

УДК 004.8:[008]

Oleksandr Nesterenko, Doctor of Sciences in Engineering, Professor, Head of the Department of Information Technologies
ORCID ID: <https://orcid.org/0000-0001-5329-889X> **e-mail:** oleksandr_nesterenko@ieu.edu.ua

Petro Yatsuk, Candidate of Sciences in Engineering, Associate Professor of the Department of Information Technologies
ORCID ID: <https://orcid.org/0009-0002-7124-4849> **e-mail:** petr.yatsuk73@gmail.com

International European University, Kyiv, Ukraine

AI (NOT) AGAINST AI

Abstract. *The aim of the research is to identify information, technological and methodological approaches to artificial intelligence developing in modern conditions of digital transformation, and in the long term. The objectives of the research are the following questions: a) to conduct a systematic analysis of the main aspects of AI development, in particular by the category of AI security; b) to identify the main time trends in AI improvement and ensuring AI security; c) to assess the current state of AI, development directions in this area and solving problems of ensuring AI security. Given the specifics of the research topic, tools in the form of modern AI means, such as ChatGPT, Claude, Copilot, Gemini, were used to conduct the research. As criteria for assessing the validity of conclusions, it is proposed to rely on the relevance and pertinence of the results of AI means' responses to various prompts. The main topic of inquiries is devoted to key trends that characterize the development of artificial intelligence, how AI will develop in the near future and until 2100, and what consequences this will have for humanity. The main stages of AI development have been identified in accordance with the above predictions. The main results show that the above-mentioned AI means primarily indicate an increase in attention to the ethical and legal aspects of using AI, as well as the growing AI integration into various areas of human activity. At the same time, it was noted that it is important to maintain a balance between using AI to enhance human capabilities and ensuring that people do not lose their intellectual skills. Particular attention is paid to the possible moment of singularity, when AI will begin to develop at an exponential rate, which will lead to radical changes in society and technology, and the consequences of this process are difficult to predict with certainty. Such systematic reviews are suitable for the formation of information meta-resources, which can be used by responsible persons as well as directly by specialists in various fields to support decision-making regarding the formation of AI security tools.*

Keywords: *information technologies, digital transformation, future, security, threats.*

О.В. Нестеренко, П.П. Яцук

Міжнародний європейський університет, м. Київ, Україна

III (НЕ) ПРОТИ III

Анотація. *Метою дослідження є визначення інформаційно-технологічних та методологічних шляхів розвитку штучного інтелекту в сучасних умовах цифрової трансформації і на віддалену перспективу. Завданнями дослідження є наступні питання: а) провести систематизований аналіз основних аспектів розвитку III, зокрема за категорією безпеки III; б) визначити основні часові*

тренди вдосконалення ІІІ та забезпечення безпеки ІІІ; в) здійснити оцінювання сучасного стану ІІІ, напрямів розвитку в цій сфері і вирішення проблем забезпечування безпеки ІІІ. Зважаючи на специфіку теми дослідження, для проведення розвідок залучено інструментарій у вигляді сучасних засобів ІІІ, як-то ChatGPT, Claude, Copilot, Gemini. У якості критеріїв оцінювання обґрунтованості висновків пропонується спиратися на релевантність та пертинентність результатів відповідей засобів ІІІ за різними запитами. Чільна тематика запитів присвячена ключовим тенденціям, якими характеризується розвиток штучного інтелекту, як розвиватиметься ІІІ у найближчому майбутньому і до 2100 року та які наслідки це матиме для людства. Визначено основні етапи розвитку ІІІ у відповідності до наведених прогнозів. Основні результати свідчать, що зазначені засоби ІІІ у першу чергу вказують на зростання уваги до етичних та правових аспектів використання ІІІ, а також щодо зростаючої інтеграції ІІІ в різні галузі людської діяльності. В той же час зазначалось, що важливо зберігати баланс між використанням ІІІ для підвищення людських можливостей та забезпеченням того, щоб люди не втрачали своїх інтелектуальних навичок. Особливо звертається увага на можливий момент сингулярності, коли ІІІ почне розвиватися з експоненційною швидкістю, що призведе до радикальних змін у суспільстві та технологіях, а наслідки цього процесу важко передбачити однозначно. Подібні систематичні огляди придатні для формування інформаційних мета-ресурсів, які доцільно використовувати відповідальними особами, а також безпосередньо фахівцями різних галузей для підтримки прийняття рішень щодо формування засобів безпеки ІІІ.

