Univerzitetska misao - časopis za nauku, kulturu i umjetnost [ISSN: 1451-3870]
Vol. 25, str. 38-56, 2026. god., web lokacija gde se nalazi rad:http://um.uninp.edu.rs
Tematska oblast: Društvene i humanističke nauke / podoblast: Pravo
Datum prijema rada: 02.12.2015. *Datum prihvatanja rada: 14.01.2026.*
UDK: 004.8:340.132 doi: 10.5937/univmis2525038M
342.72/.73

Originalni naučni rad

HUMAN RIGHTS AND CIVIL LAW RISKS OF ARTIFICIAL INTELLIGENCE: A SYSTEMATIC ANALYSIS OF SURVEILLANCE, ALGORITHMIC BIAS, AUTOMATED DECISION-MAKING, AND RESTRICTIONS ON FREEDOM OF EXPRESSION

Jelena Matijašević
University Business Academy in Novi Sad
Novi Sad, Serbia
jelena@pravni-fakultet.info, https://orcid.org/0000-0001-8068-0816

Maida Bećirović Alić
University of Novi Pazar
Novi Pazar, Serbia
maida.becirovic@uninp.edu.rs, https://orcid.org/0000-0002-9738-6581

Jasmina Nikšić
University of Novi Pazar
Novi Pazar, Serbia
j.pljakic@uninp.edu.rs, https://orcid.org/0000-0002-8190-7565

Irma Mašović Muratović
University of Novi Pazar
Novi Pazar, Serbia
i.masovic@uninp.edu.rs, https://orcid.org/0000-0003-0175-3584

Abstract

The development of artificial intelligence significantly shapes contemporary social processes, but at the same time opens up complex issues regarding the protection of fundamental human rights and the effectiveness of existing legal protection mechanisms. This paper provides a systematic analysis of the key risks that AI technologies produce in the areas of rights to privacy, equality, fair trial and freedom of expression, with particular reference to the role of civil protection as a complement to public law and regulatory instruments. The author analyzes the phenomenon of digital surveillance and the erosion of privacy due to the collection, analysis and predictive modeling of large data sets, including biometric information and data from daily user interactions, pointing out the limitations of the concept of informed consent and the need for individual private legal means of protection. Next, algorithmic bias is discussed as a mechanism that reproduces and deepens existing social inequalities, which threatens the right to equal treatment and raises the issue of civil liability for discriminatory outcomes of automated decision-making. Special attention is devoted to the application of AI in the judiciary, where non-transparent decision-making models, limited explainability and the risk of "automated suggestions" call into question the

realization of procedural guarantees of a fair trial, as well as the possibility of compensation for damages due to the violation of procedural rights. The final part discusses the risks to freedom of expression, including the consequences of automated content moderation, personalized ranking of information and the generation of synthetic content, while pointing out the importance of civil law requirements for the protection of individual rights, reputation and dignity. By analyzing the relevant literature and the current regulatory framework, the paper identifies normative gaps and emphasizes the need for a coherent legal system that, in addition to public law supervision, also provides effective private law protection mechanisms. The paper concludes that sustainable regulation of artificial intelligence is possible only through a combination of technical transparency, democratized supervision and strengthening of procedural and material guarantees, including a developed system of civil liability in the digital environment.

Keywords: artificial intelligence, human rights, civil protection, digital surveillance, fair trial rights, freedom of expression, data protection, discrimination.

RIZICI VEŠTAČKE INTELIGENCIJE PO LJUDSKA PRAVA I GRAĐANSKOPRAVNU ZAŠTITU: SISTEMATSKA ANALIZA NADZORA, ALGORITAMSKE PRISTRASNOSTI, AUTOMATIZOVANOG DONOŠENJA ODLUKA I OGRANIČAVANJA SLOBODE IZRAŽAVANJA

Apstrakt

Razvoj veštačke inteligencije značajno oblikuje savremene društvene procese, ali istovremeno otvara složena pitanja u pogledu zaštite temeljnih ljudskih prava i efektivnosti postojećih mehanizama pravne zaštite. Ovaj rad pruža sistematsku analizu ključnih rizika koje AI tehnologije proizvode u oblastima prava na privatnost, jednakost, pravično suđenje i slobodu izražavanja, sa posebnim osvrtom na ulogu građanskopravne zaštite kao dopune javnopravnim i regulatornim instrumentima. Autorka analizira fenomen digitalnog nadzora i erozije privatnosti usled prikupljanja, analiziranja i prediktivnog modelovanja velikih skupova podataka, uključujući biometrijske informacije i podatke iz svakodnevnih interakcija korisnika, ukazujući na ograničenja koncepta informisanog pristanka i potrebu za individualnim privatnopravnim sredstvima zaštite. Zatim se razmatra algoritamska pristrasnost kao mehanizam koji reprodukuje i produbljuje postojeće društvene nejednakosti, čime se ugrožava pravo na jednako postupanje i otvara pitanje građanskopravne odgovornosti za diskriminatorne ishode automatizovanog odlučivanja. Posebna pažnja posvećena je primeni AI u pravosuđu, gde netransparentni modeli odlučivanja, ograničena objašnjivost i rizik od „automatizovane sugestije“ dovode u pitanje ostvarivanje procesnih garancija pravičnog suđenja, ali i mogućnost naknade štete usled povrede procesnih prava. U završnom delu razmatraju se rizici po slobodu izražavanja, uključujući posledice automatizovanog moderiranja sadržaja, personalizovanog rangiranja informacija i

generisanja sintetičkih sadržaja, pri čemu se ukazuje na značaj građanskopravnih zahteva za zaštitu prava ličnosti, ugleda i dostojanstva pojedinca. Analizom relevantne literature i važećeg regulatornog okvira rad identifikuje normativne praznine i naglašava potrebu za koherentnim pravnim sistemom koji, pored javnopravnog nadzora, obezbeđuje i delotvorne privatnopravne mehanizme zaštite. Rad zaključuje da je održiva regulacija veštačke inteligencije moguća samo kroz kombinaciju tehničke transparentnosti, demokratizovanog nadzora i jačanja procesnih i materijalnih garancija, uključujući razvijen sistem građanskopravne odgovornosti u digitalnom okruženju.

