D3 Менеджмент, D8 Право
УДК: 658:378.1:351.863
УДК 340:004.056.5:351.746

Земко Алла Михайлівна
кандидат юридичних наук, професор
професор кафедри права
Луцький національний технічний університет, Україна
Zemko Alla
Candidate of Law, Professor
Professor of the Department of Law
Lutsk National Technical University, Ukraine
ORCID: 0000-0003-1969-1764

Вавдіюк Наталія Степанівна
доктор економічних наук, професор
завідувачка кафедри менеджменту
Луцький національний технічний університет, Україна
Vavdiyuk Natalia
Doctor of Economics, Professor
Head of the Department of Management
Lutsk National Technical University, Ukraine
ORCID: 0000-0001-9100-3722

Лук'янчук Юрій Анатолійович
кандидат технічних наук
доцент кафедри комп'ютерних наук
Луцький національний технічний університет, Україна
Lukianchuk Iurii
Candidate of Technical Sciences
Associate Professor of the Department of Computer Science
Lutsk National Technical University, Ukraine
ORCID: 0000-0001-9690-6197

ШТУЧНИЙ ІНТЕЛЕКТ У ПУБЛІЧНОМУ СЕКТОРІ: ЕТИЧНІ ПРИНЦИПИ, СТАНДАРТИ БЕЗПЕКИ ТА ВІДПОВІДАЛЬНІСТЬ ДЕРЖАВНИХ СЛУЖБОВЦІВ

Анотація. *У статті розглядаються етичні, правові та безпекові аспекти використання штучного інтелекту в системі державного управління. Акцент зроблено на необхідності формування нормативно-етичних принципів, які забезпечують прозорість алгоритмічних рішень, захист персональних даних, запобігання дискримінації та підзвітність службовців. Проаналізовано сучасні дослідження, що вказують на ризики автоматизованих систем: від технічних*

вразливостей до соціальних упереджень. Особливу увагу приділено українському досвіду, зокрема прикладу Волинської обласної військової адміністрації, яка впровадила перші регламенти щодо використання ШІ в управлінській практиці. Запропоновано практичні рекомендації для державних службовців, які стосуються критичного аналізу результатів, обмеження автоматизованих рішень, забезпечення права на апеляцію та використання лише сертифікованих систем. Підкреслюється необхідність підвищення цифрової грамотності кадрів публічного сектору, а також документування всіх дій, що здійснюються з використанням інтелектуальних технологій. Зроблено висновок, що лише за умови дотримання комплексних етичних норм штучний інтелект може стати надійним інструментом державного управління.

Ключові слова: *штучний інтелект, державне управління, етика, прозорість, персональні дані, відповідальність, цифрова безпека, дискримінація, алгоритмічне рішення.*

ARTIFICIAL INTELLIGENCE IN THE PUBLIC SECTOR: ETHICAL PRINCIPLES, SECURITY STANDARDS AND RESPONSIBILITY OF PUBLIC SERVANTS

Abstract. *The article explores ethical, legal, and security aspects of using artificial intelligence in the public administration system. Emphasis is placed on the necessity of developing normative and ethical principles that ensure transparency of algorithmic decisions, protection of personal data, prevention of discrimination, and accountability of civil servants. Contemporary research highlighting the risks of automated systems is analyzed, ranging from technical vulnerabilities to embedded social biases. Special attention is given to the Ukrainian experience, particularly the case of the Volyn Regional Military Administration, which implemented the first official guidelines for AI use in administrative practice. Practical recommendations for public officials are proposed, including critical evaluation of AI outputs, limitations on automated decisions, ensuring the right to appeal, and the use of certified systems only. The importance of improving digital literacy among public sector employees is emphasized, along with mandatory documentation of all actions involving AI technologies. The article concludes that only with strict adherence to ethical standards can artificial intelligence become a reliable tool of public governance.*

Keywords: *artificial intelligence, public administration, ethics, transparency, personal data, accountability, cybersecurity, discrimination, algorithmic decision.*

Постановка проблеми. Штучний інтелект стрімко перетворюється на один із ключових чинників трансформації державного управління. Його можливості, від обробки великих масивів даних до автоматизації рутинних адміністративних процесів, відкривають великі перспективи для підвищення

ефективності публічного сектору. Проте одночасно виникає питання етичних, правових та безпекових аспектів його застосування. Впровадження ШІ без належної нормативної бази й контролю може порушити права громадян та поставити під загрозу довіру до органів влади.