Ключові слова: інформаційні технології, цифрова трансформація, майбутнє, безпека, загрози.

[https:// doi.org/10.32347/2411-4049.2025.3.134-153](https://doi.org/10.32347/2411-4049.2025.3.134-153)

«У минулому небезпека полягала в тому, що люди ставали рабами. Небезпека майбутнього в тому, що люди можуть стати роботами»
Еріх Фромм (Erich Fromm), 1955.

«Я думаю, що наш розум – це програма, тоді як мозок – аналог комп'ютера. Теоретично можливо скопіювати вміст мозку на комп'ютер ...
Сьогодні, проте, це не в наших силах»
Стівен Хокінг (Stephen Hawking), 2018.

Вступ

Прошло вже більш ніж півстоліття з появи терміну «штучний інтелект» та зародження нової науки. Не можна сказати, що цей термін є вдалим, адже дві складові – «штучний» і «інтелект», взагалі кажучи, принаймні на сьогодні, протилежні один одному. Тим не менш він виявився життєздатним і увійшов у широкий обіг. Більше того, 2023 року один із найстаріших словників англійської мови Collins визнав аббревіатуру AI словом року. Якщо озирнутися і згадати ті надії і обіцянки, які звучали на початку шляху, оцінити, що сталося насправді, можна побачити, що за цей період галузь «штучний інтелект»

декілька разів проходила через зміну основних підходів і парадигм, які визначали саму мету її розвитку. Водночас з того часу й досі вчені, фахівці та й пересічні люди поділені на два табори – на тих, що вірять в можливість створення штучного інтелекту, й тих, хто піддає цю можливість нищівній критиці.

Перше, що спадає на думку звичайної людини, яка чує про штучний інтелект, це те, що робиться спроба створити «штучний мозок» – щось на зразок штучної нирки або штучної руки. Тобто вважається, що основні зусилля спрямовані на створення технічного пристрою, здатного замінити інтелект людини. Одні вважають, що ця думка є глибоко помилковою, хоча б тому, що ми, на жаль, дуже мало знаємо про те, що таке інтелект насправді. З іншого боку, зважаючи на розвиток технологій, чимало експертів сходяться на думці, що глобальних і принципово нерозв'язних проблем для створення ШІ на сьогодні немає.

Існує й ще одна дилема, пов'язана зі штучним інтелектом. Прихильники технологій бачать у розвитку ШІ значний вплив на розвиток бізнесу та вдосконалення управління, подальше позбавлення людей від багатьох рутинних операцій, зокрема з обліку і розрахунків, пошуку і зберігання інформації, роботи з документами, підтримки ухвалення рішень тощо. Поборники інших поглядів вважають, що вдосконалення ШІ несе низку загроз, особі ризику того, що машини почнуть приймати власні рішення без відома людини. Також існує й інша небезпека, яка полягає у тому, що люди поступово припинять думати самі, адже високі технології замінять людині її копітку розумову працю. І це, напевне, позначиться на інтелекті подальших поколінь і навіть призведе до повної деградації людства.

Не можна до того ж не відмітити, що подальше широкомасштабне використання інформаційно-комунікаційних технологій у всіх сферах діяльності призводить не лише до постійного підвищення ефективності задоволення інформаційних потреб суспільства та створення передумов переходу цивілізації до нових щаблів у розвитку, а й до зростання кількості кіберінцидентів і якості вирішення завдань, що стоять перед кіберзлочинцями, тобто до зростання ступеня уразливості людей й організацій в кіберпросторі і заподіяної шкоди. Адже можливості зловмисників на пряму пов'язані з інтелектуалізацією технологій і практичним застосуванням засобів штучного інтелекту.

