АКТУАЛЬНІ ДОСЛІДЖЕННЯ МОЛОДИХ НАУКОВЦІВ

UDC: 341.231.14:004.8:342.7
DOI: https://doi.org/10.31861/jiel.2026.1.53-64

Anastasiia Labyk,
PhD candidate,
Assistant Professor of the Department
Of International Law and Comparative Law,
Yuriy Fedkovych Chernivtsi
National University, Ukraine
e-mail: a.labyk@chnu.edu.ua
ORCID: https://orcid.org/0009-0005-1376-6771

ALGORITHMIC DISCRIMINATION AS A THREAT TO DIGITAL HUMAN RIGHTS: A THEORETICAL AND LEGAL DIMENSION

The article investigates the issues regarding the protection of digital human rights in the context of the rapid implementation of artificial intelligence systems and automated decision-making within the public sector. Particular emphasis is placed on the emergence of a new phenomenon — algorithmic discrimination, which poses a serious threat to the principle of equality and creates unprecedented challenges for traditional anti-discrimination mechanisms. Algorithmic discrimination is defined as a latent, unintentional form of unfair or biased treatment towards an individual or a group of individuals possessing one or more protected characteristics, which is caused by the functioning of automated decision-making systems and lies in the large-scale and self-reinforcing restriction of their rights as a result of processing historically biased source data or applying neutral proxy variables.

The article identifies and analyses the key characteristics of this phenomenon: the latent nature and insufficient transparency of the operation of algorithms, which hinder the detection of biased treatment until the violations acquire a systematic character; the absence of intent to restrict the rights of a specific individual or group of individuals, as well as the predominantly indirect nature of algorithmic discrimination; the use of proxy variables, which have a neutral character but, due to existing historical biases in society, can lead to unequal treatment; intersectionality or its cross-cutting nature, caused by the combination of several protected characteristics; a significantly accelerated and large-scale character compared to the traditional understanding of discrimination, caused by the automation of decision-making, as well as the "feedback loop", whereby the operation of algorithms further amplifies the existing unfair treatment of certain protected groups within society.

The article demonstrates that the opacity of algorithms precludes the effective application of the classical standards for the allocation of the burden of proof under Article 14 of the ECHR, as it deprives the applicant of the tools necessary to establish a causal link between the source data, algorithmic processes, and discriminatory treatment. The study also substantiates the necessity of expanding the human rights protection doctrine and adapting the approaches of the ECtHR, incorporating the concept of "explainable artificial intelligence", in order to guarantee equitable protection against algorithmic bias.

Keywords: *human rights, digital rights, artificial intelligence, digitalisation of international law, European Court of Human Rights, algorithmic discrimination.*

ISSN 3083-7979

Лабик Анастасія. Алгоритмічна дискримінація як загроза цифровим правам людини: теоретико-правовий вимір.

У статті досліджується проблематика захисту цифрових прав людини в умовах стрімкого впровадження систем штучного інтелекту та автоматизованого прийняття рішень у публічний сектор. Акцентовано увагу на виникненні нового явища — алгоритмічної дискримінації, що становить серйозну загрозу принципу рівності та створює безпрецедентні виклики для традиційних антидискримінаційних механізмів. Сформульовано визначення алгоритмічної дискримінації як прихованої, ненавмисної форми несправедливого чи упередженого ставлення до особи або групи осіб, які володіють однією чи кількома захищеними ознаками, що спричинена функціонуванням систем автоматизованого прийняття рішень, яка полягає у масштабному та самопідсилюваному обмеженні їхніх прав внаслідок обробки історично упереджених вихідних даних чи застосування нейтральних проксі-змінних.

У статті виокремлено та проаналізовано ключові ознаки цього явища: латентний характер та недостатня прозорість роботи алгоритмів, які перешкоджають виявленню упередженого ставлення, поки порушення не набудуть систематичного характеру; відсутність умислу на обмеження прав конкретної особи чи групи осіб, а також переважно непрямий характер алгоритмічної дискримінації; використання проксі-змінних, які мають нейтральний характер, проте через існуючі в суспільстві історичні упередження можуть призводити до нерівного ставлення; інтерсекційність або перехресний характер, зумовлений поєднанням кількох захищених ознак; суттєво пришвидшений та масштабний характер порівняно із традиційним розумінням дискримінації, зумовлений автоматизацією прийняття рішень, а також «зворотної петлі», при якому робота алгоритмів ще більше підсилює існуючу в суспільстві несправедливість поводження щодо окремих захищених груп.