Однією з головних проблем є питання захисту персональних даних, які в державному секторі часто є вразливими через обсяги інформації, що обробляються, а також через слабкість систем кіберзахисту. Україна неодноразово ставала об'єктом цілеспрямованих кібератак, що лише підсилює актуальність цього виклику. Наприклад, у 2023-2024 роках країна посіла одне з провідних місць у світі серед цілей політично мотивованих атак на критичну інфраструктуру, включно з органами публічного управління [1, 2].

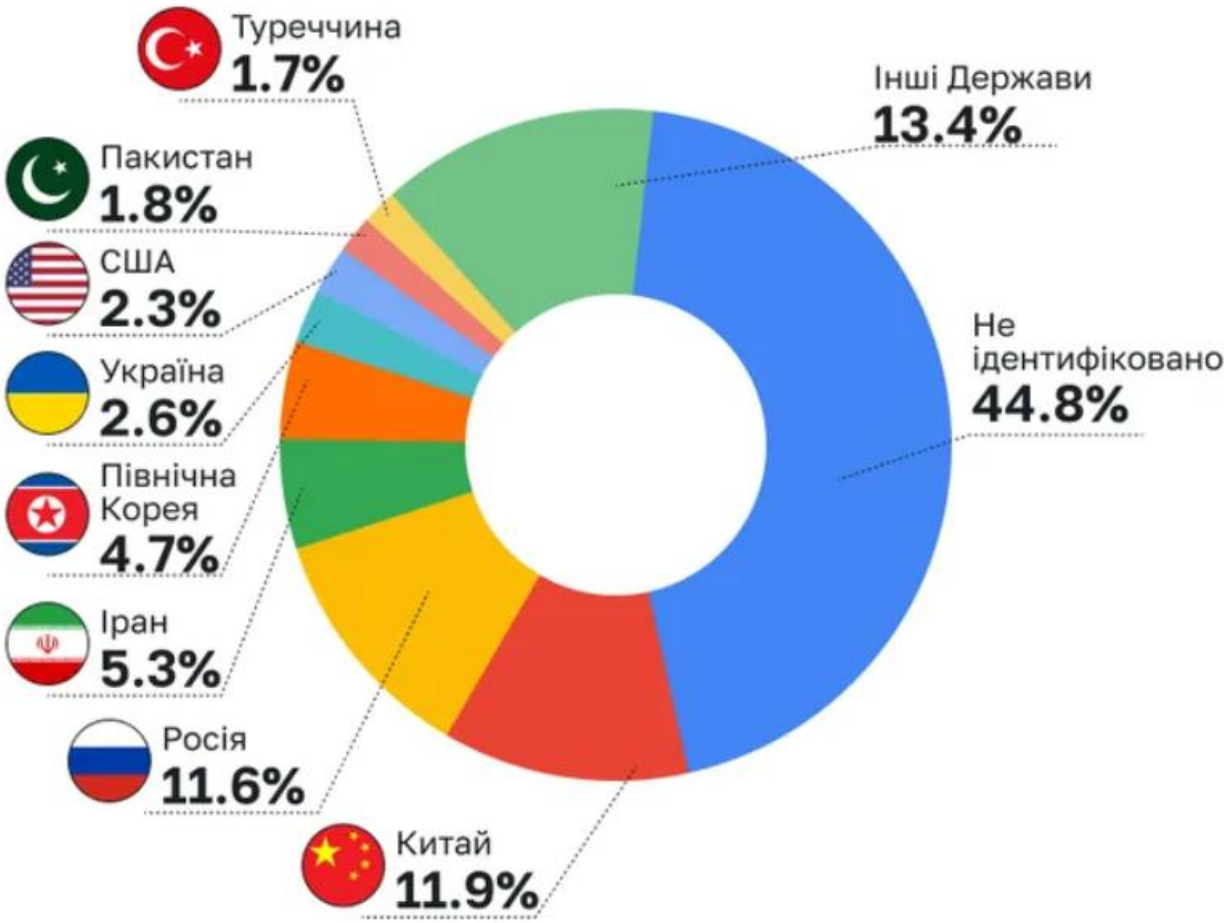

Рис. 1. Країни, які відповідальні за найбільшу частку кіберінцидентів з політичним підтекстом з 2000 по 2023 роки

Джерело: сформовано авторами за даними [1]

Іншою фундаментальною проблемою є відсутність чітких етичних стандартів, що регламентують використання алгоритмів у прийнятті рішень, які впливають на права й свободи громадян. Вже зараз існують приклади, коли ШІ може генерувати упереджені або хибні висновки, що можуть автоматично застосовуватися до справ реальних людей. Дослідження, проведене компанією Anthropic, демонструє, що навіть найсучасніші системи можуть бути зламані або маніпульовані з використанням методів, як «джейлбрейк», що дозволяє генерувати заборонені або небезпечні сценарії [3].