Ključne reči: veštačka inteligencija, ljudska prava, građanskopravna zaštita, digitalni nadzor, pravo na pravično suđenje, sloboda izražavanja, zaštita podataka, diskriminacija.

INTRODUCTION

The accelerated application of artificial intelligence in the public and private sectors has opened a new chapter in the debate on the protection of human rights. Technological progress has led to the fact that decisions affecting an individual's position are increasingly made, shaped or mediated by algorithmic systems. Such an environment generates specific and interwoven risks, especially in the areas of surveillance, bias, automated decision-making in the judiciary and management of the free flow of information. Practice shows that these risks are not incidental, but structural and stem from the way AI is integrated into modern social processes.
The purpose of this paper is to offer a systematic analysis of the most important challenges that artificial intelligence produces for human rights. Special emphasis is placed on four areas that are the most sensitive according to previous research: mass surveillance and invasion of privacy, reproduction of social inequalities through algorithmic bias, the influence of automated recommendations on the realization of procedural guarantees in procedures, and the restriction of freedom of expression through automated moderation and the manipulation of digital space. The starting hypothesis of the work is based on the assumption that the existing international and national normative frameworks are not sufficiently developed to respond to these risks in their full complexity, which leads to the violation of the essence of the guaranteed rights.

In addition to public law regulation and institutional oversight, this paper proceeds from the assumption that the effective protection of human rights in the age of artificial intelligence requires a strengthened role of civil law mechanisms. While international and national regulatory frameworks primarily operate through ex ante standards and supervisory control, they often fail to provide individualized remedies for concrete harm caused by algorithmic systems. Civil law protection, through liability for damages, injunctive relief, and the protection of personality rights, enables affected individuals to seek direct redress for violations of privacy, equality, procedural guarantees, and freedom of expression. By addressing responsibility, causation, and proportionality on a case-by-case basis, civil law complements

human rights protection by translating abstract normative guarantees into enforceable private claims. This dual perspective allows the analysis to move beyond purely regulatory responses and to examine how private law can function as a corrective and preventive instrument in the governance of artificial intelligence.

Methodologically, the paper relies on a normative-dogmatic analysis of international instruments for the protection of human rights and modern regulations governing data processing and the development of artificial intelligence. The comparative approach enables comparing different models of regulation, while the analytical-synthetic processing of scientific literature and the available practice of supervisory bodies provides a basis for identifying the key causes of risk and their manifestations in practice. Such a methodological framework allows the observed problems to be considered in their systemic dimension and to define guidelines for the improvement of human rights protection mechanisms in the digital environment.

ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS: BASIC ASSUMPTIONS

Artificial intelligence (AI) occupies an increasingly important place in modern societies, not only as a technological tool, but also as a legal and ethical challenge. In the legal context, AI is most often defined as a system capable of automating decision-making processes through the analysis of large amounts of data, with the ability to learn and adapt without constant human intervention. Such systems include machine learning, deep neural networks, algorithmic profiling and automated decision making. The legal regulation of artificial intelligence implies the need to look at these technical systems through normative frameworks that protect the basic rights and freedoms of the individual.

AI systems function through the so-called "black boxes", where neither users nor regulators often have a clear picture of the possible consequences? The connection between AI and human rights is becoming more pronounced with the expansion of its applications in areas that directly affect human dignity, identity and personal freedoms. In this sense, AI poses serious challenges to the principles of accountability, transparency and fairness, which represent the basis of any democratic legal order (Rodrigues, 2020). Doomen (2023) warns that one of the biggest obstacles in the standardization of AI is precisely the lack of a clear legal classification of these systems - whether they are just tools in the hands of humans or actors that can function independently and make decisions.

The modern world is heavily reliant on artificial intelligence - and in everyday, almost imperceptible interactions. AI is used in navigation systems, personalized recommendations on digital platforms, applications for voice and image recognition, virtual assistants, translators, disease diagnostics, all the way to system monitoring and public policy management. According to Stanford HAI data from 2024, more than 60% of the global population uses at least one AI-functionality on a daily basis, often unknowingly (Stanford HAI, 2024). Biometric systems for facial recognition, algorithmic selection of job candidates, automatic profiling of users and moderation

of content on social networks, represent a direct threat to basic rights when adequate regulatory supervision is absent. This problem is particularly pronounced when algorithmic decisions are used in the fields of health, education, policing or justice.
However, on the other hand, it must be pointed out that artificial intelligence significantly contributes to the improvement of efficiency, precision and accessibility in many areas. The integration of artificial intelligence and large databases can significantly improve the healthcare system through personalized medicine, more accurate disease detection and more efficient process management. Such technological advances enable faster diagnosis, better treatment outcomes, and reduced administrative costs with an improved patient experience (Vadisetty 2025, pp. 9). In addition, when it comes to the judicial system, AI can greatly influence faster court turnarounds through systematization of cases, personalization of education and improvement of traffic safety. However, despite these benefits, it is precisely the breadth of its application that makes AI a potential threat when it escapes human control. The key point of danger arises when the human no longer controls the technology, but the technology controls the human - either through the shaping of behavior through personalized information, or through decisions made without a clear mechanism of human supervision.
At the moment when machine decision-making takes place without the possibility of intervention, the individual loses autonomy over his own status, which opens the door to serious rights violations. The Future of Life Institute (2023) specifically warns of the risk of competent but uncontrolled AI systems, which, without malicious intent, may produce results contrary to the basic values of human dignity and freedom.
Bearing in mind the mentioned challenges, the authors of this paper direct their attention to the human rights dimensions of the development and application of artificial intelligence. A special focus will be placed on the right to privacy and data protection, the right to non-discrimination and equality, the right to a fair trial, as well as the right to freedom of expression and opinion, as the most sensitive rights that are most often questioned in the modern context due to the uncritical and unregulated use of AI technologies.
In the following subsections, each of the mentioned areas will be analyzed in detail, through the prism of theoretical interpretations, legal acts and examples from practice.

