

ISSN 2812-698X
ISSN (online) 2812-6998

REGIONAL LAW REVIEW

- ANNUAL EDITION -

BELGRADE, 2024

Održavanje konferencije „Regional Law Review“ i izdavanje ove publikacije podržalo je Ministarstvo nauke, tehnološkog razvoja i inovacija Republike Srbije.

International conference “Regional Law Review” and publishing of this collection of papers were supported by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia.

COLLECTION REGIONAL LAW REVIEW

Publishers

Institute of Comparative Law, Belgrade, Serbia

In Cooperation with

Faculty of Law, University of Pécs, Hungary
Faculty of Law, University of Ljubljana, Slovenia
Faculty of Law, University of Latvia, Riga, Latvia

For the Publisher

Prof. Jelena Čeranić Perišić, PhD

Editors

Jelena Kostić, PhD, Institute of Comparative Law, Belgrade, Serbia
Prof. Anita Rodina, PhD, Faculty of Law, University of Latvia, Riga, Latvia
Prof. Teresa Russo, PhD, Department of Legal Sciences, University of Salerno, Italy

Secretary

Bogdana Stjepanović, PhD, Institute of Comparative Law, Belgrade, Serbia

Technical editor

Ivana Radomirović, M.A., Institute of Comparative Law, Belgrade, Serbia

Editorial Board

Prof. Saša Zagorc, PhD, Faculty of Law, University of Ljubljana, Slovenia
Prof. Helga Špadina, PhD, Faculty of Law, "Josip Juraj Strossmayer" University of Osijek, Croatia
Prof. Biljana Vukoslavčević, PhD, Mediteran University in Podgorica, Montenegro
Prof. Goran Koevski, PhD, Faculty of Law, University of Skopje, North Macedonia
Prof. Marko Babić, PhD, University of Warsaw, Poland
Prof. Vesna Rijavec, PhD, Faculty of Law, University of Maribor, Slovenia
Prof. Giacomo Pailli, PhD, Faculty of Law, University of Florence, Italy
Mirjana Glintić, PhD, Institute of Comparative Law, Belgrade, Serbia

Advisory Board

Prof. Jelena Čeranić Perišić, PhD, Institute of Comparative Law, Belgrade, Serbia
Gojko Pavlović, PhD, Tax administration of the Republic of Srpska, Bosnia and Herzegovina
Prof. Slađana Jovanović, PhD, Faculty of Law, Union University in Belgrade, Serbia
Prof. Bojan Urdarević, PhD, Faculty of Law, University of Kragujevac, Serbia
Prof. Ljubinka Kovačević, PhD, Faculty of Law, University of Belgrade, Serbia
Prof. Ljubinko Mitrović, PhD, Pan-European University Apeiron in Banja Luka, Bosnia and Herzegovina

Scientific Board of the RLR Conference

Doc. Matija Damjan, PhD, Faculty of Law, University of Ljubljana, Slovenia
Prof. Nikol Žiha, PhD, Faculty of Law, "Josip Juraj Strossmayer" University of Osijek, Croatia
Prof. Ágoston Mohay, PhD, Faculty of Law, University of Pécs, Hungary
Prof. Valentina Ranaldi, PhD, "Niccolò Cusano" University of Rome, Italy
Gojko Pavlović, PhD, Tax administration of the Republic of Srpska, Bosnia and Herzegovina
Milica Matijević, PhD, Institute of Comparative Law, Belgrade, Serbia
Prof. Daniela Ježová, PhD, Faculty of Law, Comenius University in Bratislava, Slovakia

Organisational Board of the RLR Conference

Prof. Vid Jakulin, PhD, Faculty of Law, University of Ljubljana, Slovenia
Prof. Gordana Gasmir, PhD, Institute of Comparative Law, Belgrade, Serbia
Prof. Zoltán Bankó, PhD, Faculty of Law, University of Pécs, Hungary
Prof. Jelena Dujmović Bocka, PhD, Faculty of Law,
"Josip Juraj Strossmayer" University of Osijek, Croatia
Prof. Dragana Čorić, PhD, Faculty of Law, University of Novi Sad, Serbia
Ana Cović, PhD, Institute of Comparative Law, Belgrade, Serbia
Ena Gotovuša, L.L.M., Faculty of Law, University of Sarajevo, Bosnia and Herzegovina

Reviewers

- Prof. Saša Gajin, PhD, Union University in Belgrade, Serbia
Prof. Nevenko Vranješ, PhD, University of Banja Luka, Bosnia and Herzegovina
Prof. Lucia Mokra, PhD, Comenius University in Bratislava, Slovakia
Milica Matijević, PhD, Institute of Comparative Law, Belgrade, Serbia
Prof. Bojan Urdarević, PhD, University of Kragujevac, Serbia
Prof. Nataša Mrvić Petrović, PhD, Institute of Comparative Law, Belgrade, Serbia
Sanja Jelisavac Trošić, PhD, Institute of International Politics and Economics, Belgrade, Serbia
Prof. Marko Dimitrijević, PhD, University of Niš, Serbia
Ana Knežević Bojović, PhD, Institute of Comparative Law, Belgrade, Serbia
Prof. Ana Đanić Čeko, PhD, "Josip Juraj Strossmayer" University of Osijek, Croatia
Prof. Hana Kovačikova, PhD, Comenius University in Bratislava, Slovakia
Prof. Mina Zirojević, PhD, Institute of Comparative Law, Belgrade, Serbia
Prof. Jelena Čeranić Perišić, PhD, Institute of Comparative Law, Belgrade, Serbia
Prof. Marko Božić, PhD, Union University in Belgrade, Serbia
Prof. Goran Obradović, PhD, University of Niš, Serbia
Prof. Bojan Spaić, PhD, University of Belgrade, Serbia
Prof. Gordana Gasmir, PhD, Institute of Comparative Law, Belgrade, Serbia
Prof. Jelena Vučković, PhD, University of Kragujevac, Serbia
Prof. Hristina Runcheva Tasev, PhD, Faculty of Law "Iustinianus Primus", North Macedonia
Prof. Miomira Kostić, PhD, University of Niš, Serbia
Prof. Darko Simović, PhD, University of Criminal Investigation and Police Studies in Belgrade, Serbia
Doc. Branka Babović Vuksanović, PhD, University of Belgrade, Serbia
Prof. Bojan Vlaški, PhD, University of Banja Luka, Bosnia and Herzegovina
Sanja Stojković Zlatanović, PhD, Institute of Social Sciences, Serbia
Prof. Ivana Bodrožić, PhD, University of Criminal Investigation and Police Studies in Belgrade, Serbia
Prof. Srđan Golubović, PhD, University of Niš, Serbia
Jelica Gordanić, PhD, Institute of International Politics and Economics, Belgrade, Serbia
Vanja Korać, PhD, Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade, Serbia
Doc. Marija Dragičević, PhD, University of Niš, Serbia
Doc. Aleksandar Antić, PhD, University of Kragujevac, Serbia
Doc. Novak Vujičić, PhD, University of Belgrade, Serbia
Mihajlo Vučić, PhD, Institute of International Politics and Economics, Belgrade, Serbia
Marko Novaković, PhD, Institute of International Politics and Economics, Belgrade, Serbia
Dragan Prlja, PhD, Institute of Comparative Law, Belgrade, Serbia
Ranko Šovilj, PhD, Institute of Social Sciences, Belgrade, Serbia
Aleksandra Rabrenović, PhD, Institute of Comparative Law, Belgrade, Serbia

Proofreading

ABC prevodi d.o.o.

