and protection of personal data and use of technology to improve access to health care, enrich health service delivery,

[28] WHO: Working for health: An introduction to the World Health Organization, Geneva: WHO Press, 2017. ISBN: 92-4-156313-5, p. 4. Online: https://www.who.int/about/brochure_en.pdf (quoted 11.11.2021).

[29] WHO: Digital Health, 26 May 2018, A71/VR/7, para. 7 and 10. Online: https://apps.who.int/gb/ebwha/pdf_files/WHA71/A71_R7-en.pdf (quoted 11.11.2021).

[30] ITU: United Nations Activities on Artificial Intelligence (AI), International Telecommunication Union: Geneva, 2019, p. 71, ISBN 978-92-61-29601-8. Online: https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf (quoted 11.11.2021).

[31] SALATHÉ, M., WIEGAND, T., WENZEL, M., KRISHNAMURTHY, R.: Focus Group on Artificial Intelligence for Health, 2018, p. 3. Online: https://www.itu.int/en/ITU-T/focusgroups/ai4h/Documents/FG-AI4H_Whitepaper.pdf (quoted 11.11.2021).

improve the quality of service and patient safety and increase the efficiency and coordination of care.[32] Although the Declaration specifically focuses on primary healthcare, the above mentioned is relevant also outside the primary healthcare. As stated by Rasanathan and Evans, social and technological innovations can be better harnessed and applied widely, including in building capacity in primary care and community systems to tackle health security challenges revealed by COVID-19.[33] The WHO adopted in 2019 the *Global strategy on digital health: 2020–2025*, where states are encouraged to strengthen governance for digital health at global, regional and national levels. Actions to strengthen governance should include defining principles and reaching across-sectoral and international agreements for data sharing, quality and accuracy of health data and prioritization of investment plans and policy.[34] This is especially relevant when using AI in health and nursing. In its strategic objective no. 3 – Strengthen governance of digital health at global, regional and national levels, the member states, Secretariat of the WHO and partners for implementing the Global strategy on digital health are, in a short term (1–2 years), asked to support the Secretariat in establishing international health data regulation, a framework for regulating, benchmarking or certifying artificial intelligence and digital health medical devices.[35]

One of the most important documents adopted by the WHO regarding the use of AI in health is the 2021 Report *Ethics and Governance of AI in Health* in which the WHO´s Expert Group confirms six guiding principles in AI in health:

1) to protect human autonomy;
2) AI designers should safeguard privacy and confidentiality by providing patients with valid informed consent through appropriate legal frameworks;
3) calls for AI designers to ensure regulatory requirements for safety, accuracy and efficacy, including measures of quality control;

[32] WHO: Declaration of Astana, adopted at the Global Conference on Primary Health Care, 25 and 26 October 2018, p. 9. Online: https://www.who.int/docs/default-source/primary-health/declaration/gcphc-declaration.pdf (quoted 11.11.2021).

[33] RASANATHAN, K., EVANS, T.: Primary health care, the Declaration of Astana and COVID-19. In: *Bulletin of the World Health Organization*, 28 September 2020, p. 10. Online: https://www.who.int/bulletin/online_first/BLT.20.252932.pdf (quoted 11.11.2021).

[34] WHO: Global strategy on digital health 2020-2025, Geneva, 2019, ISBN 978-92-4-002092-4, p. 24. Online: https://www.who.int/docs/default-source/documents/gs4dh-daa2a9f352b0445bafbc79ca799dce4d.pdf (quoted 11.11.2021).

[35] *Ibidem*, p. 54.

4) requires information to be published or documented before the AI technology is designed or deployed;
5) to ensure inclusiveness and equity so that AI for health is accessible to the widest possible number of people, irrespective of age, gender, ethnicity or other characteristics protected under human rights codes;
6) urges designers, developers and users to transparently assess applications during actual use to determine whether AI responds adequately and appropriately to expectations and requirements.[36]

The report recommends that the WHO should work in a coordinated manner with appropriate intergovernmental organizations to identify and formulate laws, policies and best practices for ethical development, deployment and use of AI technologies for health. Furthermore, the WHO should consider issuing model legislation to be used as a reference for governments that wish to build an appropriate legal framework for the use of AI for health. With the increase in AI standards and laws around the world and diffusion of how and where AI ethics is managed, additional international oversight and enforcement may be necessary to ensure convergence on a core set of principles and requirements that meet ethical principles and human rights obligations.[37]

Although the WHO did not adopt a draft legislation on the use of AI in health and nursing, the first step has already been taken with the adoption of the "Ethics and Governance of AI in Health". It can be assumed that in the following years the WHO in cooperation with States and non-state actors will work on a model legislation in the analysed area. Furthermore, the already existing legislation, especially in the EU, the USA and other technologically developed states will influence the future international legal regulation.

15.3.2 Organization for Economic Co-Operation and Development

The OECD is an international inter-governmental organization whose one of the aims is to promote policies designed to achieve the highest sustainable economic growth and employment and a rising standard of living in Member States, while maintaining financial stability, and thus to contribute to the

[36] WHO: Ethics and Governance of Artificial Intelligence in Health: WHO Guidance, Geneva, 2021, ISBN 978-92-4-002920-0.
[37] *Ibidem*, p. 111.

development of the world economy.[38] Naturally, the OECD and its Member States focus on the new possibilities that the AI provides. The OECD supports governments by measuring and analysing the economic and social impacts of AI technologies and applications, and engaging with all stakeholders to identify good practices for public policy. The most important instrument in this regard is the OECD *Principles on Artificial Intelligence* adopted in May 2019 through the OECD Council Recommendation on Artificial Intelligence. The recommendation sets out five principles based on the values of responsible stewardship of trustworthy AI, namely:

a) AI should benefit people and the planet through inclusive growth, sustainable development and prosperity;
b) AI systems should be designed to respect the rule of law, human rights, democratic values and diversity, and should include appropriate safeguards – for example, allowing for human intervention if necessary – in order to ensure a just society;
c) there should be transparency and responsible publicity about AI systems to ensure that people understand and can challenge AI-based results;
d) AI systems must operate reliably and safely throughout their life cycles and potential risks should be continuously assessed and managed;
e) organizations and individuals developing, implementing or operating AI systems should be responsible for their proper functioning in accordance with the above principles.[39]

In the context of health and health care, as well as nursing care, it is especially important for policy to foster a digital ecosystem for AI, operationalise the OECD AI principles, establish appropriate regulation and guidance, build human capacity and invest strategically and sustainably.[40] Ensuring data quality, availability and security can be achieved by implementing the

[38] Paris Convention on the Organisation for Economic Co-operation and Development, 14 December 1960, art. 1.

[39] OECD: Principles on Artificial Intelligence, 22 May 2019, OECD/LEGAL/0449, p. 7–8. Online: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (quoted 11.11.2021).

[40] HASHIGUCHI, T., SLAWOMIRSKI, L., ODERKIRK, J.: OECD Health Working Papers No. 128: Laying the foundations for artificial intelligence in health, OECD, 3 June 2021, DELSA/HEA/WD/HWP(2021)5, para. 42. Online: https://www.oecd-ilibrary.org/docserver/3f62817d-en.pdf?expires=1638634715&id=id&accname=guest&checksum=7E4C4EF518CB36DDC790CA7CE9E22987 (quoted 11.11.2021).