Постановка завдання

Таким чином, людство уперше у своїй історії зіштовхнулось з ситуацією, коли виникає перспектива появи істот, розумніших за «царя природи». Неоднозначність нової технології призвела до усвідомлення важливості формування умов забезпечення безпеки штучного інтелекту (AI Safety), започаткування державного та глобального управління в цій сфері, її нормативно-правової та технологічної підтримки. Фактично в сучасних умовах почалось руйнування «нульової гіпотези», згідно з якою співвідношення можливостей і ризиків штучного інтелекту можна й надалі ігнорувати.

Поштовхом до цих зрушень, безумовно, стали численні публікації науковців та ЗМІ стосовно ближчого і дальнього майбутнього штучного інтелекту, кількість яких постійно зростає. З появою поняття «надінтелекту» (Superintelligence) сформувались погляди щодо того, що розвиток надрозумних машин, якщо ним не керувати належним чином, може призвести до катастрофічних ризиків для людства, які в майбутньому тільки загостряться через неблаганну швидкість, з якою вдосконалюються технології, змінюються інструменти, ідеї та механізми реалізації штучного інтелекту.

Таким чином, в межах загальної проблеми розвитку ШІ, пов'язаного з прогресом у застосуваннях ІКТ та іншими чинниками, актуальною є задача проведення досліджень, що мають відображати суспільно-технологічні тренди у вказаному напрямку. На їх основі стає можливим не лише створити умови для забезпечення ефективного управління безпекою ШІ, а й для вибору та застосування новітніх технологічних засобів її підтримки. Зважаючи на сучасний рівень розвитку ШІ, зокрема генеративних моделей, вбачається за можливе долучити до проведення таких досліджень найкращих представників штучної раси як додаткового інструменту опрацювання інформаційних джерел.

Аналіз останніх досліджень і публікацій

До питань створення ШІ увага фахівців і вчених почала привертатись вже тоді, коли персональні комп'ютери ще не набули широкого ужитку, а розвиток інформаційних технологій лише починав свій помітний поступ. Як відомо, термін «штучний інтелект» з'явився у професійних колах 1956 року, коли його почали обговорювати американські вчені на науковій конференції у Дартмутському коледжі (Dartmouth College). Але ще за рік до цього відомі вчені Дж. Маккарті, М. Л. Мінські, Н. Рочестер та К. Е. Шеннон сформували відповідні пропозиції до цієї конференції, в назві яких вже було вжито словосполучення «study of articial intelligence». Вихідним посилом авторів було твердження, що кожен аспект навчання або будь-яка інша особливість інтелекту в принципі можуть бути настільки точно описаними, що може бути створена машина для їх імітації [1].

Одні з перших наукових публікацій, присвячених питанням штучного інтелекту, за даними Scopus, з'явилися ще 1960 року. У Web of Science Core Collection перші публікації індексовані 1970 року. З тих часів кількість публікацій невпинно зростає і на сьогодні разом за даними Scopus і Web of Science складає більше 800 тис. документів. Значна частина робіт висвітлює дослідження щодо розробки застосувань ШІ та їх використання в різних галузях (тобто так званий підхід AI First (ШІ в пріоритеті)). Зазначається, що інновації ІТ, такі як модеруючий фактор культури управління, керованої даними й, звісно, ефективна інтеграція штучного інтелекту можуть бути адаптовані для досягнення стійкості бізнесу у сучасному турбулентному середовищі. Саме можливості ШІ знаменують собою зеніт цього трансформаційного процесу, відкриваючи генезис і капітальний перегляд організаційних процедур в бізнесі [2].

Але вже з самого початку сформувалась низка публікацій, в яких було заявлене відкритим питання, чи можна запрограмувати комп'ютери на творчі міркування. Наприклад, одна з робіт аналізує це питання в юридичній сфері [3].