Prepress

Branimir Trošić

Print

Birograf Comp doo Beograd

Printed in 150 copies

ISBN 978-86-82582-25-0

ISSN 2812-698X

ISSN (online) 2812-6998

doi: 10.56461/iup_rirc.2024.5

Licensed under CC licence:

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

Indexed in HeinOnline Law Journal Library



Indexed in Directory of Open Access Journals



*Aleksandar MIHAJLOVIĆ**

Institute of Comparative Law, Belgrade, Serbia

*Vesna ĆORIĆ***

Institute of Comparative Law, Belgrade, Serbia

ARTIFICIAL INTELLIGENCE AND DISCRIMINATION - STRENGTHS AND WEAKNESSES OF THE CURRENT EUROPEAN ANTI-DISCRIMINATION LEGAL FRAMEWORK***

We are witnessing an enormous development of artificial intelligence (AI) which boosts economic productivity, creates new job opportunities, and gives hope that human life will be more prosperous. On the other side, AI, as a new system, that is undiscovered and unpredictable, creates ethical and legal dilemmas and threats to human rights violations, in the context of the principle of equality, the rule of law, and democratic principles, if it is used in an inappropriate way. The subject of the paper is discrimination in the process of AI application in different fields of people's everyday lives. The aim of this investigation is to analyze provisions in the recently adopted European Union (EU) AI Act and the Council of Europe Framework Convention which are expected to prevent discriminatory treatment through an AI life cycle, and to give a bird-view of the selected cases of AI-related discrimination, as well as of the position of the Serbian national authorities in that regard. On that road, the authors will provide a critical and comparative analysis of these two instruments governing the AI application. Subsequent to that, the paper is focused on the position taken by the Serbian authorities in order to examine the level of its readiness to stay in line with the legal challenges the AI implementation brings, and illustrates a

* LLM, Research Assistant, ORCID: 0000-0001-8309-7896, e-mail: a.mihajlovic@iup.rs.

** PhD, Senior Research Associate, ORCID: 0000-0003-4240-7469, e-mail: v.coric@iup.rs.

Vesna Ćorić, PhD, was appointed as a mentor to Aleksandar Mihajlović, LLM, research assistant at the Institute of Comparative Law, by the Decision of the Scientific Council of the Institute of Comparative Law No. 771 of 05/10/2023. This paper is one of the results of the mentoring process with the mentee Aleksandar Mihajlović.

*** This paper is a result of the research conducted at the Institute of Comparative Law financed by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia under the Contract on realization and financing of scientific research of SRO in 2024 registered under no. 451-03-66/2024-03/200049.

current example which is related to the implementation of the Social Card Law. The methodological framework includes doctrinal, comparative, and descriptive methods.

Keywords: artificial intelligence, anti-discrimination, European Union AI Act, Council of Europe Framework Convention, Serbia, Social Card Law.

1. INTRODUCTION

We are witnessing an enormous development of artificial intelligence (hereinafter: AI) which looks very promising and pleasurable, especially from an economic perspective, higher productivity, higher salaries and more free time. For instance, there is a prediction that AI could contribute to the global economy up to \$15.7 trillion in 2030, where \$6.6 trillion is likely to come from increased productivity and \$9.1 trillion is likely to come from consumption-side effects (Rao & Verweij, 2017, p. 3). Besides these catchy numbers, AI is still an unknown and unpredictable area, and its implementation can provoke different legal challenges in exercising different rights, and especially in implementing the principles of equality and non-discrimination, as universal values of law.

Although the first impression regarding AI systems can be their objectivity in the decision-making process, different examples support the fact that AI is biased and can provoke discriminatory treatment toward different segments of people's lives. Another challenge is the unpredictability of AI effects throughout its lifecycle. At the moment, the initiative to regulate AI by legally binding instruments has come from the European Union (hereinafter: EU) and the Council of Europe (hereinafter: CoE), which adopted the EU Regulation on Artificial Intelligence¹ (hereinafter: EU AI Act) and the CoE Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law² (hereinafter: Framework Convention) in 2024.

The subject of the paper is discrimination in the process of AI application in different fields of people's everyday lives. The aim of this investigation is to analyze provisions in the recently adopted the EU AI Act and the Framework Convention which are expected to prevent discriminatory treatment through an AI life cycle, and to give a bird-view of the selected cases of AI-related discrimination, as well as of the position of the Serbian national authorities in that regard.

Firstly, the authors provide an analysis of definitions of AI and elements of an AI lifecycle. More specifically, the definitions introduced by the Organization for Economic Cooperation and Development (hereinafter: OECD), the CoE and the EU are to be examined. In

¹ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) Text with EEA relevance, OJ L, 2024/1689, 12.7.2024. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689 (15. 7. 2024).

² Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, CM(2024)52-final, 17 May 2024. Available at: <https://search.coe.int/cm/#{%22CoEIdentifier%22:%220900001680afb11f%22,%22sort%22:%22CoEValidationDate%20Descending%22}}> (7. 7. 2024).

the subsequent section, the authors will provide an overview of how AI programs are created and what risks are in the process of their implementation along with selected examples of how the AI application can provoke discriminatory treatment to concrete social groups. After that, the authors will have a look on the current anti-discrimination provisions in the EU AI Act and the Framework Convention. Within the last section, it will be shown where Serbia is at the moment regarding the AI development and what should be done to comply with the recently introduced the EU and the CoE framework. It is important to mention that Serbia was among the first countries in the world to adopt the Strategy of Development of Artificial Intelligence 2020-2025³ (hereinafter: AI Strategy), as well as the Ethical Guidelines for the Development, Application and Usage of Reliable and Responsible Artificial Intelligence⁴ (hereinafter: Ethical Guidelines). In this context, the authors will illustrate some current challenges which are related to the implementation of the Serbian Social Card Law⁵ (hereinafter: SCL) which applies machine learning in the administrative procedure of exercising the right to social aid. This statute seems particularly important for risk assessment of its impact on indirect algorithmic discrimination before the Serbian national authorities given that the procedure for assessing its constitutionality was initiated before the Serbian Constitutional Court. The methodological framework includes doctrinal, comparative, and descriptive methods.

2. AI DEFINITION AND ELEMENTS OF AN AI LIFECYCLE

In the beginning of this part, the authors briefly elaborate on the definition of AI, starting from a universally accepted definition given by the OECD to the European definitions offered by the EU AI Act and the Framework Convention respectively. Concurrently, the elements of an AI lifecycle are described, aiming to illustrate its complexity. The explanation of an AI lifecycle is given as a theoretical overview which is common for every AI system.

2.1. AI Definition

In the modern world, it is very difficult to distinguish the digital revolution from AI which is becoming an inevitable part of every segment of human life (Miasato & Silva Reis, 2019, p. 193). When talking about AI, Surden (2019, p. 1308) was right when said that “[...], AI systems are not intelligent thinking machines, in any meaningful sense. Rather, [...], they are often able to produce useful, intelligent results without intelligence”. In some circumstances related to handling cases in taxation, social insurance, transport tariffs etc, human intelligence can be redundant and not very productive because of the higher possibility of making risks related to manual decision-making, and the development of automation is more than desired in such cases (Sannerholm, 2021, p. 225).

³ Strategy of Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025, *Official Gazette RS*, No. 96/2019.

⁴ Ethical Guidelines for the Development, Application and Usage of Reliable and Responsible Artificial Intelligence, *Official Gazette RS*, No. 23/2023.

⁵ Social Card Law, *Official Gazette RS*, No. 14/2021.

For lawyers, it is not an easy task to predict new technological innovations and their correlation with a legal system, while technological development and inventions will anticipate necessary legal changes very precisely, because the legal environment will be *conditio sine qua non* for their further development (success) or failure (Fornasier Oliveira, 2021, p. 354). Therefore, the urgent need had been identified to regulate the AI system and its elements, and lay down their key definitions on supranational and international levels. An inevitable progress of new technologies cannot be an excuse for violations of human rights, democratic values and the rule of law (Nemitz, 2021, p. 240).⁶

The OECD's revised definition states that "*An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.*" (Grobelnik, Perset & Russell, 2024). It means that AI is a machine system without human elements. It is fed by information, and it can produce different outputs. Although AI is a non-living system, it can affect both physical (living) and virtual environment. The last part of the definition underlines that AI systems have different capacities to produce inputs, which depend on their level of complexity. The same definition was adopted by the Framework Convention⁷, and this approach illustrates its character of an international treaty, which tends to create a consensus of an internationally accepted AI definition. The EU AI Act accepted almost an identical definition⁸ as the Framework Convention and the OECD, and this approach towards uniformity in defining an AI system is commendable and can be attributed to the fact that these two instruments were drafted at the same period of time.