Доведено, що непрозорість алгоритмів унеможливлює ефективне застосування класичних стандартів покладення тягаря доказування за статтею 14 ЄКПЛ, оскільки позбавляє заявника інструментів для встановлення причинно-наслідкового зв'язку між вихідними даними, алгоритмічними процесами та дискримінаційним поводженням. Обґрунтовано необхідність розширення правозахисної доктрини та адаптації підходів ЄСПЛ з урахуванням концепції «пояснюваного штучного інтелекту» задля гарантування справедливого захисту від алгоритмічної упередженості.

Ключові слова: *права людини, цифрові права, штучний інтелект, цифровізація міжнародного права, Європейський суд з прав людини, алгоритмічна дискримінація.*

Introduction. Today, the global community faces previously unknown challenges associated with rapid digitalisation, which has permeated virtually all spheres of social life. Originating in business and the information technology sector proper, this process has gradually extended to the public sector, precipitating significant transformations in classical approaches to public administration and the field of public services. It is also important to note the large-scale impact of digitalisation on the private life of individuals, their everyday routines, leisure, and communication with one another. On the one hand, emerging technologies have significantly simplified the implementation of certain human rights, such as the right to freedom of expression or the right to access information, the exercise of which has become easier due to the availability of digital devices and access to the Internet. On the other hand, human activity in the virtual environment has rendered individuals more vulnerable in matters concerning the protection of private life, personal data, security,

ISSN 3083-7979

equality, and non-discrimination.

Statistics for 2025 demonstrate the exceptionally rapid pace of digitalisation and the implementation of artificial intelligence technologies. According to data from the Ministry of Digital Transformation, the overall level of digital literacy among the population of Ukraine has increased significantly over the past five to six years. In particular, the percentage of surveyed Ukrainians who have not acquired any digital skills decreased from 15 per cent in 2019 to 4 per cent in 2025. At the same time, the proportion of respondents possessing basic digital skills grew by 10 per cent over the same period [1].

With regard to the use of artificial intelligence, a study conducted by the Kyiv International Institute of Sociology produced the following findings: 64 per cent of the surveyed Ukrainians do not use AI at all (12 per cent of them are entirely unaware of its existence, whilst 52 per cent are familiar with it but utilise it in neither their professional nor private lives). Among the 26 per cent of respondents who are aware of and have practically applied artificial intelligence technologies, only 12 per cent actively integrate them into their daily lives [21]. Thus, despite the fact that more than 10 per cent of the survey participants have never heard of the existence of AI, a significant proportion of the population is currently gradually using it in their work or daily lives.

It is noteworthy that half of the surveyed Ukrainians have, at least once in their lives, made a decision based on responses generated by an AI assistant [1]. Such statistics appear rather alarming, particularly where such conclusions affected the exercise of rights, freedoms, or legitimate interests. In particular, the issue of the use of artificial intelligence systems in the administration of justice has become quite acute today. For instance, in July 2024, the Superior Court of Cabarrus County (North Carolina, USA) examined the issue of imposing disciplinary liability on a lawyer who had utilised ChatGPT during the preparation of procedural documents in a case. In the course of the proceedings, the court discovered citations and conclusions generated by artificial intelligence, purportedly drawn from judicial decisions which, at that time, did not exist at all [9]. A similar issue was also identified in a judgment of the Federal Circuit and Family Court of Australia of August 2024, in which the Court established a breach of the rules of professional ethics by a lawyer, following which a restriction on the practice of law was imposed upon him. The aforementioned violation consisted in the fact that the lawyer had submitted to the court a list of extracts from judicial decisions and precedents in cases similar to the one being heard; however, this list had been entirely generated by artificial intelligence [3].

Unfortunately, such instances are increasingly encountered in the practice of foreign courts. Domestic judicial practice has been no exception regarding the delegation by parties to the proceedings of their duties concerning the formulation of legal conclusions, arguments, and the substantiation of legal positions, including to AI assistants. For instance, in the ruling of the Administrative Cassation Court within the Supreme Court of 15 January 2026 in case no. 240/14153/24, the Court pointed to purportedly established legal conclusions of the Grand Chamber and the Supreme Court, specifying case numbers and dates of judgments, which had been cited in the cassation appeal lodged by the Main Directorate of the Pension Fund of Ukraine. The court's attention was drawn to the fact that the cited legal conclusions of the Grand Chamber of the Supreme Court corresponded so closely to the appellant's position that they appeared implausible. Following a thorough verification, it transpired that the analysed argumentation was a manifestation of so-called AI "hallucinations", whereby the system presents generated responses as authentic, corroborating them with references to fictitious sources [22; 20, c. 4].