OECD Health Data Governance Recommendation. This Recommendation sets out principles for national health data governance frameworks that improve data quality and accessibility while protecting privacy and data security. The Recommendation calls on governments to address unnecessary barriers to the efficient exchange and interoperability of health data, particularly those that are blocking public-private and cross-border monitoring and research. The Recommendation calls for international cooperation to develop global standards for data exchange and data terminology; and harmonising health data governance frameworks that protect data privacy and security.[41] In line with the OECD Health Data Governance and AI Principles, frameworks for health data governance should emphasise transparency, public communication and stakeholder engagement, explicitly highlighting the importance of trust.[42] Subsequently, the OECD published in June 2021 report *State of Implementation of the OECD AI Principles Insights from National AI Policies*[43] which gives practical advice for implementing the OECD AI Principles throughout each phase of the AI policy cycle, namely policy design, implantation and intelligence.

The OECD also established the AI Policy Observatory which brings together committees from across the OECD as well as a range of other stakeholders. The goal is to identify promising AI applications, map their economic and social impact and share the information as widely as possible.[44] Furthermore, the European Community-OECD database of national AI policies contains national AI strategies and AI-related policy initiatives from over 60 countries. Strategy and policy priorities include financing AI research and development institutions and projects, addressing societal challenges, promoting AI uptake by business, fostering inclusive social dialogue, equipping the population with the skills for developing and using AI and fostering a fair labour market transition for workers. Countries are using public and inclusive dialogue for trustworthy AI. Many countries are establishing

[41] *Ibidem*, para. 68.

[42] OECD: Recommendation of the Council on Health Data Governance, 13 December 2016, OECD/LEGAL/0433, p. 4. Online: http://legalinstruments.oecd.org (quoted 11.11.2021).

[43] For more see OECD: State of Implementation of the OECD AI Principles Insights from National AI Policies, June 2021, DSTI/CDEP(2020)15/FINAL. Online: https://www.oecd-ilibrary.org/docserver/1cd40c44-en.pdf?expires=1638292371&id=id&accname=guest&checksum=14A6B682AC03423B4C28FCE66A9754E9 (quoted 11.11.2021).

[44] OECD: OECD creates expert group to foster trust in artificial intelligence, 13 September 2018. Online: https://www.oecd.org/innovation/oecd-creates-expert-group-to-foster-trust-in-artificial-intelligence.htm (quoted 11.11.2021).

national AI offices that are tasked with overseeing national AI policy implementation and ensuring policy coherence.[45] Open access to public sector data continues to be a priority as national data strategies increasingly focus on AI to foster a robust digital ecosystem for AI and advance AI research and development. Policies to promote access to public data and initiatives that enable private sector data sharing include data trusts, data dams and data spaces. As part of their AI strategy, several countries have developed or are developing centralised, accessible repositories of open public datasets such as anonymised government health records and satellite data.[46] In February 2020, the OECD launched OECD.AI, a platform for policy makers to monitor developments in the AI policy landscape and the OECD.AI Network of Experts (ONE AI), a multi-stakeholder expert group that is developing practical guidance to help implement the OECD AI Principles.[47]

States are exploring approaches to ensure trustworthy AI and mitigate risks associated with the development and deployment of AI systems. In addition to exploring the application and need to adapt current legislation for AI, emerging regulatory actions for AI trustworthiness include:

1) providing soft law guidance;
2) considering hard law approaches,
3) introducing application-specific moratoriums or bans;
4) promoting controlled environments for regulatory experimentation and
5) supporting international standardisation efforts and international law efforts.

Overall, countries' initiatives still retain predominately "soft" regulatory approaches for AI, including the development of ethical frameworks and guidelines, voluntary processes, technical standards, and codes of conduct. There is, however, a trend towards the development of legislative reforms and regulations for specific applications.[48] In general, the OECD Principals on Artificial Intelligence and its initiatives in AI are mostly oriented to guide States in drafting national legal frameworks on AI which may be also used in health and nursing.

[45] *Ibidem*, p. 10.
[46] *Ibidem*, p. 11.
[47] *Ibidem*, p. 15.
[48] *Ibidem*, p. 28.

15.4 Specific Role of Regional Organizations in Regulating AI in Health and Nursing

Although several regional organisations announced their interest in the application of the AI, the proper legal framework by which the AI will be regulated is in most parts absenting. In the following subchapters, the current state of legal regulation of AI in health and nursing is analysed. Particular attention is focused on already adopted, as well as proposed legislature of the EU and AU.

15.4.1 The European Union on the path to the first legally binding document

The European Union is the most active regional organization in the area of AI. Between 2018 and 2021, organs of the EU have adopted numerous instruments, legally binding as well as non-binding, whose aim is to ensure an appropriate ethical and legal framework.[49] A particular progress was made, when the High-Level Expert Group on AI adopted the *Ethics Guidelines for Trustworthy AI*, which not only outlines the main ethical requirements re-

[49] EU: White Paper On Artificial Intelligence – A European approach to excellence and trust, European Commission, 19 February 2020, COM(2020) 65. Online: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf; EU: Resolution of the European Parliament of 12 February 2020 on automated decision-making processes: ensuring consumer protection and free movement of goods and services, 12 February 2020, 2019/2915(RSP); EU: Opinion of the European Economic and Social Committee on 'Artificial intelligence — The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society' (own-initiative opinion), OJ C 288, 31 August 2017; EU: European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics, 18 July 2018, 2015/2103(INL), OJ C 252; EU: European Parliament resolution of 12 February 2019 on a comprehensive European industrial policy on artificial intelligence and robotics, 12 February 2019, 2018/2088(INI); EU: Declaration on Artificial Intelligence Cooperation, 10 April 2018. Online: https://ec.europa.eu/jrc/communities/en/community/digitranscope/document/eu-declaration-cooperation-artificial-intelligence; European Commission. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A European strategy for data, https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf; European Commission. Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee. Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, https://ec.europa.eu/info/sites/info/files/report-safety-liability-artificial-intelligence-feb2020_en_1.pdf (quoted 11.11.2021).

garding the development and use of AI, but also provides definition of the term "artificial intelligence". In accordance with the document, the AI is trustworthy when following requirements are fulfilled:

1) human agency and oversight,
2) technical robustness and safety,
3) privacy and data governance,
4) transparency,
5) diversity, non-discrimination and fairness,
6) environmental and societal well-being, and
7) accountability.[50]

When it comes to the four legal aspects that are mentioned in Chapter 1, the EU in a way addresses all of the mentioned aspects. In 2018, the European Parliament in its resolution on civil law rules on robotics considered the possibility of granting a specific legal status to the AI, in form of the so-called e-person, who would be responsible for compensating for any damage they may cause and to make independent decisions with third parties.[51] In this context, the European Parliament has made civil law recommendations in the field of robotics, which include a proposal to examine the possibility of introducing so-called e-subjectivity for robots so that they can be held liable under civil law for the damage they cause. However, the European Economic and Social Committee does not share this view and opposes any form of legal status for robots or AI systems, as this creates an unacceptable moral hazard. The acknowledgment of legal personality of AI seems to problematic for several reasons. Firstly, the AI does not have characteristics that are usually associated with human persons, such as freedom of will, intentionality, self-awareness or a sense of personal identity.[52] Secondly, even if the legal personality of the AI were established in a similar way as the legal personality of the legal person, it would also be an inappropriate solution, as the legal

[50] EU: Ethics Guidelines for Trustworthy AI, Independent High-Level Expert Group on Artificial Intelligence Set Up by the European Commission, 8 April 2019, Brussels, p. 2. Online: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_EN.pdf (quoted 11.11.2021).