2.2. Elements of an AI Lifecycle

The complexity of AI systems is attributable to the three types of machine learning where an AI system is applied in different formats. First, *supervised learning* is basic and is based on the dataset, and trained to discover a pattern in the limited framework of concrete data. For example, this is a way how detection of spam emails works. Second, *unsupervised learning* is more complex in comparison to the former, because the program learns how to find a pattern among data and produce a concrete result. For

⁶ The development of AI can be seen as a chance for improving the state of fundamental rights, democracy and the rule of law in general, because AI will replace lawyers from doing repetitive jobs, and leave an opportunity to them to use this time to focus on important aspects of concrete areas (Kaur & Puri Gopal, 2021, p. 347).

⁷ "For the purposes of this Convention, "artificial intelligence system" means a machine-based system that for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment." (Article 2, Framework Convention).

⁸ "AI system means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments." (Article 3 paragraph 1 point (1), EU AI Act).

instance, the program is trained to detect and classify fruits, vegetables and animals based on given criteria such as a colour. Third, *reinforcement learning* represents the most complex system, which uses dataset, interacts with a virtual or real environment, searches for an optimal way to complete a task while implementing steps which maximize a chance to complete the task successfully (Leslie *et al.*, 2021, p. 9).⁹

The definition provided by the OECD, the EU and the CoE instruments shows that AI has its *lifecycle*, which amounts to a process of how it is developed and how it works in practice. In general, the literature recognizes 12 stages of an AI lifecycle, which are separated into three levels. The first level is called *design* and it includes the following four steps: 1. project planning when a project team decides if they will apply AI on a concrete task or not; 2. problem formulation which will be addressed by an AI model; 3. data extraction or procurement which should provide necessary data to train an AI model. Data can be extracted from surveys or similar methodologies, or procured which means to obtain existing datasets based on legal agreements; 4. data analysis which starts when all the necessary data is provided (Leslie *et al.*, 2021, p. 10). The second level is called *development* and the following four steps are part of it: 1. preprocessing is implemented as a phase of feeding a model and includes among other tasks, data cleaning and data wrangling; 2. model selection and training which depends on a concrete task which should be done; 3. model testing and validation; 4. model reporting where experts can detect if an AI model should be modified based on detected obstacles (Leslie *et al.*, 2021, p. 11). The third level is *the deployment of an AI model* which includes the following four steps: 1. model implementation in a real world; 2. user training for AI implementers; 3. monitoring of a model implementation; 4. updating or deprovisioning of an AI model based on results of monitoring and experiences from its implementation (Leslie *et al.*, 2021, p. 12).

While the Framework Convention has covered on the surface only the activities related to the AI lifecycle that have the potential to interfere with human rights, democracy, and the rule of law, the EU AI Act has recognized concrete steps in detail which should be followed through the development of high-risk AI systems. The EU AI Act has introduced a *risk management system* that shall be established, implemented, documented and maintained in relation to the development of high-risk AI systems. The elements of this system are: the identification and analysis of the known and reasonably foreseeable risks; the estimation and evaluation of the risks that may emerge when the high-risk AI system is used in accordance with its intended purpose, and under conditions of reasonably foreseeable misuse; the evaluation of other risks possibly arising; and the adoption of appropriate and targeted risk management measures.¹⁰ Concurrently, there are stipulated concrete rules regarding the data and data governance in a way that high-risk AI systems which make use of techniques involving the training of AI models with data shall be developed on the basis of training, validation, and testing data sets that meet the quality criteria referred to in paragraphs 2 to 5 of Article 10 whenever such

⁹ In the literature, we can find a classification of the three waves of AI: first, “*handcraft knowledge*”, as the most primitive category of AI; second, “*statistical learning*” also known as “*machine learning*”; third, “*deep learning*” as the most complex type of AI (see more at: Rejeski, Reynolds & Wright, 2018, pp. 6-7).

¹⁰ Article 9, the EU AI Act.

data sets are used.¹¹ Special attention is placed to the detection of possible biases which can produce discriminatory treatment. The EU AI Act requires also that high-risk AI systems shall be designed and developed in such a way as to ensure that their operation is sufficiently transparent to enable deployers to interpret a system's output and use it appropriately.¹² High-risk AI systems during the process of their development shall be effectively overseen by natural persons during the period in which they are in use.¹³ Prior to deploying a high-risk AI system, deployers shall perform an assessment of the impact on fundamental rights that the use of such system may produce.¹⁴ The EU AI Act has also recognized the importance of monitoring the implementation of AI systems through the notifying authority, and each Member State shall designate or establish at least one notifying authority.¹⁵ As a way to stimulate the AI development as well as to prevent possible negative effects through its development, the EU AI Act has introduced the AI regulatory sandbox.¹⁶ Its aim is to provide conditions for an innovative AI system development, training, validation and testing, where appropriate in real-world conditions, for a limited time and under regulatory supervision.

3. DISCRIMINATION BASED ON AI APPLICATION – FROM UNKNOWN TO PREDICTABLE CONSEQUENCES OF AI

AI models apply different datasets which are used for their training and based on them they are ready to make content, predictions, recommendations and conclusions. This kind of data is called “*big data*” which are large in their complexity and interrelationships, and AI models discover previously unknown connections between data elements (Schmidt & Stephens, 2019, p. 134). The opposite is “*alternative data*” which are used for decision-making or model building, but they do not have a *historical background*¹⁷ of their application

¹¹ Article 10, the EU AI Act. Prescribed practices in accordance with the paragraph 2 of the Article 10 from the EU AI Act include the following: the relevant design choices; data collection processes and the origin of data, and in the case of personal data, the original purpose of the data collection; relevant data-preparation processing operations, such as annotation, labelling, cleaning, updating, enrichment and aggregation; the formulation of assumptions, in particular with respect to the information that the data are supposed to measure and represent; an assessment of the availability, quantity and suitability of the data sets that are needed; examination in view of possible biases that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operations; appropriate measures to detect, prevent and mitigate possible biases; the identification of relevant data gaps or shortcomings that prevent compliance with this Regulation, and how those gaps and shortcomings can be addressed.

¹² Article 13, the EU AI Act.

¹³ This is called human oversight and is regulated by Article 14, the EU AI Act.

¹⁴ Fundamental Rights Impact Assessment for High-Risk AI Systems is regulated by the Article 27, the EU AI Act.

¹⁵ Article 28, the EU AI Act.

¹⁶ Article 57, the EU AI Act.

¹⁷ This kind of data is called “*historical data*”, because they are coming from some records, such as medical records, Internet search done by concrete population, criminal records etc (Huq, 2021, p. 3). This

in the decision-making process (Schmidt & Stephens, 2019, p. 134). An example can be a credit decision by a bank based on data such as the occupation of a client, rental payments, utility payments, and their educational background, where this data had not been used for a concrete credit decision or any other decision before (Schmidt & Stephens, 2019, p. 134).

AI models use machine learning algorithms which rely on different data, and based on the complexity of AI programs, it is not always possible to predict effects which AI can provoke. In the context of supervised and unsupervised machine learning it is easier to make accurate predictions, but, in case of reinforcement or so-called deep learning, it is very hard to predict all possible consequences. That is the reason why this kind of unpredictability creates AI as a “black box”. In the literature, there are two types of AI “black boxes”: strong and weak “black boxes”. First, effects of strong “black boxes” can be very severe for humans, and there is no way to determine how the AI makes decisions or predictions, what kind of information influences an AI decision, and what is the rank of importance of variables processed by the AI. On the other hand, weak “black boxes” refer to AI where engineers can predict to some extent its accuracy effects although it can still be “opaque” to humans (Bathae, 2018, p. 906). For this reason, the EU AI Act adopted the risk-based approach in the classification of AI systems (Korać, Prlja & Gasmi, 2021, p. 163) following the principle – the higher risk, the stricter rules will be applied to an AI model. It is noteworthy that the Framework Convention has not adopted the risk-based classification of AI systems, because it presumes that every AI system constitutes a risk for people and their environment.