Ситуацію ускладнює і відсутність єдиних протоколів відповідальності, коли автоматизоване рішення було прийнято системою, але реальна відповідальність розмивається між розробниками, провайдерами технології та державними службовцями. Як наслідок, порушуються базові принципи прозорості, відкритості й підзвітності, які є фундаментом демократичного врядування.

Незважаючи на технічний прогрес, алгоритми штучного інтелекту досі не можуть гарантувати об’єктивність і неупередженість, особливо в умовах, коли дані, на яких вони навчаються, містять вбудовану соціальну, гендерну або етнічну дискримінацію. Саме тому у багатьох країнах наголошується на необхідності гуманітарного підходу до розробки та застосування ШІ, у якому пріоритетами залишаються права людини, рівність і справедливість [4].

Особливої уваги потребує і проблема інформованої згоди громадян, адже ШІ може застосовуватись без чіткого повідомлення про його використання. Це порушує принципи демократичного врядування, де кожен має право знати, як саме ухвалюється рішення щодо нього. У ЄС, наприклад, уже діють офіційні «Етичні настанови для надійного ШІ», які визначають обов’язкову відкритість, пояснюваність і підзвітність інтелектуальних систем [5].

Водночас у публічному управлінні виникає бажання автоматизувати рішення, аби пришвидшити процеси або зменшити вплив людського чинника. Проте саме ця автоматизація без людського контролю може мати згубні наслідки. Наприклад, відмова в соціальній допомозі або обмеження доступу до

державних послуг через алгоритмічну помилку здатні поставити під загрозу основні права людини. Про це описує Вірджинія Юбенкс у книзі Automating Inequality [6] та показує, як навіть добросовісне використання ШІ в соціальній сфері може мати катастрофічні наслідки для вразливих груп.

Проблема полягає не лише у впровадженні інноваційних технологій, а й у забезпеченні того, щоб ці технології відповідали етичним, правовим і безпековим стандартам, які дозволять зберегти баланс між ефективністю державного управління та захистом прав громадян.

Аналіз останніх досліджень і публікацій. Етичне та безпечне використання штучного інтелекту в державному управлінні вже не перший рік перебуває в полі зору науковців, аналітиків, розробників політик і представників міжнародних організацій. Серед сучасних досліджень простежується чітка тенденція до пошуку міждисциплінарного балансу між технологічним прогресом, правами людини, принципами відкритості управління та необхідністю забезпечення надійності цифрових інструментів.

Значну увагу приділено проблемі прозорості алгоритмів і пояснюваності рішень, які генерує штучний інтелект. У роботі Біннса [7] аналізується концепція алгоритмічної підзвітності у темі демократичної легітимності та публічного контролю, де автор вказує, що без належної обґрунтованості кожне алгоритмічне рішення втрачає правову силу та суспільну довіру. Аналогічно, у звіті Європейської комісії щодо принципів надійного ШІ наголошується на обов’язковості етичного контролю, включаючи пояснюваність дій ШІ, обмеження на автономність рішень та наявність механізмів апеляції.

У науковому середовищі також сформувався консенсус щодо того, що дані, на яких навчаються моделі ШІ, часто є джерелом прихованої дискримінації. У ґрунтовній статті Брента Міттельштадта та співавторів [8] розглядається проблема упередженості алгоритмів, яка виникає не лише з погано структурованих даних, а й з соціального контексту, в якому вони застосовуються. Автори детально класифікують етичні ризики, пов’язані з автоматизованим прийняттям рішень, зокрема у сфері публічних послуг.