INTERNATIONAL NORMATIVE FRAMEWORK REGULATING ARTIFICIAL INTELLIGENCE IN THE CONTEXT OF HUMAN RIGHTS PROTECTION

The development and application of artificial intelligence (AI) have posed numerous challenges to the international community, and the establishment of a clear normative framework is crucial for the protection of human rights and fundamental freedoms. Although existing international human rights treaties are not specifically written with a focus on AI, their universal application provides a basis for regulating the impact of AI on the rights of individuals.

The international legal framework includes key instruments such as the Universal Declaration of Human Rights (1948), the International Covenant on Civil and Political Rights (1966), and the UN Charter of the Rights of the Child (1989), which protect the rights to privacy, non-discrimination, freedom of expression, and a fair trial—the rights most threatened by the application of AI technologies (UN, 1948; UN, 1966; UN, 1989).

In addition to these universal instruments, the initiatives and guidelines of international bodies, such as the UN Working Group on Artificial Intelligence (UN Human Rights Council, 2021), which promote the ethical use of AI and respect for human rights, are also important.

Although the international framework is still in the development phase and mostly relies on non-binding recommendations (soft law), it sets important guidelines for states and international actors in creating policies and regulations on artificial intelligence. The European Union takes a leading role in the normative shaping of the field of artificial intelligence, striving to establish a balance between encouraging innovation and preserving basic human rights. The most significant step in this direction is the adoption of the Regulation on Artificial Intelligence (AI Act, 2024), the first comprehensive legal act in the world dedicated to this area. The AI Act introduces a risk classification system based on proportionality, whereby low-risk applications are subject to more lenient requirements, while high-risk systems are subject to strict obligations regarding transparency, security and human oversight. This approach emphasizes that technologies that can affect fundamental rights, such as algorithms in justice, employment or health care, must be subject to the highest standards of control (Smuha, 2021).

In addition to the AI Act, the General Data Protection Regulation (GDPR, 2016/679), which guarantees EU citizens the protection of personal data, the right to privacy and transparency during automated decision-making, continues to play a key role. The GDPR in Article 22 foresees the right of an individual not to be the subject of a decision based solely on automated processing, including profiling, if that decision has legal consequences or similarly significantly affects him (Wachter, Mittelstadt & Floridi, 2017). This laid the foundation for the concept of the "right to explanation", which specifically refers to algorithmic decisions and represents an important guarantee of the protection of human rights in the digital environment.

Also, the OECD Principles for Trusted AI, adopted by many member states, define key values such as fairness, transparency and accountability, which is the basis for numerous national and regional regulations (OECD, 2019).

The European Court of Human Rights (ECtHR) plays a crucial role in the formation of protection standards, which through practice interprets and adapts the European Convention on Human Rights to the challenges of modern technologies. In the case of Big Brother Watch and Others v. the United Kingdom (2021), the Court took the position that mass collection and automated analysis of data without adequate guarantees constitutes a violation of the right to privacy under Article 8 of the Convention. This practice clearly shows that the court is ready to develop protection standards in the context of digital technologies, including AI. This confirms that regulation in the EU is not only based on legislative initiatives, but also on the

evolving interpretation of human rights through judicial practice (Leenes et al., 2018). In this way, the European Union is shaping an integrated approach that combines normative regulation (AI Act), horizontal data protection framework (GDPR) and judicial protection mechanisms (ECHR), creating a comprehensive system that aims to make artificial intelligence safe, transparent and compatible with basic human rights.

The United States approaches the regulation of artificial intelligence in a fragmented manner, following a federal system in which responsibilities are divided between the federal government and the states. Unlike the European Union, there is no single law that directly regulates AI, but the legal framework is built through a combination of guidelines, regulatory documents and initiatives at the level of federal agencies. The key document in this context is the AI Bill of Rights, published by the White House in 2023, which defines the basic principles for the development and application of AI technologies in accordance with human rights and ethical standards, including transparency, fairness, accountability and privacy protection (White House, 2023, Blueprint for an AI Bill of Rights).

In addition to the AI Bill of Rights, the NIST AI Risk Management Framework, which provides guidelines for identifying, assessing and mitigating risks in the implementation of AI systems, also plays a significant role. This framework functions as a voluntary standard, but it influences the practices of companies and regulatory bodies, directing them towards alignment with the principles of fairness, non-discrimination and accountability (National Institute of Standards and Technology, 2023).

The fragmentation of the regulatory approach is further complicated by the problems of federalism and differences in competences. At the federal level, various agencies, such as the Federal Trade Commission (FTC) and the Food and Drug Administration (FDA), regulate specific applications of AI, but there is no unified coordination that spans all sectors. This dispersion of rules can lead to uneven protection of human rights, especially when it comes to discrimination, privacy and algorithmic surveillance (Calo, 2021).

Because of this fragmented approach, the implementation of AI in the US often depends on voluntary self-regulation and industry standards, and the responsibility for protecting the rights of citizens is partially left to the market. This raises questions about the efficiency and adequacy of human rights protection, especially compared to the coherent European model, where AI systems are subject to clear ex ante regulatory and judicial standards (Smuha, 2021).

Critically observed, the European approach offers stronger guarantees of human rights protection, but some authors evaluate it as potentially restrictive for innovation and technological competitiveness. The American model, on the other hand, enables faster development of the AI industry, but at the cost of a lower degree of legal certainty and greater exposure of citizens to discrimination, non-transparent profiling and commercial surveillance. A comparative analysis shows that neither model is without flaws: the EU offers robust protection but risks regulatory overload, while the US offers innovation with significantly greater risks to human rights. The optimal solution probably lies in their combination — in the

balance between technological dynamics and strict guarantees of individual protection.