A greater amount of data, which is used for training of an AI model, provides stronger processing power and advances a mathematical algorithm¹⁸ which helps AI to operate autonomously (Greenstein, 2022, p. 299). Algorithms rely on statistical inferences which can discover discriminatory correlations (Xenidis, 2020, p. 745). Although we can perceive algorithms as objective, they are biased¹⁹ because AI models are trained by humans who have a taste for discrimination and the dataset has “historical records” of discrimination. It is important to mention that not every bias is associated with protected characteristics such as sex, age, ethnicity, and religion. Instead, such a feature can be an algorithm which classifies people based on a fact if they have cats or not (European Union Agency for Fundamental Rights, 2022, p. 24). In another case, an algorithm can contain biases related to a protected characteristic, but the final result may not be discriminatory if it does not lead to a less favourable treatment (European Union Agency for Fundamental Rights, 2022, p. 24).

There are different examples of how AI models can discriminate, for instance, in the field of a labour market it is interesting to mention the case of *the Amazon Company* when

data has some historical background where it was used for making concrete decisions and deciding about someone’s rights and/or obligations. That is the reason why this data is seen as biased.

¹⁸ In the simplest way, an algorithm can be explained as a program which gives concrete instructions to a computer in which order and how concrete task should be done (Prlja, Gasmi & Korać, 2022, p. 84).

¹⁹ There are different forms of biases: *historical bias, representation bias, measurement bias, aggregation bias, evaluation bias, deployment bias, automation and confirmation biases* etc. (see more in: Bartoletti & Xenidis, 2023, pp. 16-19).

the created automated hiring tool discriminated against women who applied for software engineering positions (Goodman, 2018; Dastin, 2018).²⁰ The program was trained by the data from CVs of male engineers who were dominant employees in these positions. The program scored female's resumes with lower points because it marked words such as "women's" or "women's rugby team" as less worthy. Concurrently, some colleges and schools were ranked with lower points as well, because the program recognized as more valuable concrete schools attended dominantly by male engineers, and which were used as data to feed the program. The Amazon Company has tried to fix this problem and modified the algorithm, but at the end it stopped using it in 2018. This example of algorithmic discrimination in the hiring process supports *the statistical discrimination theory* (Chen, 2023, pp. 2-3) which means that individual characteristics of an applicant are observed in the context of his/her membership of a concrete group of people. For instance, an employer may not have a taste for discrimination, but high economic costs of discovering the true potential of a candidate will lead him to put all female applicants in a group of less productive than male applicants because based on the statistical data female employees are often absent from job because of the pregnancy and child care obligations than males.

In Finland, the National Non-Discrimination and Equality Tribunal found that a company discriminated against a client when he applied for a credit loan (Bartoletti & Xenidis, 2023, p. 21; *Algorithms in credit scoring: Discrimination based on use of statistical data in Finland*, 2021). This example is related to the algorithm which was used to assess the credit capability of a client and took into account data related to age, gender, language and place of residence, but did not include the applicant's actual credit history. The person in this case was a Finnish male speaker and came from a rural area, and the statistical model assessed these characteristics as disadvantages. Cases of discrimination can occur also in the housing market when landlords use AI programs to assess applicants as potential tenants. Using addresses, names and ages of applicants can lead to discriminatory treatment toward people of colour, elderly people, or people who live in some areas which are associated with higher levels of committed crimes (*Housing Discrimination: Big Data, AI, and Algorithmic Models*, 2023). The last example includes the Dutch Government which faced a huge scandal when the Tax authorities used an algorithm that detected incorrect applications for child benefits and potential fraud, based on the data related to race and ethnicity of applicants (*Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms*, 2021). This program affected primarily disadvantaged families who did not have Dutch nationality and who received higher risk scores (*Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms*, 2021).

From concrete examples, we can conclude that in algorithmic discrimination, *indirect discrimination* often occurs, because *seemingly neutral datasets*, which are used to feed AI programs, contain hidden biases, which in the end have discriminatory effects (Korać, Prlja & Gasmi, 2022, p. 287). Discrimination can be based on one, two or several protected grounds, in other words, it can occur as multiple discrimination, depending on an AI program and variables which are used. There is a view that discrimination can be proved easily

²⁰ See more in: Weerts *et al.*, 2023, pp. 809-811.

when an algorithm is involved and that anti-discrimination principle must be respected before, when an AI model is designed (Kleinberg *et al.*, 2018, p. 114). Although this in fact sounds absolutely true and makes sense, the problem also occurs because AI companies lack transparency of how an algorithm is trained, which data is used, etc. (Heinrichs, 2022, pp. 143, 150-153). This problem of transparency will be presented in the part of the paper related to the implementation of the Serbian Social Card Law.

All the presented examples show that rooted discriminatory patterns are transported from humans to datasets which are used to train mathematical algorithms. The huge risk of applying AI is the violation of the principle of equality which leads to discriminatory treatment. In many cases, it is not predictable at all how an AI model will work in practice. Such a kind of unpredictability provokes a phenomenon which has been called a “black box”. The selected cases further prove that the current European anti-discrimination legal framework is significantly challenged by the AI development. Such a challenge was recognized and addressed by the EU and the CoE. A part of their response to the spreading AI application to different areas of human life has led to the creation of specific legal instruments regulating AI along with focusing on human rights protection and anti-discrimination measures.

4. THE ANTI-DISCRIMINATION PROVISIONS WITHIN THE EU AI ACT AND THE FRAMEWORK CONVENTION

The EU AI Act and the Framework Convention represent a pioneering achievement of AI regulation. Although their legal natures as well as the institutional framework under which they have been developed differ among each other, their common aim is the regulation of AI in a way which prevents negative effects on people. Both acts contain anti-discriminatory norms which have strengthened the implementation of existing anti-discriminatory provisions and introduced new rules. First, it will be analysed the impact of these two acts on the national legislative frameworks of Member States, and, after that, anti-discriminatory rules from these acts will be compared.

4.1. A Different Approach of Implementing the EU AI Act and the Framework Convention and Their Impact on National Legislation

As a response to the growing development of AI and to different risks that it brings, the EU and the CoE recently adopted two legal documents, the EU AI Act and the Framework Convention. Both legal acts represent a revolutionary step to regulate the AI application on the regional and international level. They contain provisions governing the anti-discrimination matters in the AI context and are important for both the EU member states and its candidate countries.

The EU AI Act is a regulation, and, as a binding secondary law of the EU, it will be applied uniformly toward all EU member states. Since regulations are directly applicable, they are *legally binding without any action from an individual member state, and they take effect as soon as they are published in the Official Journal of the EU. Although EU*

regulations require no implementing legislation within individual member states, there has been a widespread practice among candidate countries to align their national regulatory frameworks with provisions of EU regulations to the highest possible extent in the accession process. Therefore, it is expected that aspiring candidate countries, such as Serbia, harmonize their legislation with the EU AI Act provisions in short run (Korać, Prlja & Gasmi, 2023, p. 213).

On the other side, the Framework Convention is an international treaty that will be opened for signatories on 5 September 2024 to all interested states, including Serbia, as well as to the EU as an entity, irrespectively of its member states. Every state party shall incorporate it as a part of its national law and it is important for both the EU Member States and candidate countries to interpret the provisions of the EU AI Act and the Framework Convention jointly, paying equal and close attention to both of them. However, it is noteworthy that the enforcement of the Framework Convention is “multifaceted” (Güçlütürk Gazi, 2024) and it goes beyond national implementation and includes the establishment of at least one effective oversight mechanism, regular consultations, and discussions among the State Parties and international cooperation (Articles 23 and 26, the Framework Convention).

4.2. The EU AI Act and the Framework Convention – A Comparison of Their Anti-Discriminatory Rules

While both legal instruments contain provisions governing anti-discrimination matters in the AI context, the Framework Convention is less detailed in comparison to the EU AI Act since it sets forth a set of general obligations and principles for its state parties, and leaves specific details to domestic legislation.²¹ For that reason, the Framework Convention belongs to the category of so-called “framework agreements” (Lück-Matz, 2011).