[51] EU: European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics, 18 July 2018, 2015/2103(INL), OJ C 252, para. 59 f).

[52] UNESCO: Report of the World Commission on the Ethics of Scientific Knowledge and Technology on Robotics Ethics, 14 September 2017, Paris, SHS/YES/COMEST-10/17/2 REV., para. 201.

person is also responsible for the action of the natural person, which is not the case for AI.[53]

The protection of fundamental human rights and freedoms is a crucial legal aspect in the development and use of AI in health and nursing. The use of AI with its specific characteristics (e.g., opacity, complexity, dependency on data, autonomous behaviour) can adversely affect a number of fundamental rights enshrined in the *European Charter of Fundamental Rights*.[54] Fundamental rights not only protect individuals from state intervention, but also oblige the state to protect certain freedoms from interference by third parties. The state can fulfil these so-called "obligations to protect" by, for example, enacting appropriate legislation that applies to relations between private individuals or by creating specific approval procedures for placing goods or services on the market that could endanger the fundamental rights of its users.[55] Particular attention in regards to the use of AI in health and nursing should be focused on the respect and protection of human dignity, which is the "main pillar" of fundamental human rights and freedoms. Apart from human dignity, respect for private life and protection of personal data (Articles 7 and 8), non-discrimination (Article 21), equality between women and men (Article 23), right to freedom of expression (Article 11) and freedom of assembly (Article 12), right to an effective remedy and to a fair trial, the rights of defence and the presumption of innocence (Articles 47 and 48),[56] as well as the general principle of good administration. The right to a high level of environmental protection and the improvement of the quality of the environment (Article 37) is also relevant in relation to the health and safety of people. In case infringements of fundamental rights still happen, effective

[53] O´SULLIVAN, S. (et al.): Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. In: *The International Journal of Medical Robotics and Computer Assisted Surgery*, 2019, p. 7, ISSN: 1478-596X.

[54] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, Brussels, 21 April 2021, 2021/0106(COD), Section 3.5. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 11.11.2021).

[55] SCHNEEBERGER, D., STÖGER, K., HOLZINGER, A.: The European Legal Framework for Medical AI. In: HOLZINGER, A., KIESEBERG, P., MIN TJOA, A., WEIPPL, E. (eds.): *Machine Learning and Knowledge Extraction*, Cham: Springer, 2020, p. 210, ISBN 978-3-030-57320-1. Online: https://www.aholzinger.at/wordpress/wp-content/uploads/2020/08/The-European-Legal-Framework-For-Medical-AI.pdf (quoted 11.11.2021).

[56] Charter of Fundamental Rights of the European Union, OJ C 326, 26.10.2012.

redress for affected persons will be made possible by ensuring transparency and traceability of the AI systems coupled with strong *ex post* controls. The current proposal of Artificial Intelligence Act imposes some restrictions on the freedom to conduct business (Article 16) and the freedom of art and science (Article 13) to ensure compliance with overriding reasons of public interest such as health, safety, consumer protection and the protection of other fundamental rights when high-risk AI technology is developed and used. Those restrictions are proportionate and limited to the minimum necessary to prevent and mitigate serious safety risks and likely infringements of fundamental rights.[57]

Regarding the safety and effectiveness of medical AI, the relevant legal regulation consists of the *Medical Device Regulation* (2017/745) (MDR), which came into force on 26 May 2021[58] and the *Regulation on in vitro diagnostic medical devices* (2017/746) (RIVDMD), which will come into force in 26 May 2022.[59] The mentioned regulations are reflecting the future use of AI in healthcare. According to the Art. 2 (1) of the MDR, a "medical device" means any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes:

- diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease,
- diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability,
- investigation, replacement or modification of the anatomy or of a physiological or pathological process or state,
- providing information by means of in vitro examination of specimens derived from the human body, including organ, blood and tissue donations,

[57] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, Brussels, 21 April 2021, 2021/0106(COD), Section 3.5. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 11.11.2021).

[58] The Medical Device Regulation repealed the *Medical Device Directive* (93/42/EEC) and the Directive on active implantable medical devices (90/385/EEC).

[59] The Regulation on in vitro diagnostic medical devices will repeal the Directive on in vitro diagnostic medical devices (98/79/EC).

and which does not achieve its principal intended action by pharmacological, immunological or metabolic means, in or on the human body, but which may be assisted in its function by such means.[60] On the other hand, the MDR in Recital 19 clarifies that "software for general purposes, even when used in a healthcare setting, or software intended for lifestyle and well-being purposes is not a medical device."[61]

Medical devices are under MDR classified into four categories,[62] based on the intended purpose of the medical devices and their inherent risks.[63] The MDR introduces new implementing and classification rules for software in medical devices. In accordance with the Rule 11 in Chapter III of Annex VIII of the MDR, "software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as class IIa, except if such decisions have an impact that may cause:

- death or an irreversible deterioration of a person's state of health, in which case it is in class III or
- a serious deterioration of a person's state of health or a surgical intervention, in which case it is classified as class IIb.

Software intended to monitor physiological processes is classified as class IIa, except if it is intended for monitoring vital physiological parameters, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is classified as class IIb. All other software is classified as class I."[64] A CE marking will especially indicate the conformity with the applicable requirements set out in the MDR so that a medical device can move freely within the EU and be put into service in accordance with its intended purpose.[65] Manufacturers of medical devices shall undertake an assessment of the conformity of their devices prior to placing them on the market.[66] The applicable conformity assessment procedure is based on the classification and type of the particular medical de-

[60] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, OJ L 117, 5.5.2017, art. 2 (1).

[61] *Ibidem*, Recital 19.

[62] Classes I (low risk), IIa (medium risk), IIb (higher risk), and III (highest risk).

[63] *Ibidem*, Art. 51 (1).

[64] *Ibidem*, Rule 11 in Chapter III of Annex VIII.

[65] *Ibidem*, Recital 40 and Art. 2 (43).