RIGHT TO PRIVACY AND DATA PROTECTION

The right to privacy and data protection is one of the fundamental human rights that is seriously challenged by the development and application of artificial intelligence (AI). AI systems operate on the basis of processing large amounts of personal, sensitive, biometric and behavioral data to create algorithmic models for learning and predicting user behavior. Such processing often takes place without sufficient informed consent and transparency, which opens numerous legal and ethical dilemmas.

In today's digital society, individuals are exposed to constant surveillance through platforms, applications and sensor systems, and AI enables the analysis and connection of data in ways that were not possible before. It violates the autonomy of the individual to manage his own information and directly conflicts with the basic principles of personal freedom and dignity. Although there are regulations that protect the right to privacy, such as the General Data Protection Regulation (GDPR) — the General Data Protection Regulation of the European Union, adopted in 2016 and applied since 2018 — which aims to ensure the protection of personal data and enable people to have greater control over their own information, practice shows that these principles are often not applied consistently in the context of complex AI systems. Among other things, the GDPR stipulates the principles of "privacy by design" and "privacy by default", as well as the right to explanation and the right to object to automated decisions.

Special concern is caused by the occurrence of the so-called "predictive privacy," where algorithms draw conclusions about users from data they did not directly provide, leading to profiling that users often have no knowledge of or control over. Sectors such as healthcare, education and employment are particularly at risk, where AI systems process sensitive data, and patients and users do not have access to information about how to make decisions and control over their own data. The mass use of biometric data in public surveillance further threatens privacy, as many citizens are not even informed about the collection of such data

The verdict Glukhin v. Russia (2023) represents one of the first decisions of the ECtHR directly related to the misuse of modern surveillance technologies, including facial recognition systems. In this case, the Russian authorities identified and sanctioned the applicant on the basis of his photograph taken by public transport cameras, using algorithmic analysis of biometric data. The court found that this type of surveillance represents a serious interference with the right to privacy from Article 8 of the ECHR, especially because it is a technology that enables mass, unlimited and indiscriminate monitoring of citizens without their knowledge. The ECtHR emphasized that biometric surveillance, due to its ability to identify and monitor individuals in real time, represents a much more invasive form of control than traditional video surveillance and as such requires a clear, precise and predictable legal framework, which did not exist in Russia. The court also indicated

that such surveillance has a "chilling effect" on freedom of expression and participation in protests, which further increases its harmfulness in a democratic society. Since the authorities did not provide adequate guarantees, the surveillance was assessed as disproportionate and therefore inconsistent with the Convention, with which the ECtHR concluded that there had been a violation of Article 8.

Unlike a relationship with a lawyer, doctor or therapist, where legal institutes ensure the confidentiality of communication, conversations with AI systems are not protected by any legal privilege. This was also pointed out by the executive director of OpenAI, Sam Altman, emphasizing that everything the user shares with ChatGPT can be used as evidence in court proceedings, including records of deleted conversations that the company is legally obliged to keep (TRT Balkan, 2025). Such practice shows the existence of a serious normative vacuum in the field of privacy protection in the digital environment.

It is particularly problematic that users are often unaware that the information they exchange with AI systems does not enjoy the same degree of protection as communication with experts. This creates a situation where sensitive data can be collected and used in legal proceedings in a way that would be unimaginable when it comes to medical or legal secrets. Such a situation undoubtedly calls into question the realization of the right to privacy guaranteed by international documents such as Article 8 of the European Convention on Human Rights (ECHR, 1950), Articles 7 and 8 of the Charter of Fundamental Rights of the European Union (2000) and provisions of the General Data Protection Regulation (GDPR, 2016/679).

The practice of the European Court of Human Rights additionally confirms that digital communication falls under the protection of the right to privacy. In the Barbulescu v. Romania case, the Court pointed out that even in a professional environment there is a reasonable expectation of privacy of communication (ECtHR, 2017). It follows that the user's interaction with AI systems would have to be covered by adequate guarantees. Altman's warning therefore indicates an urgent need for the adoption of new laws that would explicitly protect the privacy of conversations with AI systems and regulate the way that data is processed and stored (TRT Balkan, 2025).

The modern development of artificial intelligence shows that the traditional concept of privacy is no longer sustainable in its classical form, because technological systems have the capacity to reconstruct personal profiles even from minimal, fragmentary data. We are of the opinion that the right to privacy is entering a phase of structural vulnerability: algorithmic inferences, mass surveillance, non-transparent decision-making models and the collection of data beyond the user's knowledge seriously limit the real autonomy of the individual. A special problem is the fact that most of the existing legal regimes start from the assumption of informed consent, although in the real digital environment such consent is often illusory, reduced to a formality without a real understanding of how data is processed. In doing so, users do not have the ability to control the algorithmic processes or influence the way the models draw conclusions about their identity, habits and potential behaviors. We believe that such a situation opens up space for the erosion of dignity and the instrumentalization of the individual as an object of technological processing.

In this sense, it is necessary to redefine the right to privacy so that it goes from passive data protection to an active model of "digital self-determination". This would require strengthening the right to explain algorithmic decisions, introducing the obligation of technical transparency and the possibility of independent supervision over systems that monitor, analyze or profile users. We believe that legislators should introduce an explicit ban on the use of data from communication with AI systems in court proceedings, except with strict procedural guarantees and clear conditions of proportionality. Additionally, we propose the establishment of a separate legal institute—the digital communications privilege—that would protect users' conversations with AI tools in a manner comparable to medical or legal privilege. Without such reforms, we risk the relationship between man and technology becoming one of one-sided surveillance in which the individual has no viable possibility of resistance or control over his own identity data.

