

DOI 10.34132/mspc2025.02.13.08

Володимир Ємельянов
Олексій Непша

РИЗИК АЛГОРИТМІЧНОЇ УПЕРЕДЖЕНОСТІ ЦИФРОВОЇ ФЕМІДИ: ПРАКТИЧНИЙ АСПЕКТ

В тезах представлено аналіз такого потенційного ризику використання штучного інтелекту в судовій владі як алгоритмічна упередженість. На підставі практичного юридичного кейсу (рішення Європейського суду з прав людини у справі Опуз проти Туреччини 2009 року) продемонстровано небезпеку подібної упередженості. Наголошено на необхідності для публічної влади України, інтегруючи штучний інтелект в правосуддя, не лише зважати на його переваги, а й оцінювати потенційні ризики.

Ключові слова: корупція, судова влада, правосуддя, штучний інтелект, алгоритмічна упередженість.

Відповідно до урядової Концепції розвитку штучного інтелекту в Україні (надалі – AI, від англійської – artificial intelligence), ухваленій у грудні 2020 року, одним із пріоритетних завдань в сфері правосуддя є винесення судових рішень у справах незначної складності (за взаємною згодою сторін) на основі результатів аналізу, здійсненого з використанням технологій AI стану дотримання законодавства та судової практики [1].

Загалом, будучи прихильниками урядових кроків щодо цифровізації правосуддя і оцінюючи їх, насамперед, як інноваційний інструмент потенційної корупційної превенції в судовій владі, вважаємо за доцільне звернути увагу на один з ризиків, що пов'язаний із імплементацією AI та носить комбінований, етико-філософський і прикладний характер. Ми ідентифікуємо його як алгоритмічну упередженість цифрового судді.

Алгоритмічна упередженість AI полягає в потенційній і високо ймовірній залежності цієї технології від світогляду й світосприйняття її «творців», розробників, адже AI не народжується з «чистого аркуша», його навчають люди й на даних, які також надають люди. Виникає ризик перетворення упередженості розробників програмного коду (будь то співробітники державних органів чи корпоративних структур) в недоброчесність цифрового судді, тому що в алгоритми свідомо чи необережно можуть закладатися як власні індивідуальні або групові чи колективні приховані заобони й тенденційність чи пріоритети.

Якщо історичні судові дані, правові звичаї і традиції, на яких навчається AI, містять системні упередження, то технологія «засвоїть» і почне відтворювати цю несправедливість, але вже під маскою об'єктивності, що тільки вкоренить несправедливість на глибокому структурному рівні.

Досить показовим може стати приклад Туреччини – країни, в якій історично існує проблема із домашнім насильством щодо жінок, а поблажливе ставлення до неї державних органів часто і справедливо критикується. Так, в червні 2009 року в справі «Опуз проти Туреччини» (заява № 33401/02) Європейський суд з прав людини (далі – ЄСПЛ) ухвалив рішення, яке стало знаковим. Згідно з фактичними обставинами протягом багатьох років чоловік заявниці Нахіде Опуз неодноразово бив її саму та її матір, погрожував їм убивством, завдавав ножових поранень та навіть намагався переїхати автомобілем. Жінки десятки разів зверталися до поліції та прокуратури, писали заяви, надавали медичні довідки про побої, натомість реакція правоохоронних органів була відверто «млявою»: кримінальні провадження або не порушували, або закривали після того, як жінки під тиском чоловіка забирали заяви. Чоловіку призначали незначні штрафи або короткі умовні терміни, що лише посилювало його відчуття безкарності. Врешті-решт, після чергового звільнення з-під варти чоловік застрелив матір заявниці. Лише після цього його засудили до довічного ув'язнення, але звільнили з-під варти до розгляду апеляції, і він знову почав погрожувати дружині. ЄСПЛ у рішенні відверто й прямо поклав відповідальність за смерть жінки на турецьку владу, вказавши, що владі було достеменно відомо про смертельну небезпеку, яка загрожувала жінкам, погрози вбивством були постійними, насильство не вщухало, натомість державні органи не вжили адекватних заходів, щоб запобігти вбивству. Смерть матері заявниці стала прямим наслідком цієї бездіяльності. ЄСПЛ також підкреслив, що пасивність і толерантність турецької судової системи до дій кривдника створили для нього «клімат безкарності». Коли кримінальні справи постійно закриваються, а покарання є сміховинним, це є сигналом для кривдника, що він може продовжувати свої дії. Судова практика