[66] *Ibidem*, Art. 52 and Annexes IX–XI.

vice.[67] For instance, class I devices have a low level of vulnerability and thus the conformity assessment procedure can generally be carried out under the sole responsibility of the manufacturers.[68] In contrast, medical devices classified into class IIa, IIb, and III which have a higher risk than class I devices entail the involvement of a notified body, a conformity assessment body designated in accordance with the MDR.[69]

When it comes to the legal framework for liability of AI, there is currently no fully harmonized EU regulatory framework. Unlike traditional product liability regimes, where a product can be characterized as "defective" due to the manufacturer's negligence, which in turn can be considered to cause harm – in the case of AI, there is no equivalent error. This is because AI has not been explicitly programmed to work in a specific way. In many cases, AI developers will not be able to provide a traditional causal explanation of AI behaviour based on their programming inputs. The complexity of large information inputs combined with ever-changing learned behaviour disrupts the traditional occasional connections between programmers' input and system behaviour.[70] It seems that there is usually a "shared" or "distributed" responsibility among robot designers, engineers, programmers, manufacturers, investors, vendors and users. At the same time, this solution weakens the aspect of responsibility. The main challenge is to avoid the possible paralyzing effect of taking and attributing responsibility. One solution to take responsibility may be to develop techniques to anticipate the impacts of robotic development as much as possible.[71] Another solution is to carefully address the necessary occurrence of unexpected consequences by considering the social introduction of robotic technologies as a "social experiment" that needs to be carried out with great care.[72]

It must be noted that EU has taken some steps to address this issue. In 2017, the European Parliament adopted the resolution on *Civil Law Rules*

[67] *Ibidem*, Art. 52.
[68] *Ibidem*, Recital 60 and Art. 52 (7).
[69] *Ibidem*, Recital 60 and Art. 2 (42).
[70] KERR, I., MILLAR, J.: Delegation, Relinquishment and Responsibility: The Prospect of Robot Experts. In: CALO, R., FROOMKIN, A. M., KERR, I.: *Robot Law*, Cheltenham, Edward Elgar, 2016, p. 106–108, ISBN: 978-1-78347-672-5.
[71] WAELBERS, K., SWIERSTRA, T.: The Family of the Future: How Technologies Can Lead to Moral Change. In: VAN DEN HOVEN, J., DOORN, J., SWIERSTRA, T. (eds.): *Responsible Innovation*, Springer, Dordrecht, 2014, p. 219–236, ISBN: 978-94-017-8956-1.
[72] VAN DE POEL: Why New Technologies Should Be Conceived as Social Experiments. In: *Ethics, Policy & Environment*, Vol. 16, No. 3, 2013, p. 352–355, ISSN: 2155-0085.

on Robotics.[73] The questions whether the current liability rules are sufficient and whether new rules are required "to provide clarity on the legal liability of various actors concerning responsibility for the acts and omissions of robots."[74] It also points out that the current scope of Council Directive concerning liability for defective products (85/374/EEC – *Product Liability Directive*) may not adequately cover the new developments in robotics.[75] The resolution emphasizes that the civil liability for damage caused by robots is a crucial issue which also needs to be analysed and addressed at Union level in order to ensure the same degree of efficiency, transparency and consistency in the implementation of legal certainty throughout the European Union for the benefit of citizens, consumers and businesses alike.[76] It thus asks the European Commission for a proposal for a legislative instrument on legal questions related to the development and use of robotics and AI foreseeable in the next 10–15 years, combined with non-legislative instruments such as guidelines and codes of conduct.[77] The resolution recommends that the European Commission should define in this legislative instrument which of the two approaches should be applied: either strict liability (i.e., which "requires only proof that damage has occurred and the establishment of a causal link between the harmful functioning of the robot and the damage suffered by the injured party") or the risk management approach (i.e., which "does not focus on the person 'who acted negligently' as individually liable but on the person who is able, under certain circumstances, to minimize risks and deal with negative impacts").[78] It also recommends an obligatory insurance scheme and an additional compensation fund to ensure that damages will be paid out in situations where no insurance cover exists.[79]

[73] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics, 18 July 2018, 2015/2103(INL), OJ C 252.

[74] *Ibidem*, Section AB.

[75] *Ibidem*, Section AH.

[76] *Ibidem*, Section 49.

[77] *Ibidem*, Section 51.

[78] *Ibidem*, Section 53–55.

[79] GERKE, S., MINSSEN, T., COHEN, G.: Ethical and legal challenges of artificial intelligence-driven healthcare. In: BOHR, A., MEMARZADEH, K.: Artificial Intelligence in Healthcare, London: Academic Press, 2020, p. 314–316, ISBN 978-0-12-818438-7. Online: https://reader.elsevier.com/reader/sd/pii/B9780128184387000125?token=35D2D384BA50A7028F25E9A1E20EEA422D0C618C7D3FA7856615047D6BE1760F5929EDE96D418C4129CBB130F287CAD2&originRegion=eu-west-1&originCreation=20211205145558 (quoted 11.11.2021).

In November 2019, the Expert Group on Liability and New Technologies—New Technologies Formation published report – *Liability for Artificial Intelligence and Other Emerging Digital Technologies*.[80] The Expert Group concludes that liability regimes are mainly regulated by the EU Member States except for strict liability of producers for defective products that is regulated by the *Product Liability Directive* at the EU level and that the Member States' liability regimes are a good starting point for new technologies and provide at least basic protection of victims.[81] On the other hand, the report also contains several points that need to be changed at national and EU levels. For instance, a person operating a permissible technology that nevertheless carries an increased risk of harm to others, for example AI-driven robots in public spaces, should be subject to strict liability for damage resulting from its operation. It also states, for instance, that a person using a technology which has a certain degree of autonomy should not be less accountable for ensuing harm than if said harm had been caused by a human auxiliary. Furthermore, manufacturers of products or digital content incorporating emerging digital technology should be liable for damage caused by defects in their products, even if the defect was caused by changes made to the product under the producer's control after it had been placed on the market. For situations exposing third parties to an increased risk of harm, compulsory liability insurance could give victims better access to compensation and protect potential tortfeasors against the risk of liability.[82] In February 2020, the European Commission published a *Report on the safety and liability implications of AI, the Internet of Things, and robotics*, in which the European Commission argues that in principle the existing Union and national liability laws are able to cope with emerging technologies.[83] However, it also identifies some challenges raised by new digital technologies such as AI that need to be addressed by adjustments in the current national and EU regulatory

[80] Expert Group on Liability and New Technologies New Technologies Formation: Liability for artificial intelligence and other emerging digital technologies, European Commission, 2019, ISBN 978-92-76-12959-2. Online: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf (quoted 11.11.2021).

[81] *Ibidem,* p. 3.

[82] *Ibidem,* p. 3–4.

[83] European Commission: Report on the safety and liability implications of AI, the Internet of things, and robotics, 19 February 2020, p. 17. Online: https://ec.europa.eu/info/sites/default/files/report-safety-liability-artificial-intelligence-feb2020_en.pdf (quoted 11.11.2021).

frameworks such as the Product Liability Directive. Although the Product Liability Directive's definition of product is broad, its scope could be further clarified to better reflect the complexity of emerging technologies and ensure that compensation is always available for damage caused by products that are defective because of software or other digital features. This would better enable economic actors, such as software developers, to assess whether they could be considered producers according to the Product Liability Directive. In close coordination with corresponding changes in the Union safety framework, the notion of "putting into circulation" that is currently used by the Product Liability Directive could be revisited to take into account that products may change and be altered. This could also help to clarify who is liable for any changes that are made to the product. [84]

Lastly, data protection and privacy are of great importance in health and nursing. Personal data used in development or use of AI in health and nursing is anonymised. However, the study of Na et al. points out, that existing algorithms could be used to re-identify 85.6% of adults and 69.8% of children in a physical activity study.[85] This finding certainly will not persuade the general public to trust manufacturers, software engineers that their personal data are safe and will not re-identified. Regarding the protection of personal data, the EU in 2016 adopted the *General Data Protection Regulation* (GDPR—2016/679), whose aim is to protect the right of natural persons to the protection of personal data.[86] It applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the EU, notwithstanding of whether the processing takes place in an EU or third country.[87] The legal framework for data protection in Europe provides detailed requirements and restrictions on the processing of personal data and contains new provisions on automated decision-making and profiling, which pose interesting challenges for robot developers. GDPR specifically addresses "data concerning health" which are defined as person-

[84] *Ibidem,* p. 14–15.