From the perspective of civil law protection, violations of privacy caused by artificial intelligence systems open the possibility for individual claims aimed at compensating both material and non-material damage. Unlike administrative or regulatory mechanisms, civil law remedies enable the injured party to directly seek redress for the infringement of personality rights, including the right to privacy, data autonomy, and personal dignity. The large-scale collection, profiling, and predictive use of personal and biometric data may constitute an unlawful interference with personality rights, giving rise to claims for damages, injunctions, and the cessation of unlawful data processing. In this context, civil liability plays a corrective and preventive role, ensuring that abstract data protection principles are translated into effective and individualized protection. Without access to civil law remedies, the protection of privacy risks remaining largely declaratory, particularly in cases where supervisory authorities are unable to respond promptly or comprehensively to complex AI-driven data practices.

THE RIGHT TO NON-DISCRIMINATION AND EQUALITY

The right to non-discrimination and equality is one of the key pillars of modern legal systems and the basis of human rights protection. In the context of artificial intelligence, this principle gains particular weight, because AI systems, although they are presented as objective and neutral, often reproduce or even reinforce existing social inequalities and prejudices. This is particularly evident in algorithmic decisions that are made on the basis of data that are not always representative or are biased due to human errors in data collection and interpretation (Barocas & Selbst, 2016).

One of the most frequently reported negative impacts of artificial intelligence on human rights is the impact on the prohibition of discrimination, that is, on the right to equal treatment. Algorithmic discrimination can occur in various sectors, as AI systems often reproduce or reinforce existing social inequalities. In the public sphere, this includes areas such as the judiciary, social benefits, the pension system or the assessment of entitlements to various forms of assistance, where automated decisions can lead to unfair outcomes for certain groups. In the private sector,

discriminatory effects appear in the processes of employment, loan approval, housing or targeted advertising, where algorithms often use patterns from historical data that contain biases (Prlja, Gasmi, Korać 2023, pp. 71).

The EU General Data Protection Regulation (GDPR) in point 71 of the Preamble and Article 22 clearly regulates automated decision-making, giving the individual the right not to be subjected to decisions based solely on automated processing that produce legal or significant consequences, such as online loan rejection or algorithmic selection in employment. This type of processing also includes profiling, i.e. the assessment of an individual's personal characteristics, but it is allowed only in special cases — when it is provided for by EU law or national law, when it is necessary to prevent fraud and tax evasion, for the execution of a contract, or when explicit consent has been given. Although the GDPR represents a key mechanism for preventing discrimination when processing personal data, its application does not include all forms of algorithmic decision-making, especially predictive models that do not allow identification of individuals. Additionally, the strict regime for processing particularly sensitive data (e.g. on health or racial origin) makes it difficult for organizations to collect the information necessary to detect and prove algorithmic discrimination, which narrows the possibility of its effective monitoring and sanctioning (Borgtesius 2018, pp. 44-45).

Although existing laws on non-discrimination, data protection and consumer protection provide certain instruments to combat such phenomena, they are not sufficiently aligned with the specifics and complexity of modern AI systems. This is why further normative improvement is necessary - both through amendments to existing regulations and through the adoption of new, general and sector-specific rules that will precisely regulate the development and application of automated decision-making. The goal of such reforms must be to create a system framework that enables the safe use of AI technologies and ensures a high degree of protection of individual rights and freedoms.

In conclusion, the analysis of discrimination in artificial intelligence systems shows that algorithmic biases are deeply rooted in the way models are developed, trained and implemented, and that they represent a serious challenge to the principles of equality and fair treatment. The theoretical framework indicates that discrimination occurs as a consequence of inadequate data sets, opaque learning processes and models that reproduce social stereotypes instead of neutralizing them. The findings of Mehrabi's research (2019) further confirm these assumptions, showing that algorithms, even when not using sensitive data, can indirectly reconstruct racial, gender, or socioeconomic characteristics through so-called proxy variables, making the bias structural rather than incidental. This research demonstrates that AI systems can favor or discriminate against entire groups through patterns that are hidden and hard to see by the human observer, thereby compromising the essence of equal treatment. Therefore, we believe that for effective protection against algorithmic discrimination, it is necessary to establish stricter oversight mechanisms, technical standards for bias evaluation, and clear procedural guarantees that ensure that AI systems are transparent, verifiable, and accountable. Only with a combination of robust legal frameworks and scientifically based methodologies is it possible to

prevent artificial intelligence from becoming a source of new inequalities and violations of basic human rights.

In addition to public law anti-discrimination frameworks, civil law mechanisms represent a crucial avenue for addressing algorithmic discrimination. Individuals affected by biased AI systems may seek compensation for discriminatory treatment through civil liability claims, particularly where automated decision-making results in unequal access to employment, credit, housing, or public services. Civil proceedings allow courts to assess discriminatory outcomes not only through intent-based standards, but also through the effects of algorithmic decision-making, which is especially important in cases involving indirect or structural discrimination. Moreover, civil law remedies enable courts to impose injunctive relief, requiring the modification or suspension of discriminatory systems. In the absence of effective civil law enforcement, algorithmic discrimination risks becoming normalized, as affected individuals are deprived of practical tools to challenge opaque and data-driven forms of unequal treatment.

RIGHT TO A FAIR TRIAL

The application of artificial intelligence in judicial systems opens up complex issues regarding the exercise of the right to a fair trial, guaranteed by Art. 6 of the European Convention on Human Rights and Art. 14 of the International Covenant on Civil and Political Rights. The key risk highlighted by the authors is the fact that AI systems often function as "black boxes", which makes it difficult to access the reasons for the decision and violates the principles of transparency and adversarial procedure (Molbæk-Steensig, 2023, pp. 11-12). If the party cannot understand on the basis of which data and samples the algorithm drew a conclusion that affects its procedural rights, it is impossible to ensure an effective possibility of contesting the decision, which directly affects the right to defense and the right to a legal remedy.

Another significant problem relates to the discriminatory outcomes of algorithmic decision-making. Matić Bošković shows that AI in justice can disproportionately negatively affect certain social groups due to historically biased datasets, especially in the areas of predictive risk, custody decision-making, or testimony credibility assessment (2024, pp. 482–486). These systems often reproduce structural social inequalities, but due to the technical framework of decision-making, they make them less visible and more difficult to challenge, thus opening up the possibility for systemic inequality in access to justice, which is contrary to the principles of equality of arms and prohibition of discrimination in the procedure.