As it appears from its title and the preamble, the Framework Convention’s purpose is to ensure that all elements of an AI lifecycle are fully consistent with *human rights, democracy and the rule of law*. On the other hand, it seems that the EU AI Act focuses primarily on *the EU internal market* and AI’s effects on it.²² However, Article 1 along with improving the functioning of the internal market, also underlines the need to “promote the uptake of human-centric [...] artificial intelligence” and to ensure *inter alia* a high level of fundamental rights (enshrined in the EU Charter of Fundamental Rights), including democracy and the rule of law. Therefore, the purposes of both legal instruments are compatible as the three European fundamental values are put at the forefront of protection, and are all inseparably treated as a “holy trinity”.

Even though the EU and the CoE are not directly and organically linked since there are structural differences between them, those fundamental values have almost identical content within both organizations (Güçlütürk Gazi, 2024). On the road to developing

²¹ The EU AI Act has 113 articles and XIII Annexes, while the Framework Convention has 36 articles. See also: Güçlütürk Gazi, 2024.

²² In that light, Lütz states that the purpose of the EU AI Act is not to ensure gender equality and non-discrimination as such. However, the author admits that “a flavour of gender equality” can be felt throughout its provisions (Lütz, 2024, p. 79).

a coherent system of human rights protection in Europe, relevant steps were taken. An illustration of such an approach may serve the provision introduced by the Lisbon Treaty stipulating that the EU shall access the European Convention for the Protection of Human Rights and Fundamental Freedoms (Ćorić & Knežević Bojović, 2020, p. 27).

Since both instruments follow the same values, they are not mutually contradictory, but rather complementing. They also underline the need to comply with provisions of other international human rights treaties. Most of those human rights instruments also contain relevant anti-discrimination norms. In that light, the Framework Convention refers *inter alia* to the necessity of application of the Universal Declaration of Human Rights, the European Convention for the Protection of Human Rights and Fundamental Freedoms, the International Covenant on Civil and Political Rights, the International Covenant on Economic, Social and Cultural Rights, and the European Social Charter. In a similar vein, the EU AI Act underlines the importance of the following human rights instruments: the United Nations (hereinafter: UN) Convention on the Rights of Persons with Disabilities, the UN Convention relating to the Status of Refugees, and the UN Convention on the Rights of the Child, as well the UNCRC General Comment No. 25 as regards the digital environment.

In addition to this body of the UN human rights instruments, the EU AI Act also gives due regard to the applicable EU regulatory framework and holds that provisions of the EU treaties, the Charter of Fundamental Rights of the EU (hereinafter: EU Charter of Fundamental Rights) and the EU secondary legislation shall be respected. When it comes to combating discrimination, the EU AI Act particularly stipulates that it does not affect the practices that are prohibited by the EU anti-discrimination law (Preamble, point 45, EU AI Act).

The approach of extensive referral to various international and supranational human rights instruments in the texts of the EU AI Act and the Framework Convention is not surprising. Namely, all the EU member states are state parties to the European Convention for the Protection of Human Rights and Fundamental Freedoms, as well as to the above UN human rights treaties. Therefore, the EU Member States are reminded through the texts of the EU AI Act and the Framework Convention of their obligation to interpret the EU AI Act in conjunction with the existing UN human rights framework, the Framework Convention, and the other relevant pieces of the European human rights' supranational framework (the EU and the CoE) in the process of development and implementation of AI lifecycles.

Considering that AI can provoke different risks, it is important that both instruments adopt *the risk-based approach* regarding the AI application. The difference is that the EU AI Act stipulates *categories of risk-systems*,²³ And the Framework Convention despite its reference to the risk-based approach, does not classify specific use of AI systems as prohibited or high-risk systems (Güçlütürk Gazi, 2024).

The EU AI Act clearly defines conditions which have to be fulfilled for one AI system to be considered a high-risk.²⁴ In addition, it contains a list of prohibited AI practices and

²³ Unacceptable, high, limited and minimal risk.

²⁴ See Article 6, EU AI Act.

a detailed classification of high-risk AI systems, which are listed in Annex III. It is argued that adequate protection against gender biases and discrimination can be only achieved in cases where AI systems are classified as a “high-risk” since the EU AI Act is of a horizontal nature and its priority is not the reduction of discrimination (Lütz, 2024, p. 81).²⁵ Therefore, an adequate identification of AI systems as high-risk systems is of key importance for ensuring fundamental rights protection in the non-discrimination realm.

Nevertheless, in the scholarly literature, it was rightly indicated that the EU AI Act does not provide sufficiently clear guidance on what falls under the scope of “high-risk” AI systems and that it is expected from the Court of Justice of the European Union to provide clarifications through its further interpretations in order to increase the legal security in the given field (Lütz, 2024, p. 82). As it was mentioned earlier, the classification of high-risk AI systems is based on Annex III’s list of use cases which are determined in a rather clear manner. They *inter alia* include employment, workers management and access to self-employment, education and vocational training, as well as access to and the enjoyment of essential private services and essential public services and benefits. The above cases are selected as important considering that gender bias and discrimination frequently occur in one of those categories (Lütz, 2024, pp. 82-83).

However, the Annex III list is subject to two types of uncertainties. First refers to the derogations which are foreseen in Article 6 paragraph 3 for the use cases and which can be introduced if one of four specified criteria is fulfilled. If one of those criteria is fulfilled and the AI system referred to in Annex III does not pose a significant risk of harm to the fundamental rights of natural persons, it will not be considered as a high risk.²⁶ In the legal doctrine, it was rightly claimed that those four criteria are not clearly determined what may seriously undermine the envisaged system of protecting fundamental rights linked to the (high-risk) AI systems. The second point that can be brought up with regard to the flexibility of the scope of high-risk AI systems relates to the option envisaged by the EU AI Act according to which the Annex III list can be amended by delegated acts if some AI system, in particular, poses a risk or an adverse impact on fundamental rights.²⁷ The reference in Article 7(1) (b) shows the importance of negative impacts on fundamental rights including the right to non-discrimination, which enables the Commission to add new AI systems to Annex III, for instance when the principle of equality between women and men is adversely affected. This opportunity for the Commission to review Annex III by delegated act, as well as its general obligation to review the AI Act on a regular basis (after three years and thereafter every four years), leads to fluidity of the exact scope of application of the notion of high-risk AI system and

²⁵ See more in: Renard-Castets & Eynard, 2023, p. 613.

²⁶ These are: (a) the AI system is intended to perform a narrow procedural task; (b) the AI system is intended to improve the result of a previously completed human activity; (c) the AI system is intended to detect decision-making patterns or deviations from prior decision-making patterns and is not meant to replace or influence the previously completed human assessment, without proper human review; or (d) the AI system is intended to perform a preparatory task to an assessment relevant for the purpose of the use cases listed in Annex III.

²⁷ Beside this condition, it is also necessary that AI systems are intended to be used in any of the areas listed in Annex III (Article 7(1) (a)).

as such may endanger achieved level of legal security within the EU regulatory framework. However, the benefits of such a flexible approach prevail over its limitations since this opportunity leaves room for further upgrades of the EU AI Act to reflect contemporary AI developments.

The Framework Convention does not contain any kind of classifications neither in the form of risk classes nor lists for high-risk AI systems. Instead, it presumes that every kind of AI can pose a potential risk to human rights, democracy, and the rule of law. It provides high-level obligations and offers a framework for the risk assessment and mitigation of its adverse impacts (Güçlütürk Gazi, 2024). In Article 10, the Framework Convention stipulates the importance of respecting *the principle of equality and non-discrimination*, including *gender equality*, as provided under applicable international and domestic law, within the lifecycle of AI systems. Concurrently, this instrument calls for the implementation of its provisions by the State Parties *without discrimination* on any ground, in accordance with their international human rights obligations (Article 17, Framework Convention).