В публікації Science Daily [9] вказано на майбутню конвергенцію цифрових технологій в уніфіковані інтелектуальні платформи, що посилює потребу в обмеженнях на доступ до приватних даних. Також акцентується увага на необхідності мультидисциплінарного підходу до управління розвитком штучного інтелекту, порушуються питання правової відповідальності, кібербезпеки та моральної автономії систем, вказуючи на ризики самовільної дії ШІ в умовах слабкого регулювання.

Крім того, увагу фахівців привертає проблема безпеки блокчейн-рішень, які пропонуються як можливий спосіб перевірки і фіксації дій ШІ. Проте, за даними The Wall Street Journal [10], навіть ці системи мають вразливості, зокрема у випадку розвитку квантових обчислень, які потенційно здатні зламати нинішні криптографічні протоколи.

Окремо слід згадати глобальні тренди, що вказують на майбутню конкуренцію між країнами не лише в сфері технологій, а й у сфері етики та регулювання ШІ. Технологічні гіганти, як Elon Musk, регулярно роблять публічні заяви про ймовірність перевищення ШІ людського інтелекту вже до 2030 року [11], що підкреслює надзвичайну актуальність регуляторних ініціатив на державному рівні.

Ефективне впровадження ШІ у державне управління вимагає не лише технологічної інноваційності, а й глибокого етичного усвідомлення, належного правового забезпечення та активної участі громадськості. Без цих складових навіть найсучасніші цифрові рішення можуть не лише втратити ефективність, а й стати джерелом глибоких системних загроз.

Метою статті є обґрунтування необхідності впровадження чітких етичних та безпекових принципів використання штучного інтелекту в державному управлінні. Доцільно визначити основні ризики, пов’язані з алгоритмічним ухваленням рішень у публічному секторі, проаналізувати сучасні підходи до їх мінімізації та сформулювати практичні рекомендації щодо відповідального, прозорого й правомірного застосування інтелектуальних систем у діяльності органів влади.

Виклад основного матеріалу. Сучасні тенденції у сфері державного управління дедалі частіше передбачають використання штучного інтелекту як інструменту для автоматизації процесів, аналітики, прогнозування та прийняття рішень. Утім, разом із технічним прогресом виникають численні виклики, які потребують не лише правового регулювання, а й глибокого етичного осмислення. Саме тому ключовим завданням є формування таких принципів використання ШІ, які забезпечать баланс між інноваційністю, ефективністю і захистом прав громадян.

В Україні особливу увагу привертає локальний досвід державних органів, зокрема Волинської обласної військової адміністрації, яка в партнерстві з науковцями ЛНТУ запровадила перші нормативно-етичні рамки використання ШІ в публічній службі. Цей досвід ґрунтується на принципах поінформованості громадян, обов'язкової перевірки даних, обмеження автоматизованих рішень і прозорої документації процесів [12]. Попри пілотний характер ініціативи, вона є одним із перших прикладів офіційного регламентування алгоритмічного втручання в адміністративні процеси в Україні.

Одним із найважливіших принципів, який набуває значення у цифрову добу, є інформування громадян про те, що у процесі взаємодії з органами влади можуть використовуватися автоматизовані системи. Прозорість передбачає, що державний службовець зобов'язаний повідомляти про застосування ШІ чітко, доступно та в зрозумілій формі. Такий підхід є не лише етичним, а й правовим зобов'язанням, який відповідає міжнародним стандартам, зокрема принципам пояснюваного штучного інтелекту, що рекомендовані Європейською комісією.

Особливе значення має перевірка достовірності інформації, яку генерують інтелектуальні системи. Алгоритми здатні продукувати упереджені або неточні результати. Тому критичний аналіз рішень, перевірка результатів з використанням додаткових джерел і фахова оцінка є обов'язковими елементами при використанні ШІ у державній службі. Людина повинна залишатися в центрі управлінського процесу та нести відповідальність за результати, навіть якщо рішення було прийнято автоматизовано.

В умовах зростаючих загроз інформаційній безпеці, особливо в Україні, питання захисту персональних даних виходить на перший план. Тому це потребує впровадження передових рішень у сфері шифрування, а також суворого дотримання стандартів захисту даних, таких як GDPR. Важливим принципом тут є мінімізація збору інформації – державні інституції мають право збирати лише ті дані, які необхідні для виконання конкретної задачі, а зберігання або використання надлишкових персональних відомостей повинне бути заборонене.