The latest research indicates that the danger lies not only in the use of AI tools, but also in the change in the epistemology of the judiciary — judicial decisions depend more and more on statistical predictions, and less and less on individual consideration of the specific circumstances of the case and judicial discretion (Mizaras et al., 2024, pp. 5–12). The authors warn that algorithmic recommendations, even when not formally binding, have a strong "automated suggestion" effect, as judges may feel professional pressure to rely on the "objectivity" of the technological system. This undermines the independence and

impartiality of the court, especially when the algorithm provides a risk score, estimates the likelihood of recidivism or suggests sentencing policy.

One of the most relevant examples of comparative judicial practice that indicates the risks of applying artificial intelligence in the context of the right to a fair trial is the case of State v. Loomis (Wisconsin Supreme Court, 2016). In this case, the accused challenged the use of the COMPAS algorithmic tool in the sentencing phase, stating that his right to a fair procedure was violated because neither the parties nor the court had insight into the way the algorithm works, nor could they review the criteria based on which the risk of re-committing a criminal offense is assessed. The court allowed the use of COMPAS, but at the same time introduced a number of restrictions: the algorithm must not be decisive or the only basis for sentencing, its assessment must be used only as a secondary factor, and judges must be aware of the methodological limitations and possible biases of this type of software. This kind of reasoning is an implicit recognition that non-transparent algorithmic systems can threaten the principles of transparency, adversariality and the right to a reasoned decision — fundamental guarantees of a fair trial. (State v. Loomis, 881 N.W.2d 749, Wis. Sup. Ct., 2016)

From the point of view of the protection of human rights, we believe that the key challenge is that AI introduces a new form of power asymmetry in the judicial procedure: technology becomes an actor that influences decisions, but without legal subjectivity, responsibility and the obligation of reasoning. We believe that this leads to the erosion of the fundamental guarantees of a fair trial — the right of access to a court, the right to a reasoned decision, the right to equality of arms, and the right to an effective remedy. Therefore, we propose to introduce minimum standards in all justice systems: mandatory explainability of algorithmic decisions, the possibility of completely turning off AI in all key moments of the procedure, a special monitoring mechanism to prevent discrimination in algorithmic models and a strict ban on the use of AI systems in deciding on criminal liability, punishment and deprivation of liberty. Only with such guarantees can AI have the function of assisting the court and not replacing it, while protecting the essential values of a fair trial and human dignity.

From a civil law standpoint, the use of artificial intelligence in judicial and quasi-judicial proceedings raises significant questions of liability for harm caused by defective or opaque algorithmic tools. Where automated systems influence judicial outcomes or procedural decisions, individuals may suffer tangible and intangible damage resulting from violations of procedural guarantees, such as the right to a reasoned decision or equality of arms. Civil law offers a framework for addressing such harm through claims against the state, judicial authorities, or private developers of AI systems, depending on the specific allocation of responsibility. The possibility of civil liability serves not only as a compensatory mechanism, but also as a deterrent against the uncritical deployment of AI in judicial contexts. Without civil law accountability, the risks associated with automated decision-making remain insufficiently internalized by institutions and technology providers.

THE RIGHT TO FREEDOM OF EXPRESSION AND OPINION

The development and application of artificial intelligence profoundly transforms the realization of the right to freedom of expression and opinion, both in its "active" dimension (the right to speak) and in its "passive" dimension (the right to receive information). De Gregorio and Dunn point out that AI systems are reshaping the public sphere by affecting the visibility, availability and ranking of information, thus indirectly determining which ideas will form and circulate in the public (De Gregorio & Dunn 2023, pp. 2–4). Automated content moderation systems also carry risks: they can increase the effectiveness of removing illegal content, but also lead to "over-removal" of legitimate speech due to detection errors, which directly threatens freedom of expression (De Gregorio & Dunn, 2023, pp. 3-5).

Llansó and Van Hoboken show that algorithmic moderation, when non-transparent and insufficiently explained, turns into a form of "private regulation of the public sphere", by which technological platforms effectively determine the limits of permissible speech without clear guarantees for users (Llansó & Van Hoboken, 2020, pp. 7–10). This problem is further deepened by the personalization and optimization of content, where recommender algorithms influence the formation of opinions, creating the so-called "filter bubbles" and "echo-chambers" that limit the pluralism of information (Llansó & Van Hoboken, 2020, pp. 13–15).

A special challenge is represented by generative models, which, as Gullo (2024) shows, can simultaneously serve as a tool of creative expression and a mechanism of disinformation. Gullo indicates that models capable of producing false, effectively "realistic" information (deepfakes) can undermine trust in public discourse and threaten the public's right to receive accurate information (Gullo, 2024, pp. 5–7). Reese adds that generative AI introduces a new kind of epistemological uncertainty: users can no longer reliably distinguish true from fabricated content, which undermines the very function of free expression in a democratic society.

At the level of human rights, the common warning of all authors is the normative gap: most of the restrictions that AI introduces to the right to freedom of expression arise from the lack of transparency of algorithms, the absence of an effective appeal procedure and the dominance of private actors who shape the public sphere according to business rather than democratic criteria. We believe that the right to freedom of expression in an AI environment can only be protected by a combination of mandatory transparency of algorithms, precise standards for automated moderation, the right of users to human review of content removal decisions, and strict regulation of generative models regarding the labeling of synthetic content. Such steps do not limit the development of the technology, but ensure that AI systems are used as instruments to enhance, not curtail, freedom of expression.