The Framework Convention does not stipulate any kind of penalties or fines for individuals or firms in cases of not complying with the norms. It is up to every State Party to introduce a mechanism for monitoring the implementation and compliance with national legal frameworks. The EU AI Act provides concrete penalties in cases of breach of the rules, and every party can introduce concrete warnings and non-monetary measures. The difference in sanctioning approach can be explained by the fact that the Framework Convention belongs to the category of “framework agreements” which sets forth only broad commitments while leaving more detailed rules to national legislation. Moreover, considering the applicable anti-discriminatory framework which has been established through the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Protocol No. 12 to this Convention, it can be expected that the European Court of Human Rights will develop in the near future relevant case-law on sanctioning violations of anti-discrimination provisions triggered by the AI system.

The European Commission established the AI Office.²⁸ Which will have a key role in the process of implementation of the EU AI Act, and every EU member state shall establish a public authority which will monitor the implementation of the EU AI Act. In Article 77 paragraph 1, the EU AI Act stipulates that “national public authorities or bodies which supervise or enforce the respect of obligations under the EU law protecting fundamental rights, including *the right to non-discrimination*, in relation to the use of high-risk AI systems referred to in Annex III shall have the power to request and access any documentation created or maintained under this EU AI Act in accessible language and format when access to that documentation is necessary for effectively fulfilling their mandates within the limits of their jurisdiction.” The AI Office through its role is expected to further clarify and develop the provisions of the EU AI Act pertaining to anti-discrimination.

²⁸ Commission decision of establishing the European Artificial Intelligence Office, Brussels, 24. 01. 2024 C(2024) 390 final. Available at: <https://digital-strategy.ec.europa.eu/en/library/commission-decision-establishing-european-ai-office> (16. 7. 2024).

5. THE SERBIAN APPROACH TO AI REGULATION AND CURRENT CHALLENGES IN THE FIELD OF DISCRIMINATION

The Republic of Serbia was among the first countries in the world to adopt the AI Strategy for the period 2020-2025, which introduced aims and measures necessary for the development of AI. The Strategy recognized the importance of regulating the relationship between the individual and the society, and targeted challenges such as the representativeness of data used for machine learning, redefinition or elimination of the need for certain professions due to the introduction of AI, the need for additional qualification of the population for future jobs and issues concerning the responsibility for consequences of autonomous system decisions based on AI, etc.

As to overcome a possible problem which is experienced in practice, the AI Strategy has recognized the importance of prevention from *discrimination based on machine learning*. Besides the general objective of the AI Strategy which is to use AI in favour of economic growth, employment and improvement of the quality of life, there are five specific objectives. The fifth specific objective which is called the *Ethical and safe application of artificial intelligence* is dedicated to *discrimination based on AI*. The AI Strategy describes that algorithmic discrimination occurs based on data which are used for the training model, such as “historic” data, gender/sex unbalanced data, or lack of the inclusiveness of all relevant data sources necessary for the development of an AI lifecycle. It is recognized that individuals who are subject to decisions made by the AI model must have *the right to an explanation* and *the right to transparency* in connection with the algorithm. For this reason, it is necessary to enable the prevention of discrimination as well as early understanding and interpretation of the model and enable explanation of the decision.

The AI Strategy recognized the necessity of adopting the Ethical Guidelines. This document was introduced in 2023 and its aim is to enable necessary conditions for science, especially in the field of AI, to develop and progress, but not to allow people, as the central figures of all processes that affect them and of which they are indirect or direct agents, to be endangered and neglected. Concurrently, AI systems that are developed must be in line with the well-being of humans, animals and the environment.

The Ethical Guidelines have reflected to some extent the EU AI Act approach and recognized as a high-risk AI system a system that is part of the safety component of a product, or is a product in itself that has a function and behaves as a safety system and as such requires an assessment of compliance with legislative norms on putting AI systems into use, by a third party. Concurrently, the Ethical Guidelines have envisaged that a high-risk AI system is a system that is listed and marked as a high-risk system in the Ethical Guidelines. The high-risk AI systems are recognized in the numerous specified fields which are not exhaustively listed. Instead, this list is an open list which leaves room for more flexible interpretation and detection of new high-risk systems if a concrete AI system shows this kind of risk. The Ethical Guidelines do not apply to systems that are prohibited under the law governing AI systems.

The following *principles* are recognized as a starting point for the creation, application, and use of AI systems that will be worthy of human trust due to their reliability

and responsibility towards humans: explainability and verifiability; dignity; prohibition of damages; and, equity. The Ethical Guidelines introduce the following *conditions* of reliable and responsible AI which are based on the above-mentioned principles: action (mediation, control, and participation) and supervision; technical reliability and safety; privacy, personal data protection, and data management; transparency; *diversity, non-discrimination and equality*; social and environmental well-being; and responsibility.

For each of the conditions, there are AI System Assessment Questionnaires which help individuals and/or organizations to identify areas for improvement and encourage them to take action to overcome perceived challenges. By filling out the Questionnaires, one gets an insight into the established measures and identifies the measures that should still be implemented for the purpose of building a reliable AI. The Questionnaires do not exclude the application of other tools and methods for assessing the fulfilment of the AI system requirements in terms of the adopted Ethical Guidelines and/or laws. They are not a guide through the legal system of the Republic of Serbia and filling them out does not release responsible subjects from legal obligations and responsibilities. The conditions related to *diversity, non-discrimination and equality* are covered by a questionnaire regarding these three values and help to identify possible negative effects of an AI model on these fields. The principles and conditions specified by the examined Serbian documents (the Strategy and the Ethical Guidelines) reflect the spirit and the provisions of the recently introduced European supranational regulatory framework. However, they lack binding character and effective enforcement, because both documents, the strategy, as a policy document, and the Ethical Guidelines constitute “soft law” instruments.

The current case of alleged algorithmic discrimination in Serbia is related to the implementation of the Social Card Law (hereinafter: SCL) which was introduced in 2021 and whose main objective is to create a more effective realization of social protection rights and services, fairer distribution of social assistance, improvement of the efficient and proactive work of authorities in the field of social protection, provision of support in defining and shaping social policy and monitoring the overall effects of social protection measures, as well as provision of up-to-date data on beneficiaries in the event of emergencies.

Based on the SCL, the social card is a unique register that contains data on the individual and related persons on social and economic status, data on the type of rights and services from social protection that a person uses or has used, as well as data on the officials who led, that is, decided on individual rights. There are around 135 personal processed data of every social aid user or applicant, and the Centre for Social Work uses the data from the system to make a decision if a person has or does not have a right to receive social aid. The Initiative for Economic and Social Rights (hereinafter: Initiative A11), as a Serbian NGO, notified complaints from people who lost the right to social aid because the centres for social work use primarily the algorithm which processes their personal data and data of related persons to make a final decision about a concrete applicant. Initiative A11 asked the responsible Ministry of Labor, Employment, Veterans and Social Affairs to publish the structure of the algorithm and the lifecycle of the program, but

this request was refused. At the moment, there is a case before the Constitutional Court of the Republic of Serbia because Initiative A11 has initiated the procedure for assessing the constitutionality of the SCL. This is the first case of possible indirect algorithmic discrimination before a Serbian authority and it remains to be seen in the following period how the Constitutional Court will adjudicate and to which extent it will take into account the recently adopted EU and the CoE framework governing the operation of AI systems.²⁹

In general, the existing Law on the Prohibition of Discrimination can be applied, as an umbrella anti-discrimination act, to cases including algorithmic discrimination, because Article 4 Paragraph 2 stipulates that “*everyone*³⁰ shall be obligated to respect the principle of equality, that is to say, the prohibition of discrimination”. It is also important to underline that, although the current AI Strategy has recognized discrimination which can occur in the process of AI implementation, it does not mention anywhere in the text the Law on the Prohibition of Discrimination in the context of the existing legal framework in Serbia (for instance, it mentions the Law on Personal Data Protection). We expect that this approach will be changed in a new strategy and that the Law on the Prohibition of Discrimination will be placed as a starting point in the part related to the recognition of preventing algorithmic discrimination.