фактично сприяла катуванню заявниці та її матері. Проаналізувавши загальну ситуацію з домашнім насильством у Туреччині, ЄСПЛ підкреслив, що переважна більшість жертв – це жінки; загальна реакція правоохоронних органів та судів на такі випадки є пасивною, і така дискримінаційна пасивність судової системи є доказом того, що насильство толерується державою [3; 4].

Описана справа та рішення ЄСПЛ в ній стало ідеологічним і юридичним підґрунтям для розробки Конвенції Ради Європи про запобігання насильству стосовно жінок і домашньому насильству та боротьбу із цими явищами (Стамбульської конвенції), що була вчинена в травні 2011 року в турецькому місті Стамбул [2].

Примітною стала та обставина, що зазначений міжнародний документ був започаткований саме в Туреччині. І, здавалося б, турецька влада дослухалася до цивілізованого світу, визнала проблему та зробила кроки до її розв'язання. Проте печальним парадоксом стало те, що вже в березні 2021 року на підставі указу турецького президента Р. Ердогана країна вийшла зі Стамбульської конвенції, а на початку січня 2023 року Державна рада Туреччини, яка de facto виконує роль вищого адміністративного суду, остаточно визнала законним рішення президента про денонсацію цієї міжнародної угоди [5].

Як бачимо, з одного боку, проблема домашнього насильства щодо жінок у Туреччині дійсно глибока й має історичні корні, а, з іншого, існує якщо не колективне схвалення, то, принаймні, суспільна думка про допустимість подібних дій. Органи правопорядку і, що найгірше, – національні суди, не вбачають в протиправних насильницьких діях, потерпілими від яких стають переважно жінки, якоїсь значної суспільної небезпеки, і таким «правовим» ставленням фактично сприяють вчиненню правопорушень.

Тому стає одночасно і очевидно і страшно від прогнозу, які саме вихідні дані можуть бути закладені в програмний код AI-судді турецькими чоловіками-розробниками, які ґрунтуватимуть своє бачення на описаній «правовій традиції». Як наслідок, такі AI-алгоритми «органічно» і з бездоганною послідовністю відтворюватимуть закладену упередженість, навіть не усвідомлюючи її протиправну та антилюдську сутність.

За такого, критично важливим стає забезпечення дотримання одного з принципів використання AI в судових системах – принципу якості та безпеки під час обробки судових рішень і даних, – що покликаний гарантувати надійність даних, які вводяться в програмне забезпечення, їхнє надходження із перевірених та сертифікованих джерел, повну прозорість цього процесу, цілісність, безпечність та резистентність до загроз того середовища, де функціонуватимуть алгоритми. Все це означає, що обов'язковою передумовою «запуску» цифрового судді має стати ретельний аудит всіх вхідних навчальних даних та очищення їх від упереджень.

Разом із тим, навчання судді-AI та пов'язана з ним алгоритмізація правосуддя мають інший бік: дійсно розробники програмного забезпечення можуть надати AI масив даних у виді судових рішень, виходячи з яких будуть сформовані певні моделі для майбутніх рішень самого цифрового судді. Але як поведе себе електронний суддя з унікальними чи нетиповими справами, що виходитимуть за межі його «розуміння»? Адже часто в юриспруденції «диявол криється в деталях», і незначна, на перший погляд, подробиця здатна кардинально змінити суть справи та рішення в ній. Натомість AI може проігнорувати такий нюанс, якщо він не зустрічався в попередніх навчальних даних.

Отже, інтегруючи AI в правосуддя, публічна влада повинна зважати не лише на переваги цієї дійсно потужної технології, які, безумовно, є численними і вагомими, а й оцінювати ризики її використання та завчасно розробляти методи їхньої мінімізації. В протилежному випадку AI не тільки не вирішить проблему корупціогенності в судовій владі, а радше – модернізує та поглибить її.