Civil law protection plays a particularly important role in safeguarding freedom of expression in environments shaped by artificial intelligence. Automated content moderation, algorithmic ranking, and generative models can cause reputational harm, unjustified suppression of lawful speech, or the dissemination of false and

damaging information. In such cases, individuals may rely on civil law remedies to protect their personality rights, including claims for the removal of unlawful content, compensation for non-material damage, and preventive injunctions. Civil courts are uniquely positioned to balance competing interests, such as freedom of expression, protection of reputation, and commercial interests of platforms, on a case-by-case basis. The availability of civil law remedies is therefore essential to prevent private technological actors from becoming de facto arbiters of permissible speech without accountability.

CONCLUSION

The conducted analysis shows that the development and application of artificial intelligence is not a technical or exclusively regulatory issue, but a deep human rights issue that touches the very essence of the modern democratic order. The paper identified four key areas of risk - surveillance and endangering the right to privacy, algorithmic bias and violations of the principle of equality, the application of AI in the judiciary and the impact on the right to a fair trial, as well as restrictions on freedom of expression through automated moderation and generative models. In all these areas, the common denominator is the fact that artificial intelligence acts as a multiplier of existing social inequalities and normative deficiencies, turning potential individual injuries into a structural problem.

In the part of the paper devoted to the right to privacy, it was pointed out that mass surveillance, predictive analytics and extensive processing of personal and biometric data lead to the classical concept of informed consent becoming factually illusory. In practice, users do not have real control over their own data, nor do they have insight into how conclusions about them are drawn and further used. Similarly, the analysis of the right to non-discrimination shows that algorithmic decisions, based on inadequate or historically burdened data sets, lead to the reproduction and deepening of social inequalities, whereby bias is no longer an exception, but part of the very logic of the system's functioning. This undermines the essence of equal treatment, as individuals and groups are discriminated against based on patterns that are hard to see and even harder to prove.

When it comes to the right to a fair trial, the paper pointed out the danger of gradually shifting the focus from individual consideration of the case to statistical predictions of risk and behavior. In such a framework, the risk increases that judges, prosecuting authorities or other actors of the procedure give too much weight to algorithmic recommendations, which can lead to the erosion of the right to a reasoned decision, equality of arms and an effective legal remedy. In the sphere of freedom of expression and opinion, it was especially emphasized that automated moderation of content, personalization of information and the appearance of synthetic, believable, but potentially misleading content are changing the very structure of the public sphere. This not only threatens the individual rights of speech and access to information, but also the collective conditions for an informed democratic discussion.

A comparative presentation of the European and American models of artificial intelligence regulation additionally confirms the initial hypothesis that the existing

normative frameworks are not sufficiently aligned with the scope and nature of these risks. The European model, based on the AI Act, the GDPR and the practice of the European Court of Human Rights, offers a more coherent and human rights-oriented approach, but faces challenges of enforcement, technical feasibility and a potential slowdown in innovation. The American model, based on fragmented regulation, voluntary standards and a strong role of the market, enables rapid technological expansion, but at the cost of insufficiently uniform protection of rights and reliance on self-regulation by actors who are also carriers of commercial interests. This dichotomy clearly shows that neither approach is without flaws and that a deeper convergence is needed between the demand for innovation and the obligation to protect human rights.

Based on everything presented, we are of the opinion that for the effective protection of human rights in the age of artificial intelligence, it is necessary to move from partial and sectoral interventions to a systemic, risk-driven approach. Such an approach must include strengthening the right to explanation and objection, technical and institutional transparency, independent oversight of the development and implementation of AI systems, as well as clear prohibitions in areas where the risk to an individual's dignity, freedom and physical integrity is unacceptably high. We consider the introduction of new institutes, such as the privilege of digital communication and standardized procedures for the assessment of algorithmic bias, to be particularly significant, which would reflect the specifics of the digital environment.

From a civil law perspective, the analysis demonstrates that private law mechanisms play a central role in ensuring effective human rights protection in the context of artificial intelligence. Civil liability, personality rights protection, and injunctive relief provide individuals with concrete and individualized remedies that go beyond abstract regulatory compliance. Unlike public law frameworks, which primarily operate through ex ante standards and institutional supervision, civil law responds ex post to actual harm, enabling courts to assess responsibility, causation, and proportionality in specific cases. In this sense, civil law functions as both a corrective and preventive instrument, internalizing the social costs of algorithmic risks and compelling technology developers, deployers, and public authorities to account for the human rights impacts of AI systems. Without a coherent and adaptable civil law framework, the enforcement of human rights in the digital environment remains fragmented and insufficiently responsive to the realities of automated decision-making.

In conclusion, artificial intelligence is not inherently a threat to human rights, but it will not become compatible with them by default. In order for its development to remain aligned with the fundamental values of the modern legal order, the law must respond promptly and decisively to the challenges of surveillance, algorithmic bias, automated decision-making in judicial contexts, and restrictions on freedom of expression. While public law regulation, constitutional guarantees, and administrative oversight are indispensable, they are insufficient to address individualized and concrete harm caused by AI-driven systems. Effective protection therefore requires a robust framework of civil law remedies, including liability for

damages, injunctive relief, and the protection of personality rights, capable of translating abstract human rights standards into enforceable individual claims. Without such mechanisms, there is a real risk that technological progress will erode decades of human rights protection, transforming individuals from subjects of rights into objects of continuous algorithmic processing and management, and undermining human dignity, legal certainty, and trust in the digital legal order.