[85] NA, L. et al.: Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning. In: *JAMA Network Open*, 21 December 2018, p. 7–8. Online: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2719130 (quoted 11.11.2021).

[86] 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), OJ L 119, 4.5.2016, Art. 1 (2).

[87] *Ibidem,* Art. 2, 3 (1).

al data related to the physical or mental health of a natural person, including the provision of healthcare services, which reveal information about his or her health status.[88] According to Art. 9 (1) of the GDPR, the processing of special categories of personal data such as genetic data, biometric data, and data concerning health is prohibited.[89] Art. 9 (2) of the GDPR, on the other hand, contains a list of exceptions. For instance, the prohibition in Article 9 (1) of the GDPR shall usually not apply in cases where the data subject has given explicit consent for one or more specified purposes or where the processing is necessary for reasons of public interest in the area of public health or for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.[90] Another relevant provision is Art. 21 in which is stated that the data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.[91] In addition, data subjects have the right of access to the personal data concerning them that are being processed and the information about the existence of automated decision-making, including profiling, (...) and (...) meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.[92] It is also likely that companies that are controllers under the GDPR must carry out a data protection impact assessment for new AI-based technologies that shall be deployed in the clinical space.[93] GDPR explicitly states that a data protection impact assessment shall especially be required in cases of a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person or pro-

[88] *Ibidem,* Art. 4 (15).
[89] *Ibidem,* Art. 9 (1).
[90] *Ibidem,* Art. 9 (2)(a), (i), and (j).
[91] *Ibidem,* Art. 22 (1).
[92] GERKE, S., MINSSEN, T., COHEN, G.: Ethical and legal challenges of artificial intelligence-driven healthcare. In: BOHR, A., MEMARZADEH, K.: Artificial Intelligence in Healthcare, London: Academic Press, 2020, p. 321, ISBN: 978-0-12-818438-7. Online: https://reader.elsevier.com/reader/sd/pii/B9780128184387000125?token=35D2D384BA50A7028F25E9A1E20EEA422D0C618C7D3FA7856615047D6BE1760F5929EDE96D418C4129CBB130F287CAD2&originRegion=eu-west-1&originCreation=20211205145558 (quoted 11.11.2021).
[93] *Ibidem,* p. 323.

cessing on a large scale of special categories of data.[94] The assessment shall contain at least:

a) a systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller;
b) an assessment of the necessity and proportionality of the processing operations in relation to the purposes;
c) an assessment of the risks to the rights and freedoms of data subjects referred to in paragraph 1; and
d) the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with this Regulation taking into account the rights and legitimate interests of data subjects and other persons concerned.[95]

Currently, the nearest to AI legal regulation is the European Union which announced on 21 April 2021 the draft of Artificial Intelligence Act (AI Act). The proposed Act is especially relevant to the future development, introduction and use of the AI in health and nursing. The AI Act lays down:

a) harmonised rules for placing on the market, putting into service and the use of artificial intelligence systems in the Union;
b) prohibitions of certain artificial intelligence practices;
c) specific requirements for high-risk AI systems and obligations for operators of such systems;
d) harmonised transparency rules for AI systems intended to interact with natural persons, emotion recognition systems and biometric categorisation systems, and AI systems used to generate or manipulate image, audio or video content;
e) rules on market monitoring and surveillance.[96]

[94] 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), OJ L 119, 4.5.2016, Art. 35 (3) (a).

[95] *Ibidem,* Art. 35 (7).

[96] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, Brussels, 21 April 2021, 2021/0106(COD), Art. 1. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 11.11.2021).

The Commission proposes to ban completely AI systems that manipulate persons through subliminal techniques or exploit the fragility of vulnerable individuals, and could potentially harm the manipulated individual or third person; serve for general purposes of social scoring, if carried out by public authorities; or are used for running real time remote biometric identification systems in publicly accessible spaces for law enforcement purposes.[97] The above mentioned draft aims to regulate high-risk AI which will include most of the health and nursing AI applications. High-risk AI systems are defined as those that are part of a product falling under the EU product safety regulation (such as toys or medical devices) or belong to a list of stand-alone high-risk AI systems laid down by the proposal, such as AI systems assessing the creditworthiness of individuals or used in the context of recruitment.[98] When it comes to the requirements for High-risk AI systems, Art. 9 states that a risk management system shall be established, implemented, documented and maintained in relation to high-risk AI systems and it shall consist of a continuous iterative process run throughout the entire lifecycle of a high-risk AI system, requiring regular systematic updating.[99] The proposal also focuses on data and data governance, technical documentation of the High-risk AI systems, record-keeping, transparency and provision of information to users, accuracy, robustness ad cybersecurity. High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use.[100] Among the proposed obligations of providers of High-risk AI systems is to have a quality management system in place, to draw-up the technical documentation of the high-risk AI system and, among others, to ensure that the high-risk AI system undergoes the relevant conformity assessment procedure, prior to its placing on the market or putting into service.[101]According to Art. 21, providers of high-risk AI systems which consider or have reason to consider that a high-risk AI system which they have placed on the market or put into service is not in conformity with this Regulation shall immediately take the necessary corrective actions to bring that system into conformity, to withdraw it or to recall it, as appropriate. They shall inform the distribu-

[97] *Ibidem,* Art. 5.
[98] *Ibidem,* Art. 6 (1).
[99] *Ibidem,* Art. 9 (1), (2).
[100] *Ibidem,* Art. 14 (1).
[101] *Ibidem,* Art. 16.

tors of the high-risk AI system in question and, where applicable, the authorised representative and importers accordingly.[102] Special provisions are proposed for product manufacturers, importers, distributors, and users of High-risk AI systems. Each Member State shall designate or establish a notifying authority responsible for setting up and carrying out the necessary procedures for the assessment, designation and notification of conformity assessment bodies and for their monitoring.[103]

15.4.2 The African Union

The African Union represents another regional inter-governmental organization interested in the development, use and future regulation of AI. Africa faces several known challenges in developing AI such as a dearth of investment, a paucity of specialised talent, and a lack of access to the latest global research. These hurdles are being whittled down, albeit slowly, thanks to African ingenuity and to investments by multinational companies such as IBM Research, Google, Microsoft, and Amazon, which have all opened AI labs in Africa. Innovative forms of trans-continental collaboration such as Deep Learning Indaba, which is fostering a community of AI researchers in Africa, and Zindi, a platform that challenges African data scientists to solve the continent's toughest challenges, are gaining ground.[104] Currently, there are several national initiatives geared toward enabling autonomous data collection and data availability systems are key to AI's progress in Africa.[105] Also, there is a need to have mechanisms in place that ensure data is correct and accurate. In addition, data protection and regulations are needed in order to avoid misuse.[106]

[102] *Ibidem,* Art. 21.