The aforementioned examples demonstrate the necessity for the adequate legal

ISSN 3083-7979

regulation of the relations arising in connection with open access to artificial intelligence technologies and their large-scale implementation in the field of public administration. It is evident that the absence of ethical norms and clear legislative restrictions in this field may precipitate unjustified interferences with an individual's private life and the violation of their fundamental rights. One manifestation of such violations may be identified as algorithmic discrimination — a recently emerged phenomenon, directly linked to automated decision-making processes. The present paper endeavours to examine the concept and characteristic features of algorithmic discrimination, as well as to identify the key issues precluding its complete elimination.

The objective of the article is a comprehensive theoretical and legal analysis of the essence and key characteristics of algorithmic discrimination as a novel challenge to digital human rights in the context of the development of artificial intelligence technologies.

Literature Review. The issue of the impact of automated systems on the implementation of fundamental human rights and freedoms is the focus of attention for a broad range of domestic and foreign scholars. The foundational principles for understanding the hidden nature of algorithms and the "black box" concept were established in the works of researchers such as F. Pasquale, C. O'Neil, and L. Edwards. The specific features of the transformation of European anti-discrimination law under the influence of artificial intelligence technologies have been comprehensively examined by F. Zuiderveen Borgesius, R. Xenidis, and M. Veale. Their works demonstrate the ineffectiveness of classical human rights protection mechanisms in combating algorithmic bias. Furthermore, particular attention has been devoted to the issue of algorithmic discrimination by J. Adams-Prassl, R. Binns, A. Kelly-Lyth, N. Rébé, B. L. Garrett, C. Rudin, and others.

With regard to domestic scholars, the issue of algorithmic discrimination has been partially addressed in the works of O. Orlov, Yu. Hurtovyi, S. Karandin, Yu. Tkach, V. Odnokolov, T. Petrenko, and others.

At the same time, it is essential to note that the overwhelming majority of studies focus on the technical and ethical aspects of the functioning of artificial intelligence. Meanwhile, issues concerning the legal nature and distinctive features of algorithmic discrimination remain insufficiently explored.

Results and Discussion. It is important to note that the phenomenon of algorithmic discrimination has emerged relatively recently. Nevertheless, contemporary research demonstrates its significant negative impact on the ability of certain social groups, united by a particular protected characteristic, to fully realise their fundamental rights and freedoms on an equal footing with others. Algorithmic discrimination is exceedingly difficult to identify: often, certain patterns become discernible only when they acquire a large-scale character. Regrettably, expectations regarding the absolute objectivity of automated decision-making have proven unfounded. The practice of applying AI systems has confirmed that algorithms are also susceptible to discrimination. Scholars attribute this issue to two categories of causes. The first is identified as the fact that algorithmic systems can inherit bias from a human operator, as source data frequently contain stereotypes, the under-representation of a particular group, or the intentional provision of incomplete or erroneous data. The second cause pertains to the operation of the systems themselves, which may disregard the metrics of small minority groups in order to generate more accurate results for the overall sample [16, c. 1]. Scholars cite the COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) software, implemented in certain states of the United States, as a prominent example of biased treatment. The system is frequently

ISSN 3083-7979

accused of bias on the grounds of race and gender, which results in more frequent refusals of early release from serving a sentence for black individuals compared to white individuals, as well as for men in comparison to women [14, c. 2]. A study carried out in California in 2021 revealed significant disparities in the treatment of individuals of different races concerning the provision of mortgage loans. The analysis of the responses, which were based on underwriting algorithms, demonstrated alarming statistics: prospective borrowers of colour were refused loans 40 to 80 per cent more frequently than white applicants. A similar issue was also identified during the monitoring of a programme used by the United States healthcare system on a commercial basis to determine the morbidity risk within the population. The primary criterion was financial healthcare expenditure: higher expenditure indicated a higher morbidity risk. However, the analysis of the practical application of the respective software revealed a systemic error that significantly underestimated the needs of African Americans. Their lower medical expenditure was more frequently attributable to low income, an inferior quality of medical care, and other factors, rather than to a lower morbidity risk [4]. The aforementioned examples demonstrate the prevalence of the phenomenon of algorithmic discrimination, which engenders high risks of human rights violations as a consequence of the use of automated decision-making systems.

In the works of foreign legal scholars, algorithmic discrimination is defined as unfair or biased treatment towards a group of individuals possessing a certain common characteristic, caused by the application of artificial intelligence technologies in decision-making processes [15, c. 3]. At the same time, it should be noted that such an interpretation is incomplete and fails to fully reflect the legal nature of this phenomenon. In order to better understand it, we shall first outline the key characteristics of algorithmic discrimination, on the basis of which we will be able to formulate its most comprehensive definition.