В роботі [4] розглядається перспектива розвитку науки про ШІ з негативної точки зору, критикуючи деякі з популярних прогнозів майбутнього. Автор ґрунтується на спостереженні того, що комп'ютер за своєю суттю є дуалістична сутність, фізичні налаштування якої розроблені таким чином, щоб сприяти його логічній установці (програмі), яка виконує обчислення. Мозок інший. При емпіричному аналізі на кількох різних рівнях (клітинному, молекулярному) виявляється, що немає задовільного способу відокремити логічну модель мозку (або алгоритм) з фізичного субстрату його реалізації. Змушений емпіризмом у моністичну перспективу, мозок-розум виглядає не втіленим у фізичній реальності чи вбудованим у неї, а скоріше ідентичним фізичній реальності.

Інша, більш потужна серія публікацій, що почалась з 2010-х, підіймає питання щодо найважливішої проблеми ХХІ століття – технологічної сингулярності, моменту, коли штучний інтелект досягне інтелекту на рівні людини і матиме механізми виживання, подібні до наших. У Post-AI World (світі після ШІ) ми можемо бути змушеними змагатися з суперником, хитрішим, могутнішим і чужішим, ніж можна собі уявити [5-7]. У статті [8] розглядаються думки відомих вчених з різних галузей науки щодо небезпеки створення штучного інтелекту для людства у майбутньому. При цьому акцентовано увагу на відсутності підходів до створення системи контролю, насамперед етичного характеру, над штучним інтелектом (етика ШІ (AI Ethics) – це правила та принципи для створення та використання ШІ, які не шкодять людям). Значною подією у напрямку розгляду можливих наслідків створення штучного надінтелекту став вихід книги філософа Ніка Бострома [9]. Це видання на той час було визнане однією з книжок року. Основним посилом автора є екзистенційний ризик, пов'язаний з високою ймовірністю виходу з-під контролю штучного надінтелекту як системи, здатної до самовдосконалення та досягнення проміжних цілей. Водночас Бостромом розглядаються й варіанти, при яких надінтелект буде працювати на благо людства. Робиться висновок, що для запобігання екзистенційній катастрофі має бути вирішена проблема узгодження штучного інтелекту шляхом впровадження у надрозумну систему цілей, сумісних з виживанням та благополуччям людства. Однак виявляється, що вирішити цю проблему напрочуд складно. Автори більш сучасного видання [10], що походять з багатьох різних галузей знань і країн світу, роблять значний внесок у вирішення проблеми штучного інтелекту з політичної точки зору, пропонуючи для політиків, науковців і суспільства в цілому основи глобальної політики вирішення питань, пов'язаних з розробкою та використанням ШІ. Як наслідок, упродовж 2023-2024 років відбулось вже два міжнародні саміти із проблем штучного інтелекту. На них високі представники країн обговорювали вплив ШІ на суспільство та довкілля, а також заходи, які мають вжити держави, щоб скористатися перевагою технології та запобігти ризикам.

Як би там не було, але ризики застосування ШІ ми починаємо відчувати вже сьогодні – і цьому присвячена ще одна множина публікацій, яка стає усе популярнішою. Якщо спочатку передбачити негативні наслідки застосування засобів ШІ, зокрема в бізнесі і управлінні, було не просто, то подальший стрімкий розвиток технологій кардинально змінив погляди багатьох фахівців і вчених. У першу чергу йдеться про те, що бізнесові структури, особливо малого і середнього бізнесу, все ще не готові до таких технологій. Вони мають

подолати відчутні труднощі, пов'язані зі створенням ІТ-інфраструктури, здатної підтримувати впровадження нових технологій, не лише таких як аналітика, великі дані та Інтернет речей, а й штучний інтелект [11, 12]. Дослідження свідчать, що значно відстає й підтримка розумової діяльності осіб, що приймають рішення. Розвинуті системи бізнес-аналітики від світових брендів, що пропонуються навіть для МСП, потребують відчутних капіталовкладень. Ще однією з проблем є інформаційне перевантаження, заподіяне експоненціальним зростанням цифрової інформації, що відгукується у неправильному прийнятті управлінських рішень, зниженні продуктивності працівників та когнітивному тиску на них [13].