REFERENCES

1. Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104(3), 671–732.
2. Borgesius, F. Z. (2018). Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe Study, 1–58.
3. Calo, R. (2021). Artificial intelligence policy: A primer and roadmap. UC Davis Law Review, 55(2), 899–942.
4. De Gregorio, G., & Dunn, P. (2023). Artificial Intelligence and Freedom of Expression. Artificial Intelligence and Human Rights (pp. 76–90). Oxford University Press.
5. Doomen, J. (2023). The artificial intelligence entity as a legal person. Information & Communications Technology Law, 15(2), 112–134.
6. Gullo, S. (2024). Deepfakes and the erosion of informational integrity. Journal of Media and Interdisciplinary Studies, 16(1), 1–18.
7. Leenes, R., Bayamlıoğlu, E. (2018). The 'rule of law' implications of data-driven decision-making: a techno-regulatory perspective. Law Innovation and Technology 10(2):1-19
8. Llansó, E., Van Hoboken, J., Leerssen, P., Harambam, J. (2020). Artificial Intelligence, Content Moderation, and Freedom of Expression. Information Society, 36(1), 1–18.
9. Matić Bošković, M. (2024). Implications of EU AI regulation for criminal justice. Regional Law Review, 49(3), 470–490.
10. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2019). A survey on bias and fairness in machine learning. ACM Computing Surveys, 52(6), 1–35.
11. Nikiforiadis, A., Tsavdari, D., Mizaras, V., & Ayfantopoulou, G. (2023). Identifying barriers and expectations in MaaS: Users' and stakeholders' perspective. Future Transportation, 3(4), 1240–1252.
12. Molbæk-Steensig, H., & Quemy, A. (2023). Artificial intelligence and fair trial rights. Artificial Intelligence and Human Rights, 265–280.
13. Prlja, D., Gasmi, V., & Korać, M. (2023). Algoritamska dikriminacija. Uporednopravni izazovi u savremenom pravu (10), 59–73.
14. Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. Journal of Responsible Technology
15. Smuha, N. A. (2021). Beyond the individual: governing AI's societal harm. Internet Policy Review, 10(3).
16. TRT Balkan. (2025, July 31). Disclosure: Your secrets are not safe with ChatGPT. TRT World.

17. Vadisetty, R., Polamarasetti, A., Guntupalli, R., Rongali, S. K., Raghunath, V., Jyothi, V. K., & Kudithipudi, K. (2021). Legal and ethical considerations for hosting GenAI on the cloud. International Journal of AI, BigData, Computational and Management Studies, 2(2), 28–34.
18. Stanford HAI. (2024). AI Index Report 2024. Stanford University.

LEGAL ACTS, REGULATIONS AND OFFICIAL DOCUMENTS

19. European Union. (2000). Charter of Fundamental Rights of the European Union.
20. European Union. (2016). General Data Protection Regulation (GDPR) (EU) 2016/679.
21. European Union. (2024). Artificial Intelligence Act (AI Act).
22. European Court of Human Rights. (2017). Barbulescu v. Romania.
23. European Court of Human Rights. (2021). Big Brother Watch and Others v. the United Kingdom.
24. United Nations. (1948). Universal Declaration of Human Rights.
25. United Nations. (1966). International Covenant on Civil and Political Rights.
26. United Nations. (1989). Convention on the Rights of the Child.
27. UN Human Rights Council. (2021). Working Group on Artificial Intelligence – Ethical Guidelines.
28. White House – Office of Science and Technology Policy. (2023). Blueprint for an AI Bill of Rights.
29. National Institute of Standards and Technology. (2023). AI Risk Management Framework.
30. OECD. (2019). OECD Principles on Artificial Intelligence.
31. TRT Balkan. (2025). Izveštaj o privatnosti i ChatGPT komunikaciji.

CASE LAW

32. European Court of Human Rights (ECtHR). (2017). Barbulescu v. Romania, Application no. 61496/08.
33. European Court of Human Rights (ECtHR). (2021). Big Brother Watch and Others v. the United Kingdom, Applications nos. 58170/13, 62322/14 and 24960/15.
34. European Court of Human Rights
35. Zakharov v. Russia, Application no. 47143/06, Judgment of 4 December 2015.
36. Supreme Court of Wisconsin
37. State v. Loomis, 881 N.W.2d 749, Supreme Court of Wisconsin, Decision of July 13, 2016.

REZIME

Ovaj rad pruža sistematsku analizu ključnih rizika koje savremeni sistemi veštačke inteligencije proizvode po ostvarivanje i zaštitu ljudskih prava. Polazeći od činjenice da se AI sve intenzivnije integriše u javne politike, pravosuđe, tržišne procese i upravljanje informacijama, rad identifikuje četiri najkritičnije oblasti narušavanja

prava: digitalni nadzor, algoritamsku pristrasnost, automatizovano donošenje odluka u pravosudnim postupcima i ograničavanje slobode izražavanja. U prvom delu analiziraju se mehanizmi masovnog prikupljanja i obrade podataka, uključujući biometrijske tehnologije i prediktivne modele, koji dovode do erozije prava na privatnost i stvaranja strukturalnih asimetrija moći. Zatim se razmatra reprodukcija društvenih nejednakosti kroz pristrasne modele mašinskog učenja, sa posebnim osvrtom na diskriminatorne ishode u zapošljavanju, kreditnom bodovanju i javnom odlučivanju. Posebna pažnja posvećena je primeni AI u pravosuđu, gde netransparentnost algoritama, nedostatak objašnjivosti i rizik od automatizovanog uticaja na sudske odluke ozbiljno dovode u pitanje garancije pravičnog suđenja. U poslednjem delu analiziraju se uticaji automatizovane moderacije sadržaja, personalizovanog rangiranja informacija i generativnih modela na slobodu izražavanja i integritet javnog diskursa. Rad kombinuje normativno-dogmatsku, komparativnu i analitičko-sintetičku metodologiju, pri čemu upoređuje evropski i američki pristup regulaciji veštačke inteligencije. Evropski model, zasnovan na ex ante zaštiti prava i strogim standardima transparentnosti, suprotstavljen je fragmentiranom i tržišno orijentisanom američkom pristupu, što ukazuje na globalnu potrebu za koherentnijim regulatornim okvirom. Zaključak rada ističe da su ovi rizici strukturne prirode i da efikasna zaštita ljudskih prava zahteva kombinaciju tehničke objašnjivosti, demokratskog nadzora, jačanja procesnih garancija i međunarodne harmonizacije pravila. Samo takav pristup može obezbediti da razvoj veštačke inteligencije ostane u granicama koje poštuju dostojanstvo, autonomiju i osnovna prava pojedinca.