[103] *Ibidem,* Art. 30 (1).

[104] CANDELON, F., EL BEDRAOUI, H., MAHER, H.: Developing an Artificial Intelligence for Africa strategy, OECD Development Matters Blog, 9 February 2021. Online: https://oecd-development-matters.org/2021/02/09/developing-an-artificial-intelligence-for-africa-strategy/#more-14109 (quoted 11.11.2021).

[105] For more information see GWAGWA, A., KRAEMER-MBULA, E., RIZK, N.RUTENBERG, I., DE BEER, J.: Artificial Intelligence (AI) Deployments in Africa: Benefits, Challenges and Policy Dimensions. In: *The African Journal of Information and Communication*, Vol. 26, 2020, p. 12–13, ISSN 2077-7213. Online: http://www.scielo.org.za/pdf/ajic/v26/02.pdf (quoted 11.11.2021).

[106] NAYEBARE, M.: Artificial intelligence policies in Africa over the next five years. In: *The ACM Magazine for Students*, Vol. 26, No. 2, 2019, p. 54, ISSN 0001-0782. Online: https://dl.acm.org/doi/10.1145/3368075 (quoted 11.11.2021).

The first step towards ethical and legal regulation of AI within the AU has been taken on 26 October 2019, when ministers responsible for communication, and information and communication technologies have adopted the Sharm El Sheikh Declaration, that puts special focus on the African Digital Transformation Strategy and the African Union Communication and Advocacy Strategy, as well as the Union's Brand and Communication Style Guideline. In accordance with the paragraph 15 of the Declaration, the Working Group on Artificial Intelligence has been established in late December 2019. The objectives include the creation of a common African stance on AI, the development of an Africa wide capacity building framework and the establishment of an AI think tank to assess and recommend projects to collaborate on in line of Agenda 2063 and UN Sustainable Development Goals.[107] Currently, the Working Group is working on an African strategy for AI and reaching a unified African position on AI. Beside the Working Group, the African Union's High-level Panel on Emerging Technologies[108] is preparing broad guidance on the use of AI to promote economic development and its use in various sectors, including health care.

As was mentioned before, one of the key legal aspects of the use of AI is protection of personal data. In 2014, the AU Convention on cyber security and personal data protection, also known as the Malabo Convention, was adopted. The objective of the Convention is setting the essential rules for establishing a credible digital environment and address the gaps affecting the regulation and legal recognition of electronic communications and electronic signature; as well as the absence of specific legal rules that protect consumers, intellectual property rights, personal data and information systems and privacy online. It also aims to set up a minimum standards and procedures to reach a common approach on the security issues in Africa and address the need for harmonized legislations necessary to enhance cooperation in the area of cyber security in Member States of the African Union. Instead of establishing a unified legal framework for all Member States, it guides them towards establishing their own cybersecurity and data protection laws.[109] In

[107] Sharm E Sheikh Declaration, 22–26 October 2019, Sharm El Sheikh, AU/STC-CICT-3/MIN//Decl., para. 15. Online: 37590-2019_sharm_el_sheikh_declaration_-_stc-cict-3_oct_2019_ver2410-10pm-1rev-2.pdf (au.int) (quoted 11.11.2021).

[108] For more on the African Union´s High-level Panel on Emerging Technologies, please see: https://www.nepad.org/microsite/african-union-high-level-panel-emerging-technologies-apet (quoted 11.11.2021).

[109] TURIANSKI, Y.: Africa and Europe: Cyber Governance Lessons. In: *Policy Insights 77*, January 2020, p. 8. Online: https://media.africaportal.org/documents/Policy-Insights-77-

accordance with the Art. 8, each State Party shall commit itself to establishing a legal framework aimed at strengthening fundamental rights and public freedoms, particularly the protection of personal data, and punish any violation of privacy without prejudice to the principle of free flow of personal data.[110] The Convention also addresses the health data, which is defines as all information relating to the physical and mental state of the data subject, including genetic data. The scope of application of the Convention relating to personal data protection is limited to:

a) any collection, processing, transmission, storage or use of personal data by a natural person, the State, local communities, and public or private corporate bodies;
b) any automated or non-automated processing of data contained in or meant to be part of a file, with the exception of the processing defined in Article 9.2 of this Convention;
c) any processing of data undertaken in the territory of a State Party of the African Union;
d) any processing of data relating to public security, defence, research, criminal prosecution or State security, subject to the exceptions defined by specific provisions of other extant laws.[111]

Each State Party shall establish an authority in charge of protecting personal data, which shall be an independent administrative authority with the task of ensuring that the processing of personal data is conducted in accordance with the provisions of this Convention.[112] Furthermore, the Convention sets obligations relating to conditions governing personal data processing. The basic principles governing the processing of personal data include:

1) Principle of consent and legitimacy of personal data processing,
2) Principle of lawfulness and fairness of personal data processing,
3) Principle of purpose, relevance and storage of processed personal data,
4) Principle of accuracy of personal data,

turianskyi.pdf (quoted 11.11.2021).

[110] African Convention on Cyber Security and Personal Data Protection, 27 June 2014, Malabo, Art. 8. Online: https://au.int/sites/default/files/treaties/29560-treaty-0048_-_african_union_convention_on_cyber_security_and_personal_data_protection_e.pdf (quoted 11.11.2021).

[111] *Ibidem,* Art. 9 (1).

[112] *Ibidem,* Art. 11 (1), (2).

5) Principle of transparency of personal data processing, and
6) Principle of confidentiality and security of personal data processing.[113]

Personal data controllers are in accordance with Art. 20 obliged to process personal data confidentially. The data controller must take all appropriate precautions, according to the nature of the data, and in particular, to prevent such data from being altered or destroyed, or accessed by unauthorized third parties.[114] Although the wording of the Malabo Convention is general, we argue, that when it comes into force, it will provide a decent starting point for the protection of personal data used by AI in health and nursing.