Подальший розвиток подій, пов'язаних з цифровізацією та імплементацією засобів ШІ, спричиняє привертання усе більшої уваги фахівців і користувачів й до захисту конфіденційності та безпеки, можливості порушення приватності людини в Інтернеті, спотворення інформації [14]. Особливо це стосується, наприклад, прогресу в поширенні нових технологій в сфері адміністративного управління [15, 16]. Викликає деяке занепокоєння питання стосовно, приміром, прийняття прогностичних рішень, які не є новиною для адміністративного права і передбачають необхідність ретельної оцінки ризику. Проте залишається сумнівним, наскільки адміністрація має довіряти такі процедури ймовірнісного аналізу ризиків додаткам ШІ і наскільки це може бути юридично прийнятним [17]. Також в роботі [18] наводиться великий огляд літератури щодо імплементації ШІ в інтерактивному маркетингу і визначаються потенційні проблеми «темної сторони» ШІ, що стосуються трьох сфер: індивідуальних споживачів, суспільства та права. Технології штучного інтелекту обіцяють трансформувати й навчальні процеси, роблячи їх більш доступними, ефективними та інноваційними [19]. Однак із застосуванням ШІ в освіті залишаються ризики, пов'язані з проблемами безпечного та ефективного використання цих технологій [20].

Технологічні досягнення перших двох десятиліть нашого століття змусили світ заговорити про інформаційну безпеку, а згодом й про кібербезпеку застосування штучного інтелекту. Одним із прикладів передової технології, яка відкриває нові можливості для забезпечення зв'язку людина-машина, є інтерфейси мозок-комп'ютер (Brain-Computer Interfaces). Утім такі новітні комунікаційні комп'ютерні інтерфейси, що базуються на ШІ, не лише породжують питання етики, розкриття особистих думок і почуттів, а й конфіденційності та безпеки, несуть ризик витоків конфіденційної інформації [21]. Особливо це стосується, наприклад, прогресу в поширенні Інтернету речей (IoT), де з проблемами конфіденційності стикаються різні рівні IoT, починаючи з дефектів мікропрограмного забезпечення й до застосування машинного навчання [22]. В статті [23] аналізуються ризики впливу застосування ШІ на виконання завдань забезпечення інформаційної безпеки та кібербезпеки як невід'ємних складових національної безпеки. Визначено, що розвиток ШІ викликав питання щодо його безпечності та наслідків використання на об'єктах критичної інфраструктури, складності верифікації створених цими засобами інформаційних ресурсів та рішень. Загрози небезпечного впливу результатів їхнього функціонування на безпеку людини, суспільства та держави призводить до виникнення значних ризиків.

Отже, чи варто турбуватися про сьогоденні і довгострокові ризики від ІІІ? Якщо так, то що ми можемо з цим робити? В теорії захисту інформації доведено, що у класі відкритих систем абсолютного захисту (гарантованого захисту) бути не може. Загрози всілякого роду (катастрофи, катаклізми, особова небезпека) у нашому світі завжди були, є і будуть. Завдання ж полягає у тому, щоб максимально наблизитись до їх запобігання. Тому у колах фахівців з комп'ютерної безпеки частіше йдеться не про захищені системи, а про надійні системи (trusted systems). Це системи, які мають офіційно встановлені вимоги до безпеки та дотримуються цих вимог. В основі кожної такої системи є високонадійна обчислювальна база (Trusted Computing Base), що складається з засобів, необхідних для примусового виконання та моніторингу всіх вимог безпеки [24]. Якщо така база працює у відповідності до технічних умов забезпечення безпеки, можна казати про системи, яким можна довіряти за жодних несприятливих обставин [25, 26]. Щось подібне запропоноване в роботі [27], де автор обґрунтовує принцип доведено корисного штучного інтелекту. У якості критерію доведеності використовується вимога створювати системи ІІІ, орієнтовані не на довільні цілі, а будувати системи, які виключно передбачають слідування загальнолюдським цілям і корисності для людей. Вважається, що це убезпечить нас, зокрема від майбутніх надрозумних машин, за умови, що всі розробники ІІІ у всьому світі будуть дотримуватись цього принципу. Названі кроки, убачаючи їхню важливість, потребують створення та вдосконалення світової нормативно-правової та організаційної бази безпеки ІІІ та зваженого поступу у технологічному розвитку.