The first characteristic of algorithmic discrimination, which has already been mentioned above, should be considered its hidden, latent nature. This feature is frequently elucidated through the term artificial intelligence "black box", which demonstrates the problem of understanding algorithms and the explainability of the decisions adopted. Unlike traditional forms of discrimination, where the inequality of treatment or the unreasonableness of an adopted decision is obvious, algorithmic bias or unfairness is considerably more difficult to identify. This is associated with the low predictability of automated decisions: it is often exceedingly difficult or even impossible to foresee the conclusion the system will reach as a result of processing the source data. In this regard, theorists distinguish two models of comprehensible artificial intelligence: interpretable and explainable AI. The former entails the provision of clear information regarding the logic of the algorithms and the decision-making process, that is, precisely which source data are taken as a basis and how the system for their processing operates. The latter model entails the explainability of the adopted decision post factum [8, c. 2].

The European Union was among the first to take significant steps towards ensuring human rights in automated decision-making processes that produce legal effects. Thus, the General Data Protection Regulation enshrines the vital principle of transparency in the activities of authorised entities (Recital 39), the right of the data subject to be informed about the use of automated systems when decisions are made concerning them, the logic involved, the impact of such decisions, and their possible consequences (Article 15(1)(h)), as well as the right not to be subject to a decision based solely on automated processing and the right to challenge it (Article 22) [12]. With the entry into force of the 2024 EU Artificial Intelligence Act, the ‘right to an explanation of individual decision-making’ was established. It provides an individual, in respect of whom a decision has been taken using 'high-risk' artificial

ISSN 3083-7979

intelligence systems, with the opportunity to request clear and meaningful explanations of the role of the artificial intelligence system in the decision-making procedure and the main elements of the decision taken (Art. 86) [13].

It is interesting that such a latent character of algorithmic discrimination gives rise to significant procedural obstacles in protecting the human right not to be subjected to discrimination. Indeed, according to the case-law of the European Court of Human Rights, the burden of proof regarding the existence of discriminatory treatment rests with the applicant [11, c. 280]. Only in rare cases, when a certain sphere falls within the exclusive competence of state authorities, or where proving discrimination appears impossible or excessively difficult, can this burden be shifted to the Government of the State accused of violating Article 14 of the Convention for the Protection of Human Rights and Fundamental Freedoms (the Court reached such a conclusion in the cases of Anguelova v. Bulgaria, Cînța v. Romania, etc.) [6; 7]. The 1950 Convention or the protocols thereto do not explicitly enshrine a prohibition of algorithmic discrimination, which leads to the conclusion that the rules of proof established for Article 14 of the ECHR apply. This, in turn, creates legal gaps in cases where the operator of AI systems themselves cannot explain the logic behind the operation of the algorithms. The 'black box' model effectively deprives the applicant of the opportunity to prove inequality of treatment compared to other persons, as it cannot ensure the transparency of the causal link between the unfair treatment and a specific protected characteristic of the individual. In this context, the question arises as to whether the Strasbourg Court is ready to depart from its traditional approach to proof in order to ensure and protect individual rights related to the operation of AI systems — a question that remains open today.

Despite existing international legal guarantees, the 'black box' problem persists, as in the operation of most modern automated decision-making systems, the analysis of factual data constitutes only a part of the respective operation. The other part consists of prediction, on which individual decisions affecting the exercise of a person's rights and freedoms are often based [5, c. 224]. Either way, the only component that can realistically be clearly regulated is the input data. However, even strict control over the content of the inputted information is not the lever that can eradicate algorithmic discrimination. The input data may not explicitly contain criteria that cause unequal treatment, but certain structural biases that have developed historically may be embedded within it [10, c. 3]. This leads us to the next characteristic feature of algorithmic discrimination, which is the absence of intent to commit it.

Researchers in this field point out that the most common type of discriminatory treatment caused by the operation of artificial intelligence is precisely indirect discrimination. It results from a general state policy implemented without taking into account the needs or specific characteristics of particular social groups. The relevant generalised data, which primarily reflect the needs of the dominant category of the population, become the basis for the operation of algorithms, thereby causing discriminatory treatment towards minorities. Given that intent is a purely human characteristic, it would be unfair to define algorithmic discrimination as intentional in every case. An exception to this rule could only be the deliberate programming of algorithms to process input data differently depending on a specific characteristic of a person (for example, sex or racial origin). Evidently, such actions are quite rare and significantly simplify the procedure of proving discriminatory treatment. However, indirect discrimination occurs most frequently, driven by historical biases, as previously mentioned, or by incorrect labelling or imbalanced sampling, which indirectly lead to an unfair decision

or treatment [2, c. 150]. The lack of transparency in the operation of algorithms only exacerbates the problems of proving indirect algorithmic discrimination, as the motives for the decision taken, its grounds, and the actual reasoning and processing that made the adoption of such a decision possible become unclear.