As an EU candidate country, Serbia should follow the future development of the EU AI Act and improve its legal framework on AI (Prlja, Gasmi & Korać, 2021, p. 126), which includes also the update of the Ethical Guidelines in accordance with the further development and implementation of the analyzed the EU and the CoE legal instruments. In the case of ratification of the Framework Convention, the Serbian authorities will be obliged to further improve the current legislation. At the moment, the AI Strategy and the Ethical Guidelines as soft law instruments provide only a vision and a framework for safe AI development without having any legally binding force.

6. CONCLUSION

AI systems have a lifecycle in which effects on people and the environment are not predictable at all. That is the reason why AI can provoke risk for violations of human rights, democracy and the rule of law. The level of the risk depends on the complexity of an AI system. The accelerating development of AI initiated the adoption of two legal instruments, the EU AI Act and the Framework Convention, under the auspice of the EU and the CoE. Although there are some differences between these two acts, they share the common aim of regulating the AI application in a safe manner and limiting its negative effects to the minimum possible extent.

Although the regulatory approach of these two instruments seems to look different, their shared value is the protection of fundamental human rights, the rule of law

²⁹ See more about this case on: (Anti) Social Card. Available at: <https://antisocialnekarte.org/en> (17. 7. 2024).

³⁰ Based on the Law on the Prohibition of Discrimination, everyone means also any legal entity registered or operating on the territory of the Republic of Serbia.

and democracy, and they both call on the substantive application of international and domestic human rights norms. While the EU AI Act will be implemented directly by the EU member states, the Framework Convention creates a foundation for how State Parties to this international treaty should adapt their national legislative frameworks. Furthermore, these two instruments supplement each other and support the overall idea of creating a safe environment for the AI application. At this stage, because the EU AI Act has just entered into force and the Framework Convention was recently adopted and will be opened for signature in Vilnius (Lithuania) on 5 September 2024 on the occasion of a conference of Ministers of Justice, we cannot predict at all possible challenges which will occur during its application, but it is absolutely clear that the anticipated interpretation of the EU AI Act by the Court of Justice of the European Union through its case-law will additionally show the relevance, efficiency and effectiveness of this instrument. Although supranational courts already dealt with issues deriving from the use of AI, it seems that throughout the implementation of the recently adopted legal instruments, both courts will clarify and strengthen the given protection, particularly in the field of combating discrimination. Among others, it is expected that the Court of Justice of the European Union will bring needed clarification when it comes to the interpretation of when an AI system should be classified as a high-risk system.

Discrimination based on AI applications, the so-called algorithmic discrimination, can occur in different spheres of everyday life, and this usually happens in an indirect form. This is because algorithms seem to be neutral and bias-free, but in reality, they reflect the biases which exist around us and which are just incorporated by humans into the AI lifecycle. Current international and European supranational human rights instruments, together with the EU AI Act and the Framework Convention, create a strong and secure base for protection from discrimination in the context of AI applications.

The Republic of Serbia was among the first countries in the world to adopt the AI Strategy and the Ethical Guidelines. By doing so, it has recognized the importance and necessity of regulating safe AI development and application. Although these two national instruments are not legally binding, they provide a solid base for the self-assessment of safe AI development and its application. Serbia, as an EU candidate country, should start the harmonization of domestic law with the EU AI Act. At least, it is expected that a new strategy and the Ethical Guidelines are to be carefully reviewed in order to determine to which extent they have to be amended to comply with both instruments, the EU AI Act and the Framework Convention. If Serbia ratifies the Framework Convention, this will introduce a concrete obligation to adapt the domestic law in accordance with the Framework Convention's provisions. At this moment, it is important to underline, that the AI Strategy has not mentioned anywhere in the text the Law on the Prohibition of Discrimination, as the main anti-discrimination legal act, although it has recognized a problem of algorithmic discrimination. The authors are looking forward to the new strategy and the revised Ethical Guidelines which will place in the centre the proper application of the Law on the Prohibition of Discrimination and will hopefully reflect the spirit and wording of two recently adopted European legal instruments.

LIST OF REFERENCES

Monographs

- Bartoletti, I. & Xenidis, R. 2023. *Study on the impact of artificial intelligence systems, their potential for promoting equality, including gender equality, and the risks they may cause in relation to non-discrimination*. Strasbourg: The Council of Europe. Available at: <https://rm.coe.int/study-on-the-impact-of-artificial-intelligence-systems-their-potential/1680ac99e3> (10. 7. 2024).
- European Union Agency for Fundamental Rights. 2022. *Bias in Algorithms – Artificial Intelligence and Discrimination – Report*. Luxembourg: Publications Office of the European Union. Available at: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf (10. 7. 2024).
- Leslie, D. et al. 2021. *Artificial intelligence, human rights, democracy, and the rule of law: a primer*. The Council of Europe. Available at: <https://rm.coe.int/primer-en-new-cover-pages-coe-english-compressed-2754-7186-0228-v-1/1680a2fd4a> (7. 7. 2024). <https://doi.org/10.2139/ssrn.3817999>
- Prlja, D. Gasmi, G. & Korać, V. 2021. *Veštačka inteligencija u pravnom sistemu EU*. Beograd: Institut za uporedno pravo. Available at: <https://iup.rs/wp-content/uploads/2021/05/2021-Ve%C5%A1ta%C4%8Dka-inteligencija-u-pravnom-sistemu.pdf> (22. 9. 2024).
- Prlja, D. Gasmi, G. & Korać, V. 2022. *Ljudska prava i veštačka inteligencija*. Beograd: Institut za uporedno pravo. Available at: <https://iup.rs/wp-content/uploads/2022/09/Ljudska-prava-i-ve%C5%A1ta%C4%8Dka-inteligencija-Prlja-Gasmi-Korac.pdf> (10. 7. 2024).
- Rao, S. A. & Verweij, G. 2017. *Sizing the prize – PwC’s Global Artificial Intelligence Study: Exploiting the AI Revolution – What’s the real value of AI for your business and how can you capitalize?* PwC. Available at: <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf> (7. 7. 2024).
- Rejeski, D., Reynolds, L. & Wright, S. 2018. *When Software Rules: Rule of Law in the Age of Artificial Intelligence*. Washington, D. C: Environmental Law Institute. Available at: <https://www.eli.org/sites/default/files/eli-pubs/when-software-rules-web.pdf> (8. 7. 2024).
- Renard-Castets, C. & Eynard, J. 2023. *Un Droit de L’intelligence Artificielle: Entre Règles Sectorielles et Régime Général: Perspectives Comparées*. Belgium: Bruylant.
- Sannerholm, R. 2021. Responsibility and Accountability: AI, Governance, and the Rule of Law. In: Collona, L. & Greenstein, S. (eds.), *Law in the Era of Artificial Intelligence*. Stockholm: Stiftelsen Juridisk Fakultetslitteratur (SJF) and The Swedish Law and Informatics Research Institute (IRI) and Law Faculty, Stockholm University, pp. 223-246. Available at: <https://irilaw.org/2022/02/16/new-publication-nordic-yearbook-of-law-and-informatics-2020-2021/> (8. 7. 2024). <https://doi.org/10.53292/208f5901.40d940a1>
- Weerts, H. et al. 2023. Algorithmic Unfairness through the Lens of EU Non-Discrimination Law: Or Why the Law is not a Decision Tree. In: *2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT’23)*, June 12–15, 2023, Chicago, IL, USA. New York : ACM, pp. 805-816. <https://doi.org/10.1145/3593013.3594044>