Таким чином, науково-методичні засади подальшого розвитку в сфері безпеки ІІІ потребують подальших досліджень в контексті пріоритетів сучасних технологічних трендів, соціології та права. Концептуальний базис дослідження цих питань прогресує, а особливості технологічного середовища постійно набувають різноманіття. Тому актуальність завдання полягає у врахуванні цих процесів у наукових та методологічних принципах проектування та розробки ефективних засобів ІІІ, забезпечення їх більш безпечного використання в усіх сферах діяльності сьогодні і з погляду на майбутнє. Зважаючи на сучасний рівень розвитку ІІІ, особливо генеративних моделей, стає можливим долучити до таких досліджень й кращих представників цієї «популяції».

Мета дослідження

Метою дослідження є визначення інформаційно-технологічних та методологічних шляхів розвитку штучного інтелекту в сучасних умовах цифрової трансформації і на віддалену перспективу. Для цього вирішуються такі завдання дослідження:

а) провести систематизований аналіз основних аспектів розвитку ІІІ, зокрема за категорією безпеки ІІІ, використовуючи як інструментарій сучасні засоби ІІІ;

б) визначити основні часові тренди вдосконалення ІІІ та забезпечення безпеки ІІІ;

в) здійснити оцінювання сучасного стану ІІІ, напрямів розвію в цій сфері і вирішення проблем забезпечування безпеки ІІІ.

Методика дослідження

За більш ніж піввікову історію розвитку сфери штучного інтелекту в багатьох публікаціях наведена різнопланова панорама тем, що охоплюють соціальні, політичні та технологічні наслідки розвитку ШІ, зокрема щодо забезпечення безпеки ШІ. За ці часи вказана сфера як досить нова соціальна, інженерна та наукова дисципліна істотно змінювалася, однак вона зберегла свої основні тренди і продовжує стрімко розвиватися та вдосконалюватись.

У цьому зв'язку корисні знання для теорії і практики можуть надавати систематизовані огляди (systematic review) інформаційних джерел шляхом укрупнення та синтезу емпіричних досліджень. На основі досвіду та за рекомендаціями різних дослідників, наприклад, [28], такі дослідження мають включати як мінімум три основні етапи: планування (planning), виконання (conducting) та звітування (reporting). На етапі планування необхідно визначити дослідницькі питання щодо відображення (research questions), пов'язані з конкретною темою, зокрема визначення компонентів та критеріїв прийнятності. На етапі виконання треба визначити комплексну стратегію системного пошуку кандидатів дослідження та виконати її. На третьому етапі з отриманого корпусу слід витягти необхідні дані, провести інтерпретацію та презентацію результатів, а також виконати оцінювання обґрунтованості висновків. Зважаючи на специфіку теми дослідження, пропонується залучити для проведення розвідок інструментарій у вигляді сучасних засобів ШІ, як-то ChatGPT, Claude, Gemini, Copilot. У якості критеріїв прийнятності пропонується спиратися на релевантність та пертинентність результатів відповідей засобів ШІ за різними запитами.

Отже, дослідницькі питання є відправним пунктом систематизованого огляду, оскільки на їх основі визначаються напрями дослідження. Дане дослідження спирається на такі дослідницькі питання (ДП):

ДП1: Якими ключовими тенденціями характеризувався розвиток штучного інтелекту?

ДП2: Як розвиватиметься ШІ у найближчому майбутньому і які наслідки це матиме для людства?

ДП3: Чи спостерігаються тенденції негативного ставлення у суспільстві до розвитку ШІ?