Another important characteristic of algorithmic discrimination, which is closely related to the previous one, is the use of proxy variables in the machine learning process. Thus, in most cases, the operator does not embed specific protected characteristics into the operation of algorithms, such as social origin, gender, racial or national origin, but uses criteria that at first glance are completely free from bias. At the same time, a deeper analysis of the practice of their application may reveal unequal treatment in respect of a certain social group. Such neutral criteria may relate, for example, to a person's place of residence, their postcode, etc., from which their nationality or race can be inferred; career breaks (especially if they are of approximately the same duration), which may indicate childcare leave (hence, the data subject is female), and others [19, c. 13]. Thus, the use of proxy variables further complicates the process of proving discriminatory treatment in the absence of explicit protected characteristics embedded into the operation of algorithms.

The next feature is the intersectionality of algorithmic discrimination. The current regulatory framework governing the prohibition of discrimination, at both national and international levels, focuses on a single protected characteristic. That is to say, when speaking of unequal treatment, we usually mean sex, skin colour, religious beliefs, political opinions, or location, etc. In the operation of automated decision-making systems, such an approach loses its relevance, and existing human rights protection mechanisms are no longer capable of fully guaranteeing justice for individuals endowed with multiple protected characteristics. Despite the fact that artificial intelligence technologies are inherently self-learning, they were nonetheless primarily shaped by humans and reflect existing social relations, which are complex in nature and cannot be artificially structured into layers or groups depending on a specific trait. Possessing characteristics such as age, sex, or skin colour, an individual may become a victim of so-called 'intersectional' discrimination, which is based on a combination of several characteristics [17, c. 740]. For instance, in the judgment of the Grand Chamber of the Court of Justice of the EU of 18 March 2014 in the case "Z v A Government Department, The Board of Management of a Community School", the Court found no discrimination against a woman with a disability who was refused maternity leave. Her disability was caused by the absence of a uterus, which compelled her to resort to surrogacy. Given that the woman was not pregnant, the CJEU saw no violation in the refusal, separating such characteristics as sex and disability rather than considering them in combination [18]. In this situation, a profound analysis of the reasons that led to the inability to become pregnant and give birth to a child was not carried out; that is, the existence of grounds for refusing to provide social guarantees led indirectly to such a protected characteristic as the applicant's disability. Undoubtedly, today more than ever, there is a timely need to improve the system of legal guarantees and to detail the regulation of the operation of algorithms in the context of multiple or intersectional discrimination.

Another characteristic feature of algorithmic discrimination lies in its self-reinforcement, which is referred to as the 'feedback loop effect'. This feature demonstrates structural errors that are embedded into the operation of artificial intelligence systems and acquire an even greater scale in the course of their operation, leading to the deepening of existing inequality. A vivid example of such self-justification is the deployment of software designed to identify areas with a higher crime rate that require stricter control and the

ISSN 3083-7979

presence of law enforcement officers. Learning from data obtained from the reports of police officers who more frequently patrolled areas compactly populated by national minorities or persons of a different racial or ethnic origin, the system automatically identified these areas as more dangerous. In the prediction process, it was precisely these areas that were classified as requiring stricter police supervision. Regardless of the fact that the crime rate could be approximately the same across all monitored territories, it was precisely the greater presence of law enforcement authorities that led to more frequent recording of offences in these specific areas, which once again led to the confirmation of their 'dangerous' status [2, c. 151]. In fact, the feedback loop effect is inherent in discrimination based not on system errors, but on historical prejudices or stereotypes embedded within it, which are characteristic of the society of a particular state or region. In other words, where unequal treatment of members of a particular group already exists, the automation of decision-making affecting the legal status of an individual further exacerbates this inequality.

Finally, it is worth mentioning such characteristics of algorithmic discrimination as its scalability and automation. Compared to the bias inherent in a specific individual (for example, an employer who refuses to employ members of national minorities, persons over the age of 45, or single mothers), which is predominantly local in nature and affects the interests of a small number of people, algorithmic injustice can take on a nationwide scale. This is particularly the case when such biases occur in the operation of algorithms that make decisions on the granting or denial of social benefits or loans, or within predictive justice, which is becoming increasingly widespread with the development of information and communication technologies and artificial intelligence.