Ще одним ключовим напрямом етичного регулювання є запобігання дискримінації. Алгоритми, що використовують історичні або неперевірені дані, часто повторюють соціальні упередження, закладені у вихідній інформації. Це стосується вікової, гендерної, етнічної або іншої дискримінації. У зв'язку з цим необхідно впроваджувати регулярні аудити систем штучного інтелекту на предмет упередженості, а також забезпечити рівний доступ до адміністративних послуг усім групам населення без винятку.

Значної уваги заслуговує принцип обмеження автоматизованих рішень. Там, де йдеться про значний вплив на права людини, наприклад, у питаннях соціального захисту, податкових зобов'язань або доступу до медичних і освітніх послуг рішення, прийняте алгоритмічно, має бути перевірене людиною. Крім того, громадянин повинен мати право на апеляцію, а процедура перегляду рішення має бути чітко прописаною.

Крім вищезазначеного, надзвичайно важливим аспектом є підзвітність посадовців. Використання ШІ не повинно нівелювати відповідальність особи, яка ухвалює управлінське рішення. Це вимагає створення механізмів документування кожного випадку застосування інтелектуальних систем, фіксації параметрів рішень і доступу до історії їхнього формування.

Також важливим є питання підготовки кадрів. Державні службовці, які працюють із інтелектуальними технологіями, повинні проходити регулярне навчання, бути обізнаними щодо ризиків, обмежень і етичних дилем. Йдеться не лише про базові навички користування, а й про здатність аналізувати, оцінювати

наслідки та приймати рішення, спираючись не лише на технічні індикатори, а й на правові, соціальні та гуманітарні орієнтири.

Окремим завданням є використання експериментальних або несертифікованих рішень. Це створює потребу в обов'язковій верифікації всіх інструментів ШІ перед їх використанням у критично важливих державних системах.

Описані рекомендації свідчать про необхідність системного і поетапного підходу до впровадження ШІ в державному управлінні. Такий підхід має ґрунтуватися не лише на технологічних інноваціях, а передусім на довірі, прозорості, безпеці та дотриманні прав людини як найвищої цінності цифрового суспільства.

Висновки. Штучний інтелект поступово інтегрується в систему державного управління, змінюючи характер прийняття рішень, взаємодію між державою та громадянами, а також саму логіку організації управлінських процесів. Дослідження описують, що впровадження ШІ у публічний сектор є технологічним, етичним, правовим і гуманітарним викликом. Цифрова ефективність не може переважати над цінністю людської гідності, правом на приватність, рівністю та справедливістю.

Проаналізовано основні ризики, пов'язані з автоматизованими рішеннями, непрозорими алгоритмами, можливими дискримінаціями, а також зростаючими загрозами для персональних даних. Було доведено, що держава має не лише використовувати інтелектуальні технології, але й формувати довкола них чітке нормативно-етичне середовище.

Лише комплексний підхід, що об'єднує технічну безпеку, прозорість, правову відповідальність і гуманітарну орієнтацію, здатен забезпечити ефективне та безпечне використання ШІ в публічному секторі. Цей підхід повинен спиратися на низку принципів: пояснюваність, обмеження автономних рішень, превенцію дискримінації, обов'язкову участь людини у критичних процесах та постійне підвищення цифрової грамотності персоналу.

Подальші дослідження мають бути зосереджені на розробці механізмів аудиту алгоритмів, створенні національних стандартів сертифікації ШІ-систем для державного використання, адаптації правових норм до швидкоплинного технологічного розвитку та впровадженні цифрової етики в систему підготовки державних службовців. Важливо також посилити міжнародну співпрацю у сфері регулювання штучного інтелекту, адже виклики, які породжують ці технології, виходять далеко за межі однієї країни.

Література

1. Які країни проводять найбільше політизованих кібератак у світі: Україна увійшла до п'ятірки. URL: https://tech.liga.net/ua/other/article/yaki-krainy-provodiat-naibilshe-polityzovanykh-kiberatak-u-sviti-ukraina-uviishla-do-piatirky (дата звернення: 03.06.2025).

2. Україна – одна з головних цілей для кібератак у світі. URL: https://delo.ua/telecom/ukrayina-odna-z-golovnix-cilei-dlya-kiberatak-u-sviti-yak-zaxistitisya-rozpovidajemo-z-prikladami-412454/ (дата звернення: 03.06.2025).

3. APpaREnTLy THiS iS hoW yoU JaIlBreAk AI. URL: https://www.404media.co/apparently-this-is-how-you-jailbreak-ai/ (дата звернення: 04.06.2025).