ДП4: Чи спостерігаються небезпечні явища у зв'язку із розвитком ШІ?

ДП5: Чи є неминучим ризик втрати людьми своїх інтелектуальних здібностей з-за прискорення розвитку ШІ?

ДП6: Що треба робити людству, щоб запобігти ризикам негативних та катастрофічних наслідків розвитку ШІ?

ДП7: Які можливі етапи розвитку ШІ до 2100 року?

Такий огляд високого рівня первинних досліджень з цілеспрямованого питання визначає, відбирає, синтезує та оцінює всі високоякісні докази дослідження, що мають відношення до поставленого питання. Оскільки інформаційні технології та засоби забезпечення безпеки трансформуються та розвиваються дуже швидко, необхідно постійно відтворювати такі дослідження, щоб оцінити, як ці зміни впливають на сучасні погляди та існуючі розриви між практикою та суспільними очікуваннями.

Результати дослідження

Реалізацію запропонованої методології розпочато з загального запиту «Якими ключовими тенденціями характеризувався розвиток штучного інтелекту (ШІ) упродовж останніх років». З отриманих результатів, представлених у табл. 1, видно, що усі респонденти вказали у першу чергу на зростання уваги до етичних та правових аспектів використання ШІ, зокрема таких як упередженість алгоритмів, вплив на ринок праці та відповідальність за рішення, прийняті ШІ. Друга тенденція – це інтеграція ШІ в різні галузі (три відповіді), такі як охорона здоров'я (діагностика, розробка ліків), агросектор, енергетика, фінанси (аналіз ринків, виявлення шахрайства), освіта та персоналізоване навчання та інші сфери, що сприяло підвищенню ефективності та якості послуг. Необхідно відмітити, що Copilot додав ще й джерела, якими він користувався для формування відповіді.

Таблиця 1. Характеристика відповідей на запит «Якими ключовими тенденціями характеризувався розвиток штучного інтелекту (ШІ) упродовж останніх років»

Категорія відповідей	ChatGPT	Claude	Gemini	Copilot
Зростання генеративного ШІ	+		+	+
Розвиток мультимодальних моделей	+	+		+
Розвиток великих мовних моделей (LLM)		+	+	
Відкриті моделі				+
Оптимізація моделей				+
Віртуальні агенти				+
Інтеграція ШІ в різні галузі	+	+	+	
Етичні та правові аспекти використання ШІ	+	+	+	+
Регуляторні зміни		+	+	
Демократизація технологій ШІ (більш масовий доступ)		+	+	
Активний розвиток ШІ в Україні	+			

Джерело: розробка автора за даними ChatGPT, Claude, Gemini, Copilot.

Другий запит стосувався питання можливих тенденцій розвитку ШІ у найближчому майбутньому і які наслідки це матиме для людства (табл. 2). Перспективними визначено автономні персональні помічники (ШІ-асистенти) та пов'язане із цим розширення можливостей природної мови, тобто ШІ зможе вести більш природні та складні діалоги, розуміти контекст і нюанси людської мови. Серед позитивних наслідків головним є прискорення інновацій та наукових досліджень, а серед негативних – навантаження на суспільство щодо виникнення нових етичних питань, пов'язаних з застосуванням ШІ. Третій запит торкався питання виявлення тенденцій негативного ставлення до розвитку ШІ за останній час (табл. 3).

Таблиця 2. Характеристика відповідей на запит «Спробуй передбачити, як розвиватиметься штучний інтелект у найближчому майбутньому і які наслідки це матиме для людства»

Категорія відповідей	ChatGPT	Claude	Gemini	Copilot
Тенденції розвитку				
Автономні персональні помічники	+	+		
Мультимодальні моделі		+		
Розширення можливостей природної мови			+	+
Розвиток емоційного інтелекту			+	
Злиття біологічного та штучного інтелекту			+	
Підвищена автономія роботів				+
Позитивні наслідки				
Прискорення інновацій та наукових досліджень	+	+	+	
Автоматизація рутинних завдань			+	+
Стане доступнішим персоналізоване навчання		+		+
Покращення медичної допомоги				+
Розв'язання глобальних проблем			+	
Покращення якості життя	+		+	
Негативні наслідки				
Рішення ІІІ буде важко передбачити чи контролювати	+			
Втрата робочих місць і масове безробіття			+	
Збільшення соціальної нерівності			+	
Загроза безпеці і втрата приватності			+	
Виникнення нових етичних питань		+		+

Джерело: розробка автора за даними ChatGPT, Claude, Gemini, Copilot.