Summarising the characteristics outlined above, we have concluded that there is a need to expand the traditional doctrinal definition of discrimination to fully reflect the legal nature and specific features of algorithmic discrimination. In view of this, we propose the following definition. Algorithmic discrimination is a hidden, predominantly unintentional form of unfair or biased treatment of an individual or a group of persons possessing one or more protected characteristics, caused by the functioning of automated decision-making systems, which consists in the large-scale and self-reinforcing restriction of their rights resulting from the processing of historically biased source data or the application of neutral proxy variables.

Conclusions. Thus, large-scale digitalisation and the rapid implementation of artificial intelligence technologies in the fields of public administration, justice, business, and many other areas have led to the formation of new challenges for international human rights law. One of these is the phenomenon of algorithmic discrimination, which constitutes biased treatment of persons possessing protected characteristics as a result of the implementation of automated decision-making systems. On the basis of the research conducted, a number of characteristic features have been identified that distinguish it from other types of discriminatory treatment. These include: a) its hidden and predominantly indirect nature, due to the insufficient transparency of algorithms programmed to make legally significant decisions; b) its unintentional nature, demonstrating the absence of intent on the part of either the software as such or its operator to restrict the rights of a specific social group; c) the use of neutral proxy variables, which complicates proving the causal link between the input data and unfair treatment; d) intersectionality, meaning biased treatment of an individual resulting from a combination of several protected characteristics that lead to unfair treatment only in their entirety; e) the feedback loop effect of algorithmic discrimination, in which bias in automated decision-making exacerbates existing social inequality; f) the automation of processes leading to the occurrence of systemic errors that

ISSN 3083-7979

cause discriminatory treatment. The analysis of these features allows for the conclusion regarding the extreme relevance of the issues under study and the need to improve international and national regulatory frameworks to ensure the equality of individuals in the context of the automation of decision-making processes that have legal consequences. This creates the necessary doctrinal foundation for the development of new standards of proof and the implementation of the principles of explainable artificial intelligence into international and domestic human rights practice.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. 42 % дорослих і 70 % підлітків користуються ШІ в Україні: результати дослідження Дія.Освіта. Прес-офіс Міністерства цифрової трансформації України, 2025. URL: https://thedigital.gov.ua/news/education/42-doroslykh-i-70-pidlitkiv-korystuiutsia-shi-v-ukrayini-rezultaty-doslidzennia-diiaosvita (дата звернення: 14.05.2026).
2. Adams-Prassl J., Binns R., Kelly-Lyth A. Directly discriminatory algorithms. The modern law review. 2023. Vol. 86 (1). P. 144-175. URL: https://onlinelibrary.wiley.com/doi/full/10.1111/1468-2230.12759 (date of access: 14.05.2026).
3. Australian lawyer who submitted fake AI cases allowed to stay in practice. The Law Society Gazette. September 5, 2025. URL: https://www.lawgazette.co.uk/news/australian-lawyer-who-submitted-fake-ai-cases-allowed-to-stay-in-practice/5124347.article (date of access: 14.05.2026).
4. Bains C. The legal doctrine that will be key to preventing AI discrimination. Brookings Institution. September 13, 2024. URL: https://www.brookings.edu/articles/the-legal-doctrine-that-will-be-key-to-preventing-ai-discrimination/ (date of access: 14.05.2026).
5. Berk R. A. Artificial intelligence, predictive policing, and risk assessment for law enforcement. Annual Review of Criminology. 2021. Vol. 4 (1). P. 209-237. URL: https://lnk.ua/zN2K7yY47 (date of access: 14.05.2026).
6. Case of Anguelova v. Bulgaria : Judgement of the European Court of Human Rights. June 13, 2002. Application no. 38361/97. HUDOC – European Court of Human Rights. URL: https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-60505%22]} (date of access: 14.05.2026).
7. Case of Cînța v. Romania : Judgement of the European Court of Human Rights. February 18, 2020. Application no. 3891/19. HUDOC – European Court of Human Rights. URL: https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-201533%22]} (date of access: 14.05.2026).
8. Garrett B. L., Rudin C. Interpretable algorithmic forensics. Proceedings of the National Academy of Sciences. 2023. Vol. 120. № 41. P. 1-9. URL: https://www.pnas.org/doi/full/10.1073/pnas.2301842120 (date of access: 14.05.2026).
9. Generative Artificial Intelligence and its Use in the Superior Court of Cabarrus County (District 25). Administrative Order of the Superior Court of Cabarrus County (District 25) dated July 23, 2024. URL: https://nccriminallaw.sog.unc.edu/wp-content/uploads/sites/4/2026/04/Administrative-Order-Generative-AI-July-23-2004.pdf (date of access: 14.05.2026).
10. Perets O., Stagno E., Ben Yehuda E. and others. Inherent bias in electronic health records: a scoping review of sources of bias. ACM Transactions on Intelligent Systems and Technology. 2025. P. 1-26. URL: https://doi.org/10.1101/2024.04.09.24305594 (date of access: 14.05.2026).
11. Rébé N. Artificial Intelligence: Crime, War, and Justice. Ethics International Press, 2023. 345 p. URL: https://books.google.com.ua/books?hl=uk&lr=&id=juQWEQAAQBAJ&oi=fnd&pg=PA267&dq=+algorithmic+discrimination+echr&ots=20xvzsLHvK&sig=GsBkrGDO847L8_82HSPaQ9VUdpo&redir_esc=y#v=onepage&q=algorithmic%20discrimination%20echr&f=false (date of access: 14.05.2026).