Regarding the protection of human rights related to the use of AI, the African Commission on Human and Peoples' Rights adopted *resolution No. 473 on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa*. The resolution calls on State Parties to ensure that the development and use of AI, robotics and other new and emerging technologies is compatible with the rights and duties in the African Charter and other regional and international human rights instruments, in order to uphold human dignity, privacy, equality, non-discrimination, inclusion, diversity, safety, fairness, transparency, accountability and economic development as underlying principles that guide the development and use of AI, robotics and other new and emerging technologies.[115] Furthermore, it calls on State Parties to work towards a comprehensive legal and ethical governance framework for AI technologies, robotics and other new and emerging technologies so as to ensure compliance with the African Charter and other regional treaties.[116] Characteristic of the resolution is the notion that all AI technologies, robotics and other new and emerging technologies which have far reaching consequences for humans must remain under meaningful human control in order to ensure that the threat that they pose to fundamental human rights is averted. The emerging norm of maintaining meaningful human control over AI tech-

[113] *Ibidem,* Art. 13.
[114] *Ibidem,* Art. 20–21.
[115] African Commission on Human and Peoples' Rights: 473 Resolution on the need to undertake a Study on human and peoples' rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa – ACHPR/Res. 473 (EXT.OS/ XXXI) 2021, 25 February 2021, para. 1. Online: https://www.achpr.org/sessions/resolutions?id=504 (quoted 11.11.2021).
[116] *Ibidem,* para. 4.

nologies, robotics and other new and emerging technologies should be codified as a human rights principle.[117]

Overall, it may be concluded that in terms of regional ethical and legal framework the AU is at the beginning of its drafting. The same applies to the legal regulation of AI in health and nursing. Although, there is a legal framework on the protection of personal data and protection of human rights, the use AI has its specific implications which are not all regulated by the above mentioned instruments. Also, only eight AU Member States have ratified the Malabo Convention, therefore, the Convention is still not in force.[118] However, it should be noted that there are several non-state initiatives, such as the Artificial Intelligence for Development in Africa, or Knowledge for All Foundation whose aim is produce AI policy research that will inform and facilitate the development of public policies and regulations that promote the inclusive benefits of AI, while mitigating the potential costs and risk.

Conclusion

Artificial Intelligence is and will play an important part in the future of health and nursing. In a few years time it seems that it will be normal be diagnosed or treated by the AI systems. However, it is particularly difficult to predict the future international legal regulation. From the past experience we can, however, predict that the development of international legal regulation of AI will be conducted in a fragmented manner, and will focus only on specific aspects of the use of AI. On the international level, there are numerous initiatives on the universal, specialized or regional level which to some extent overlap. For the most part, international community adopts documents on ethics of AI in health and nursing, but only a limited number of documents contain legal regulations. The majority of adopted documents is of legally non-binding nature. Each of the analysed levels (universal, specialized, and regional) has its own characteristics. The UN, at the universal level, focuses mainly on digitalization and to a limited extent to AI. Currently, there is no specific initiative that would concern international legal regulation of AI in health and nursing care. At the specialized level, the main body focusing on AI in health and nursing, as well as ethical and legal regulation is the WHO. We consider it natural, due to its specialized agenda. Although cur-

[117] *Ibidem,* para. 5.

[118] The AU Convention on cyber security and personal data protection will come into force when 15 AU Member States ratify the Convention.

rently there is no draft of an international legal framework in this area a key first step was taken, when the WHO published its *Ethics and Governance of Artificial Intelligence in Health* which addresses key ethical aspects associated with the use of AI in health and nursing. We remain hopeful that future activities of the WHO will contribute to the model legislation for AI in health and nursing that will provide States with much needed guideline to eradicate legal gaps. At the regional level, the greatest progress can be seen within the EU, where the Member States are bound by primary and secondary sources of EU law. The current state of the protection of human rights, protection of personal data, safety and effectiveness of AI seems to be appropriate for the use of AI in health and nursing care. However, there are still legal gaps concerning liability issues. The proposal of Artificial Intelligence Act includes several important safeguards to ensure safe and effective use of AI in health and nursing.

Can we expect a comprehensive international legal regulation for AI in health and nursing? Based on the present analysis, the international community is still far from a comprehensive international legal regulation for AI in health and nursing, however, several initiatives of international organizations show that States are opened to model legal regulations that would be drafted in cooperation with non-state actors. Since the use of AI in health and nursing is on a rise, the international community does not have much time for filling existing legal gaps in this area.

Bibliography

1. African Commission on Human and Peoples‘ Rights: 473 Resolution on the need to undertake a Study on human and peoples’ rights and artificial intelligence (AI), robotics and other new and emerging technologies in Africa – ACHPR/Res. 473 (EXT.OS/ XXXI) 2021, 25 February 2021. Online: https://www.achpr.org/sessions/resolutions?id=504 (quoted 11.11.2021).
2. African Convention on Cyber Security and Personal Data Protection, 27 June 2014, Malabo. Online: https://au.int/sites/default/files/treaties/29560-treaty-0048_-_african_union_convention_on_cyber_security_and_personal_data_protection_e.pdf (quoted 11.11.2021).
3. CANDELON, F., EL BEDRAOUI, H., MAHER, H.: Developing an Artificial Intelligence for Africa strategy, OECD Development Matters Blog, 9 February 2021. Online: https://oecd-development-matters.org/2021/02/09/developing-an-artificial-intelligence-for-africa-strategy/#more-14109 (quoted 11.11.2021).

4. European Commission: White Paper On Artificial Intelligence – A European approach to excellence and trust, 19 February 2020, COM(2020) 65.
5. European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics, 18 July 2018, 2015/2103(INL), OJ C 252.
6. Expert Group on Liability and New Technologies New Technologies Formation: Liability for artificial intelligence and other emerging digital technologies, European Commission, 2019, 70 p., ISBN 978-92-76-12959-2. Online: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf (quoted 11.11.2021).
7. GERKE, S., MINSSEN, T., COHEN, G.: Ethical and legal challenges of artificial intelligence-driven healthcare. In: BOHR, A., MEMARZADEH, K.: *Artificial Intelligence in Healthcare*, London: Academic Press, 2020, p. 295–336, ISBN: 978-0-12-818438-7. Online: https://reader.elsevier.com/reader/sd/pii/B9780128184387000125?token=35D2D384BA50A7028F25E9A1E20EEA422D0C618C7D3FA7856615047D6BE1760F5929EDE96D418C4129CBB130F287CAD2&originRegion=eu-west-1&originCreation=20211205145558 (quoted 11.11.2021).
8. HASHIGUCHI, T., SLAWOMIRSKI, L., ODERKIRK, J.: OECD Health Working Papers No. 128: Laying the foundations for artificial intelligence in health, OECD, 3 June 2021, DELSA/HEA/WD/HWP(2021)5. Online: https://www.oecd-ilibrary.org/docserver/3f62817d-en.pdf?expires=1638634715&id=id&accname=guest&checksum=7E4C4EF518CB36DDC790CA7CE9E22987 (quoted 11.11.2021).
9. High-Level Expert Group on Artificial Intelligence: Ethics Guidelines for Trustworthy AI, Independent High-Level Expert Group on Artificial Intelligence Set Up by the European Commission, 8 April 2019, Brussels. Online: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_EN.pdf (quoted 11.11.2021).
10. Charter of Fundamental Rights of the European Union, OJ C 326, 26.10.2012.
11. ITU: United Nations Activities on Artificial Intelligence (AI), International Telecommunication Union: Geneva, 2019, ISBN: 978-92-61-29601-8, 88 p. Online: https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf (quoted 11.11.2021).
12. KERR, I., MILLAR, J.: Delegation, Relinquishment and Responsibility: The Prospect of Robot Experts. In: CALO, R., FROOMKIN, A. M., KERR, I.: *Robot Law*, Cheltenham, Edward Elgar, 2016, p. 102–127. ISBN: 978-1-78347-672-5.
13. McCARTHY, J., MINSKY, M., ROCHESTER, N., SHANNON, C.: A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. In: *AI Magazine*, Vol. 27, No. 4, 2006, p. 12–14. ISSN: 0738-4602.
14. NA, L. et al.: Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning. In: JAMA Network Open, 21 December 2018, 13 p. Online:

https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2719130 (quoted 11.11.2021).