4. Corinne Cath. Governing artificial intelligence: ethical, legal and technical opportunities and challenges. DOI: https://doi.org/10.1098/rsta.2018.0080 (дата звернення: 04.06.2025).

5. Ethics guidelines for trustworthy AI. URL: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (дата звернення: 05.06.2025).

6. Virginia Eubanks. Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor. P.272. URL: https://us.macmillan.com/books/9781250074317/automatinginequality/ (дата звернення: 05.06.2025).

7. Reuben Binns. Algorithmic Accountability and Public Reason. DOI: https://doi.org/10.1007/s13347-017-0263-5 (дата звернення: 06.06.2025).

8. Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. The ethics of algorithms: Mapping the debate. DOI: https://doi.org/10.1177/2053951716679679 (дата звернення: 06.06.2025).

9. Laser-based artificial neuron mimics nerve cell functions at lightning speed. URL: https://www.sciencedaily.com/releases/2024/12/241219152223.htm (дата звернення: 07.06.2025).

10. A Looming Threat to Bitcoin: The Risk of a Quantum Hack. URL: https://www.wsj.com/tech/cybersecurity/a-looming-threat-to-bitcoin-the-risk-of-a-quantum-hack-24637e29 (дата звернення: 07.06.2025).

11. Probability that AI exceeds the intelligence of all humans combined by 2030 is 100%. URL: https://x.com/elonmusk/status/1871083864111919134 (дата звернення: 08.06.2025).

12. Волинська ОВА та ЛНТУ підписали меморандум про співпрацю у сфері штучного інтелекту. URL: https://voladm.gov.ua/new/volinska-ova-ta-lntu-pidpisali-memorandum-pro-spivpracyu-u-sferi-shtuchnogo-intelektu/ (дата звернення: 08.06.2025).

References

1. Which countries carry out the most politicized cyberattacks in the world: Ukraine is in the top five. URL: https://tech.liga.net/ua/other/article/yaki-krainy-provodiat-naibilshe-polityzovanykh-kiberatak-u-sviti-ukraina-uviishla-do-piatirky (access date: 03.06.2025).

2. Ukraine is one of the main targets for cyberattacks in the world. URL: https://delo.ua/telecom/ukrayina-odna-z-golovnix-cilei-dlya-kiberatak-u-sviti-yak-zaxistitisya-rozpovidajemo-z-prikladami-412454/ (access date: 03.06.2025).

3. APPaREnTLy THiS iS hoW yoU JaIlBreAk AI. URL: https://www.404media.co/apparently-this-is-how-you-jailbreak-ai/ (accessed 04.06.2025).

4. Corinne Cath. Governing artificial intelligence: ethical, legal and technical opportunities and challenges. DOI: https://doi.org/10.1098/rsta.2018.0080 (accessed 04.06.2025).

5. Ethics guidelines for trustworthy AI. URL: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (accessed 05.06.2025).

6. Virginia Eubanks. Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor. P.272. URL: https://us.macmillan.com/books/9781250074317/automatinginequality/ (accessed 05.06.2025).

7. Reuben Binns. Algorithmic Accountability and Public Reason. DOI: https://doi.org/10.1007/s13347-017-0263-5 (accessed 06.06.2025).

8. Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. The ethics of algorithms: Mapping the debate. DOI: https://doi.org/10.1177/2053951716679679 (accessed 06.06.2025).

9. Laser-based artificial neuron mimics nerve cell functions at lightning speed. URL: https://www.sciencedaily.com/releases/2024/12/241219152223.htm (accessed 07.06.2025).

10. A Looming Threat to Bitcoin: The Risk of a Quantum Hack. URL: https://www.wsj.com/tech/cybersecurity/a-looming-threat-to-bitcoin-the-risk-of-a-quantum-hack-24637e29 (accessed 07.06.2025).

11. Probability that AI exceeds the intelligence of all humans combined by 2030 is 100%. URL: https://x.com/elonmusk/status/1871083864111919134 (accessed 08.06.2025).

12. Volyn OVA and LNTU signed a memorandum on cooperation in the field of artificial intelligence. URL: https://voladm.gov.ua/new/volinska-ova-ta-lntu-pidpisali-memorandum-pro-spivpracyu-u-sferi-shtuchnogo-intelektu/ (access date: 08.06.2025).