ISSN 3083-7979

12. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). № 32016R0679. April 27, 2016. URL: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32016R0679 (date of access: 14.05.2026).
13. 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). № 32024R1689. June 13, 2024. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (date of access: 14.05.2026).
14. Rodà A. The COMPAS case: an educational journey for explaining fairness in AI-based applications. CEUR Workshop Proceedings. 2025. Vol. 3961. P. 1-9. URL: https://ceur-ws.org/Vol-3961/paper4.pdf (date of access: 14.05.2026).
15. Varona D., Suárez J. L. Discrimination, bias, fairness, and trustworthy AI. Applied Sciences. 2022. Vol. 5826. Art. 12. P. 1-13. URL: https://www.mdpi.com/2076-3417/12/12/5826 (date of access: 14.05.2026).
16. Wang X., Zhang Y., Zhu R. A brief review on algorithmic fairness. Management System Engineering. 2022. Vol. 1. P. 1-13. URL: https://link.springer.com/article/10.1007/s44176-022-00006-z (date of access: 14.05.2026).
17. Xenidis R. Tuning EU equality law to algorithmic discrimination: Three pathways to resilience. Maastricht Journal of European and Comparative Law. 2020. Vol. 27 (6). P. 736-758. URL: https://journals.sagepub.com/doi/full/10.1177/1023263X20982173 (date of access: 14.05.2026).
18. Z v A Government Department, The Board of Management of a Community School : Judgment of the Court (Grand Chamber). May 12, 2014. Case C-363/12. EUR-Lex. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62012CA0363&qid=1774239213111 (date of access: 14.05.2026).
19. Zuiderveen Borgesius F. Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe, Directorate General of Democracy, 2018. 51 p. URL: https://arxiv.org/pdf/2510.13465 (date of access: 14.05.2026).
20. В. Крат. Штучний інтелект і судочинство. 6 грудня 2024 року. 18 с. URL: https://court.gov.ua/storage/portal/supreme/prezentacii_2024/Prezent_Shtychn_intel_Sydoch.pdf (дата звернення: 14.05.2026).
21. Використання штучного інтелекту населенням України : пресреліз підготовлений президентом КМІС Володимиром Паніотто. Київський міжнародний інститут соціології, 2025. URL: https://kiis.com.ua/?lang=ukr&cat=reports&id=1528&page=1 (дата звернення: 14.05.2026).
22. ШІ та правнича діяльність: як реагують суди. Національна асоціація адвокатів України, 5 березня 2026 р. URL: https://unba.org.ua/publications/11425-shi-ta-pravnicha-diyal-nist-yak-reaguyut-sudi.html (дата звернення: 14.05.2026).

REFERENCES:

1. Pres-ofis Ministerstva tsyfrovoi transformatsii Ukrainy. (2025). 42 % doroslykh i 70 % pidlitkiv korystuiutsia ShI v Ukraini: rezultaty doslidzhennia Diia.Osvita [42% of adults and 70% of teenagers use AI in Ukraine: Results of the Diia.Osvita study]. https://thedigital.gov.ua/news/education/42-doroslykh-i-70-pidlitkiv-korystuiutsia-shi-v-ukrayini-rezultaty-doslidzennia-diiaosvita.
2. Paniotto, V. (2025). Vykorystannia shtuchnoho intelektu naselenniam Ukrainy: presreliz pidhotovlenyi prezydentom KMIS Volodymyrom Paniotto [Use of artificial intelligence by the population of Ukraine: Press release prepared by KIIS president Volodymyr Paniotto]. Kyivskyi mizhnarodnyi instytut sotsiolohii.