References

1. Kontsepsiia rozvytku shchuchnogo intelektu v Ukraini. Shvaleno rozporiadzhenniam Kabinetu Ministriv Ukrainy vid 02.12.2020 r. No. 1556-r. [Concept of Artificial Intelligence Development in Ukraine. Approved by the Resolution of the Cabinet of Ministers of Ukraine dated 02.12.2020 No. 1556-r.]. *zakon.rada.gov.ua*. Retrieved from: <https://zakon.rada.gov.ua/laws/show/1556-2020-p#Text> [in Ukrainian].
2. Konventsiiia Rady Yevropy pro zapobihannia nasyilstvu stosovno zhinok i domashnomu nasyilstvu ta borotbu iz tsymy yavyshchamy (Stambulska konventsiiia) vid 11.05.2011 r. [Council of Europe Convention on Preventing and Combating Violence against Women and Domestic Violence (Istanbul Convention) dated 11.05.2011]. *zakon.rada.gov.ua*. Retrieved from: https://zakon.rada.gov.ua/laws/show/994_001-11#Text [in Ukrainian].
3. Opuz v. Turkey – 33401/02. Judgment 9.6.2009 [Section III]. EUROPEAN COURT OF HUMAN RIGHTS. *hudoc.echr.coe.int*. Retrieved from: [https://hudoc.echr.coe.int/fre#{%22itemid%22:\[%22002-1449%22\]}](https://hudoc.echr.coe.int/fre#{%22itemid%22:[%22002-1449%22]}) [in English].
4. Rishennia YESPL u spravi «Opuz proty Turechchyny» (OPUZ proty TUREChChYNY, zaiava No. – 33401/02) [YESPL Decision in the case of «Opuz v. Turkey» (OPUZ v. TURKEY application no. – 33401/02). (n.d.). *jurfem.com.ua*. Retrieved from: https://jurfem.com.ua/wp-content/uploads/2022/01/Огляд-ЄСПЛ_січень-2022.pdf [in Ukrainian].

5. Sud Turechchyny ostatochno skhvalyv vykhid zi Stambulskoi konventsii [Turkish court finally approves withdrawal from Istanbul Convention]. (n.d.). www.eurointegration.com.ua. Retrieved from: <https://surl.li/uuwqdt> [in Ukrainian].

DOI 10.34132/mspc2025.02.13.08

Volodymyr Yemelianov
Oleksii Nepsha

RISK OF ALGORITHMIC BIAS OF DIGITAL THEMIS: A PRACTICAL ASPECT

The theses present an analysis of such a potential risk in using artificial intelligence in the judiciary as algorithmic bias. Based on a practical legal case (the European Court of Human Rights judgement within trial Opuz v. Turkey in 2009), the danger of such bias is demonstrated. It is emphasized that the public authorities of Ukraine, when integrating artificial intelligence into justice, should not only consider its advantages, but evaluate potential risks as well.

Keywords: corruption, judiciary, justice, artificial intelligence, algorithmic bias.

Відомості про авторів / Information about the authors

Володимир Ємельянов, доктор наук з державного управління, професор, в.о. завідувача кафедри публічного управління та адміністрування, Юридичного факультету, Чорноморського національного університету імені Петра Могили, м. Миколаїв, Україна. E-mail: d_idu@ukr.net, orcid: <http://orcid.org/0000-0002-2995-8445>.

Volodymyr Yemelyanov, Doctor of Public Administration, Professor, Acting Head of the Department of Public Administration and Management, Faculty of Law, Petro Mohyla Black Sea National University, Mykolaiv, Ukraine. E-mail: d_idu@ukr.net, orcid: <http://orcid.org/0000-0002-2995-8445>.

Олексій Непша, аспірант кафедри публічного управління та адміністрування, Юридичного факультету, Чорноморського національного університету імені Петра Могили, м. Миколаїв, Україна. E-mail: nepsha@i.ua, orcid: <https://orcid.org/0009-0005-7836-6234>.

Oleksii Nepsha, PhD-student of the Department of Public Management and Administration, Faculty of Law, Petro Mohyla Black Sea National University, Mykolaiv, Ukraine. E-mail: nepsha@i.ua, orcid: <https://orcid.org/0009-0005-7836-6234>.