15. NAYEBARE, M.: Artificial intelligence policies in Africa over the next five years. In: The ACM Magazine for Students, Vol. 26, No. 2, 2019, ISSN: 0001-0782, p. 50–54. Online: https://dl.acm.org/doi/10.1145/3368075 (quoted 11.11.2021).
16. OECD: OECD creates expert group to foster trust in artificial intelligence, 13 September 2018. Online: https://www.oecd.org/innovation/oecd-creates-expert-group-to-foster-trust-in-artificial-intelligence.htm (quoted 11.11.2021).
17. OECD: Principles on Artificial Intelligence, 22 May 2019, OECD/LEGAL/0449, p. 11. Online: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (quoted 11.11.2021).
18. OECD: Recommendation of the Council on Health Data Governance, 13 December 2016, OECD/LEGAL/0433, 10 p. Online: http://legalinstruments.oecd.org (quoted 11.11.2021).
19. OECD: State of Implementation of the OECD AI Principles Insights from National AI Policies, June 2021, DSTI/CDEP(2020)15/FINAL, 91 p. Online: https://www.oecd-ilibrary.org/docserver/1cd40c44-en.pdf?expires=1638292371&id=id&accname=guest&checksum=14A6B682AC03423B4C28FCE66A9754E9 (quoted 11.11.2021).
20. O´SULLIVAN, S. (et al.): Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery, The International Journal of Medical Robotics and Computer Assisted Surgery, 2019, Vol. 15, No. 1, ISSN: 1478-596X, p. 3–20.
21. Paris Convention on the Organisation for Economic Co-operation and Development, 14 December 1960.
22. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, Brussels, 21 April 2021, 2021/0106(COD). Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (quoted 11.11.2021).
23. RASANATHAN, K., EVANS, T.: Primary health care, the Declaration of Astana and COVID-19. In: *Bulletin of the World Health Organization*, 28 September 2020, 15 p. Online: https://www.who.int/bulletin/online_first/BLT.20.252932.pdf (quoted 11.11.2021).
24. 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), OJ L 119, 4.5.2016.
25. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, OJ L 117, 5.5.2017.

26. RIGBY, M.: Ethical Dimensions of Using Artificial Intelligence in Health Care. In: *AMA Journal of Ethics*, Volume 21, Number 2, February 2019, p. 121–124. Online: https://journalofethics.ama-assn.org/sites/journalofethics.ama-assn.org/files/2019-01/fred1-1902_1.pdf (quoted 11.11.2021).
27. SALATHÉ, M., WIEGAND, T., WENZEL, M., KRISHNAMURTHY, R.: Focus Group on Artificial Intelligence for Health, 2018, 10 p. Online: https://www.itu.int/en/ITU-T/focusgroups/ai4h/Documents/FG-AI4H_Whitepaper.pdf (quoted 11.11.2021).
28. SAMARGHITEAN, C., VIHINEN, M.: Medical expert systems. In: Current Bioinformatics, Vol. 3, No. 1, 2008, p. 56–65, ISSN: 1574-8936. Online: https://www.eurekaselect.com/82021/article (quoted 11.11.2021).
29. Sharm E Sheikh Declaration, 22–26 October 2019, Sharm El Sheikh, AU/STC-CICT-3/MIN//Decl. Online: https://au.int/sites/default/files/decisions/37590-2019_sharm_el_sheikh_declaration_-_stc-cict-3_oct_2019_ver2410-10pm-1rev-2.pdf (quoted 11.11.2021).
30. SCHNEEBERGER, D., STÖGER, K., HOLZINGER, A.: The European Legal Framework for Medical AI. In: HOLZINGER, A., KIESEBERG, P., MIN TJOA, A., WEIPPL, E. (eds.): Machine Learning and Knowledge Extraction, Cham: Springer, 2020, p. 209–226, ISBN: 978-3-030-57320-1. Online: https://www.aholzinger.at/wordpress/wp-content/uploads/2020/08/The-European-Legal-Framework-For-Medical-AI.pdf (quoted 11.11.2021).
31. TURIANSKI, Y.: Africa and Europe: Cyber Governance Lessons. In: Policy Insights 77, January 2020, 14 p.. Online: https://media.africaportal.org/documents/Policy-Insights-77-turianskyi.pdf (quoted 11.11.2021).
32. UNESCO: Report of the World Commission on the Ethics of Scientific Knowledge and Technology on Robotics Ethics, 14 September 2017, Paris, SHS/YES/COMEST-10/17/2 REV.
33. UN General Assembly: Road map for digital cooperation: implementation of the recommendations of the High-level Panel on Digital Cooperation: Report of the Secretary General, 29 May 2020, A/74/821. Online: https://undocs.org/en/A/74/821 (quoted 11.11.2021).
34. United Nations: Our Common Agenda: Report of the Secretary-General, United Nations, New York, 2021, ISBN: 978-92-1001012-2, 86 p. Online: https://www.un.org/en/content/common-agenda-report/assets/pdf/Common_Agenda_Report_English.pdf (quoted 11.11.2021).
35. VAN DE POEL: Why New Technologies Should Be Conceived as Social Experiments. In: *Ethics, Policy & Environment*, Vol. 16, No. 3, 2013, p. 352–355, ISSN: 2155-0085.
36. WAELBERS, K., SWIERSTRA, T.: The Family of the Future: How Technologies Can Lead to Moral Change. In: VAN DEN HOVEN, J., DOORN, J., SWIERSTRA, T. (eds.): *Responsible Innovation*, Springer, Dordrecht, 2014, p. 219–236, ISBN: 978-94-017-8956-1.

37. WHO: Declaration of Astana, adopted at the Global Conference on Primary Health Care, 25 and 26 October 2018, 12 p. Online: https://www.who.int/docs/default-source/primary-health/declaration/gcphc-declaration.pdf (quoted 11.11.2021).
38. WHO: Ethics and Governance of Artificial Intelligence in Health: WHO Guidance, Geneva, 2021, ISBN 978-92-4-002920-0, 165 p.
39. WHO: Global strategy on digital health 2020-2025, Geneva, 2019, ISBN 978-92-4-002092-4, 60 p. Online: https://www.who.int/docs/default-source/documents/gs4dh-daa2a9f352b0445bafbc79ca799dce4d.pdf (quoted 11.11.2021).
40. WHO: Working for health: An introduction to the World Health Organization, Geneva: WHO Press, 2017, ISBN: 92-4-156313-5, 26 p. Online: https://www.who.int/about/brochure_en.pdf (quoted 11.11.2021).

Author details

Mgr. Lucia BAKOŠOVÁ, PhD.
Pavol Jozef Šafárik University in Košice
Faculty of Law
Institute of International Law and European Law
lucia.bakosova@upjs.sk