ISSN 3083-7979

https://kiis.com.ua/?lang=ukr&cat=reports&id=1528&page=1.
3. Superior Court of Cabarrus County (District 25). (2024, July 23). Generative artificial intelligence and its use in the Superior Court of Cabarrus County (District 25) [Administrative order]. https://nccriminallaw.sog.unc.edu/wp-content/uploads/sites/4/2026/04/Administrative-Order-Generative-AI-July-23-2004.pdf.
4. Australian lawyer who submitted fake AI cases allowed to stay in practice. (2025, September 5). The Law Society Gazette. https://www.lawgazette.co.uk/news/australian-lawyer-who-submitted-fake-ai-cases-allowed-to-stay-in-practice/5124347.article.
5. Natsionalna asotsiatsiia advokativ Ukrainy. (2026, March 5). ShI ta pravnycha diialnist: yak reahuiut sudy [AI and legal practice: How courts react]. https://unba.org.ua/publications/11425-shi-ta-pravnicha-diyal-nist-yak-reaguyut-sudi.html.
6. Krat, V. (2024, December 6). Shtuchnyi intelekt i sudochynstvo [Artificial intelligence and judicial proceedings]. https://court.gov.ua/storage/portal/supreme/prezentacii_2024/Prezent_Shtychn_intel_Sydoch.pdf.
7. Wang, X., Zhang, Y., & Zhu, R. (2022). A brief review on algorithmic fairness. Management System Engineering, 1, 1–13. https://link.springer.com/article/10.1007/s44176-022-00006-z.
8. Rodà, A. (2025). The COMPAS case: An educational journey for explaining fairness in AI-based applications. CEUR Workshop Proceedings, 3961, 1–9. https://ceur-ws.org/Vol-3961/paper4.pdf.
9. Bains, C. (2024). The legal doctrine that will be key to preventing AI discrimination. Brookings Institution. https://www.brookings.edu/articles/the-legal-doctrine-that-will-be-key-to-preventing-ai-discrimination/.
10. Varona, D., & Suárez, J. L. (2022). Discrimination, bias, fairness, and trustworthy AI. Applied Sciences, 12(12), Article 5826. https://www.mdpi.com/2076-3417/12/12/5826.
11. Garrett, B. L., & Rudin, C. (2023). Interpretable algorithmic forensics. Proceedings of the National Academy of Sciences, 120(41), 1–9. https://www.pnas.org/doi/full/10.1073/pnas.2301842120.
12. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). (2016, April 27). https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32016R0679.
13. 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). (2024, June 13). https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.
14. Rébé, N. (2023). Artificial intelligence: Crime, war, and justice. Ethics International Press. https://books.google.com.ua/books?hl=uk&lr=&id=juQWEQAAQBAJ&oi=fnd&pg=PA267&dq=+algorithmic+discrimination+echr&ots=20xvzsLHvK&sig=GsBkrGDO847L8_82HSPaQ9VUdpo&redir_esc=y#v=onepage&q=algorithmic%20discrimination%20echr&f=false.
15. European Court of Human Rights. (2002, June 13). Case of Anguelova v. Bulgaria (Application no. 38361/97). HUDOC. https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-60505%22]}.
16. European Court of Human Rights. (2020, February 18). Case of Cînța v. Romania (Application no. 3891/19). HUDOC. https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-201533%22]}.
17. Berk, R. A. (2021). Artificial intelligence, predictive policing, and risk assessment for law enforcement. Annual Review of Criminology, 4(1), 209–237. https://lnk.ua/zN2K7yY47.

ISSN 3083-7979

18. Perets, O., Stagno, E., Ben Yehuda, E., et al. (2025). Inherent bias in electronic health records: A scoping review of sources of bias. ACM Transactions on Intelligent Systems and Technology, 1–26. https://doi.org/10.1101/2024.04.09.24305594.
19. Adams-Prassl, J., Binns, R., & Kelly-Lyth, A. (2023). Directly discriminatory algorithms. The Modern Law Review, 86(1), 144–175. https://onlinelibrary.wiley.com/doi/full/10.1111/1468-2230.12759.
20. Zuiderveen Borgesius, F. (2018). Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe, Directorate General of Democracy. https://arxiv.org/pdf/2510.13465.
21. Xenidis, R. (2020). Tuning EU equality law to algorithmic discrimination: Three pathways to resilience. Maastricht Journal of European and Comparative Law, 27(6), 736–758. https://journals.sagepub.com/doi/full/10.1177/1023263X20982173.
22. Court of Justice of the European Union. (2014, May 12). Z v A Government Department, The Board of Management of a Community School (Case C-363/12). EUR-Lex. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62012CA0363&qid=1774239213111.

Дата першого надходження/Date of first submission: 31.03.2026
Прийнято до друку/Accepted: 30.04.2026
Опубліковано/Published: 01.06.2026

ISSN 3083-7979