Scholarly Articles

- Bathae, Y. 2018. The Artificial Intelligence Black Box and the Failure of Intent and Causation. *Harvard Journal of Law & Technology*, 31(2), pp. 889-938. Available at: <https://jolt.law.harvard.edu/assets/articlePDFs/v31/The-Artificial-Intelligence-Black-Box-and-the-Failure-of-Intent-and-Causation-Yavar-Bathae.pdf> (8. 7. 2024).
- Chen, Z. 2023. Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications*, 10(567), pp. 1-12. Available at: <https://www.nature.com/articles/s41599-023-02079-x> (10. 7. 2024). <https://doi.org/10.1057/s41599-023-02079-x>
- Fornasier Oliveira de, M. 2021. Artificial Intelligence and Democratic Rule of Law. *Revista de Estudos Constitucionais, Hermenêutica e Teoria do Direito (RECHTD)*, 13(3), pp. 351-369. <https://doi.org/10.4013/rechtd.2021.133.06>
- Ćorić, V. & Knežević Bojović, A. 2020. Autonomous Concepts and Status Quo Method: Quest for Coherent Protection of Human Rights before European Supranational Courts. *Strani pravni život*, 64(4), pp. 27-40. <https://doi.org/10.5937/spz64-30165>
- Greenstein, S. 2022. Preserving the rule of law in the era of artificial intelligence (AI). *Artificial Intelligence and Law*, 30, pp. 291-323. <https://doi.org/10.1007/s10506-021-09294-4>
- Heinrichs, B. 2022. Discrimination in the age of artificial intelligence. *AI & Society*, 37(1), pp. 143-154. <https://doi.org/10.1007/s00146-021-01192-2>
- Huq, Z. A. 2021. Artificial Intelligence and the Rule of Law. *Public Law and Legal Theory Working Paper Series*, No. 764, pp. 1-14. Available at: https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=2194&context=public_law_and_legal_theory (8. 7. 2024).
- Kaur, I. & Puri Gopal, C. 2021. Impact of Artificial Intelligence on Legal Industry. *International Journal of Law Management & Humanities*, 4(2), pp. 346-354.
- Kleinberg, J. et al. 2018. Discrimination in the Age of Algorithms. *Journal of Legal Analysis*, 10, pp. 113-174. <https://doi.org/10.1093/jla/laz001>
- Korać, V., Prlja, D. & Gasmi, G. 2021. Challenges brought on by Artificial Intelligence. *Archaeology and Science*, 17, pp. 159-165. https://doi.org/10.18485/arhe_apn.2021.17.10
- Korać, V., Prlja, D. & Gasmi, G. 2022. Influence of Artificial Intelligence on Human Rights. *Archaeology and Science*, 18, pp. 279-292. https://doi.org/10.18485/arhe_apn.2022.18.19
- Korać, V., Prlja, D. & Gasmi, G. 2023. European Approach to Artificial Intelligence. *Archaeology and Science*, 19, pp. 209-216. https://doi.org/10.18485/arhe_apn.2023.19.14
- Lütz, F. 2024. The AI Act, gender equality and non-discrimination: what role for the AI office? *ERA Forum*, 25, pp. 79-95. <https://doi.org/10.1007/s12027-024-00785-w>
- Miasato, A. & Silva Reis, F. 2019. Artificial Intelligence as an Instrument of Discrimination in Workforce Recruitment. *Acta Universitatis Sapientiae, Legal Studies*, 8(2), pp. 191-212. <https://doi.org/10.47745/AUSLEG.2019.8.2.04>
- Nemitz, P. 2021. Democracy through law The Transatlantic Reflection Group and its manifesto in defence of democracy and the rule of law in the age of “artificial intelligence”. *European Law Journal*, 29(1-2), pp. 237-248. <https://doi.org/10.1111/eulj.12407>

- Schmidt, N. & Stephens, B. 2019. An Introduction to Artificial Intelligence and Solutions to the Problem of Algorithmic Discrimination. *Conference on Consumer Finance Law (CCFL) Quarterly Report*, 73(2), pp. 130-144.
- Surden, H. 2019. Artificial Intelligence and Law: An Overview. *Georgia State University Law Review*, 35(4), pp. 1305-1337.
- Xenidis, R. 2020. Tuning EU equality law to algorithmic discrimination: Three pathways to resilience. *Maastricht Journal of European and Comparative Law*, 27(6), pp. 736–758. <https://doi.org/10.1177/1023263X20982173>

Legal Sources

- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)Text with EEA relevance, *OJ L*, 2024/1689, 12.7.2024 Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689 (15. 7. 2024).
- Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, CM(2024)52-final, 17 May 2024. Available at: [https://search.coe.int/cm/#{%22CoEIdentifier%22:\[%220900001680afb11f%22\],%22sort%22:\[%22CoEValidationDate%20Descending%22\]}](https://search.coe.int/cm/#{%22CoEIdentifier%22:[%220900001680afb11f%22],%22sort%22:[%22CoEValidationDate%20Descending%22]}) (7. 7. 2024).
- Strategy of Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025, *Official Gazette RS*, No. 96/2019.
- Ethical Guidelines for the Development, Application and Usage of Reliable and Responsible Artificial Intelligence, *Official Gazette RS*, No. 23/2023.
- Social Card Law, *Official Gazette RS*, No. 14/2021.
- Law on the Prohibition of Discrimination, *Official Gazette RS*, No. 22/2009 and 52/2021.
- Commission decision of establishing the European Artificial Intelligence Office, Brussels, 24. 01. 2024 C(2024) 390 final. Available at: <https://digital-strategy.ec.europa.eu/en/library/commission-decision-establishing-european-ai-office> (16. 7. 2024).

Online Sources

- Algorithms in credit scoring: Discrimination based on use of statistical data in Finland. Equinet. 8 July 2021. Available at: <https://ai.equineteurope.org/library/algorithms-credit-scoring-discrimination-based-use-statistical-data-finland> (10. 7. 2024).
- (Anti) Social Card. Available at: <https://antisocialnekarte.org/en> (17. 7. 2024).
- Dasti, J. *Insight* – Amazon scraps secret AI recruiting tool that showed bias against women. 11 October 2018. Available at: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G/> (10. 7. 2024).
- Dutch childcare benefit scandal an urgent wake-up call to ban racist algorithms. Amnesty International. 25 October 2021. Available at: <https://www.amnesty.org/>

- en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/ (10. 7. 2024).
- Goodman, R. *Why Amazon's Automated Hiring Tool Discriminated Against Women*. 12 October 2018. Available at: <https://www.aclu.org/news/womens-rights/why-amazons-automated-hiring-tool-discriminated-against> (10. 7. 2024).
- Grobelnik, M., Perset, K. & Russell S. 2024. *What is AI? Can you make a clear distinction between AI and non-AI systems?* 6 March 2024. Available at: <https://oecd.ai/en/work/definition> (7. 7. 2024).
- Güçlütürk Gazi, O. *Understanding Council of Europe's Draft Framework Convention on AI, Human Rights, Democracy, and Rule of Law*. 17 January 2024. Available at: <https://www.holisticai.com/blog/europe-committee-artificial-intelligence-draft-framework-convention> (25. 7. 2024).
- Lück-Matz, N. 2011. *Framework Agreements*. In: Oxford Public International Law. Available at: <https://opil.ouplaw.com/display/10.1093/law:epil/9780199231690/law-9780199231690-e703#:~:text=A%20framework%20agreement%20can%20be,agreements%20between%20the%20parties%2C%20usually> (16. 8. 2024).
- Smith NMTC Associates. *Housing Discrimination: Big Data, AI, and Algorithmic Models*. SMITH NMTC ASSOCIATES, LLC. 10 October 2023. Available at: <https://smith-nmtc.com/housing-discrimination-big-data-ai-and-algorithmic-models/> (10. 7. 2024).

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

34(082)

343(082)

349::007(082)

INTERNATIONAL Conference Regional Law Review (5 ; 2024 ; Beograd)

[Fifth International Conference] Regional Law Review, Belgrade, 2024 : annual edition / [editors Jelena Kostić, Anita Rodina, Teresa Russo]. - Belgrade [etc.] : Institute of Comparative Law [etc.], 2024 (Beograd : Birograf comp). - [X], 283 str. ; 24 cm. - (Collection Regional law review, ISSN 2812-698X)

"In front of you is the fifth volume of RLR collection of papers..."

--> foreword. - Tiraž 150. - Str. VII: Foreword / editors. -

Napomene i bibliografske reference uz tekst. - Bibliografija uz svaki rad.

ISBN 978-86-82582-25-0

1. Kostić, Jelena, 1981- [уредник] [аутор додатног текста]

а) Право -- Зборници б) Кривично право -- Зборници в)

Информациона технологија -- Право -- Зборници

COBISS.SR-ID 156233737