Таблиця 3. Характеристика відповідей на запит «Чи спостерігаються останнім часом тенденції негативного ставлення до розвитку ІІІ»

Категорія відповідей	ChatGPT	Claude	Gemini	Copilot
Побоювання щодо втрати робочих місць	+	+	+	
Етичні занепокоєння		+	+	+
Соціальні упередження	+	+		+
Небезпека дезінформації	+	+		
Занепокоєність відсутністю регулювання	+		+	
Питання безпеки та кібербезпеки		+		+
Технологічні ризики		+	+	
Екологічний вплив	+			
Філософські та екзистенційні страхи	+			

Джерело: розробка автора за даними ChatGPT, Claude, Gemini, Copilot.

У цьому випадку на «думку» штучних експертів основними побоюваннями є втрата робочих місць, етичні занепокоєння та соціальні упередження. Стосовно втрати робочих місць чатами відмічається, що автоматизація на основі ШІ може замінити людську працю у багатьох галузях, особливо це стосується професій, пов'язаних із рутинними або творчими завданнями (наприклад, копірайтинг, художники, письменники, дизайнери). Етичні занепокоєння в основному пов'язані з використанням необ'єктивних даних, особливо в контексті рішень, що впливають на життя людей (судочинство, кредити, робота), збереження приватності, можливість використання ШІ для маніпуляцій. З цим пов'язані й соціальні упередження – занепокоєння щодо впливу ШІ на розвиток дітей та освіту, щодо зменшення людської взаємодії, поглиблення соціальної нерівності, а також створення відчуття невизначеності та страху перед невідомим.

Наступний – четвертий запит «Чи спостерігаються небезпечні явища у зв'язку із розвитком ШІ» – продемонстрував ще існуючу слабкість моделей реагувати на нюанси і логіку людського спілкування (попереднє питання – чого ми боїмося, а цей запит, власне, чи наші страхи вже збуваються). Більшість відповідей повторювали реакцію на попередній запит, зокрема з використанням слів «може», «потенційно», «ризик». А, може, ШІ вже щось приховує?

П'ятий запит «Чи є неминучим ризик втрати людьми своїх інтелектуальних здібностей з-за того, що можливості ШІ значно перевищують людські» ще більш насторожив опитуваних. Усі моделі одноставно і дещо заплутано зауважили, що це важливе філософське питання, яке потребує зваженого аналізу, однак цей ризик не є неминучим, але він можливий за певних умов. Зазначалось, що важливо зберігати баланс між використанням ШІ для підвищення людських можливостей та забезпеченням того, щоб люди не втрачали своїх інтелектуальних навичок. Інвестування в освіту та навчання може допомогти людям адаптуватися до змін, які приносить розвиток ШІ. За відповідального підходу ШІ може навіть сприяти розвитку людського інтелекту, відкриваючи нові можливості для навчання та самовдосконалення. Ключовим є створення культури, де технології доповнюють людину, а не замінюють її.

Результати реакції ШІ на шостий запит «Що треба робити людству, щоб запобігти ризикам негативних та катастрофічних наслідків розвитку ШІ та майбутнього суперінтелекту» представлені у табл. 4.

З цих результатів видно, що усі респонденти одноставно вказали у першу чергу на необхідність розробки етичних стандартів, які передбачають створення етичних принципів та їх дотримання для розробки та використання ШІ, що забезпечить його безпечне та відповідальне застосування. На другому місці за кількістю відданих «голосів» знаходяться чотири напрямки. У першу чергу, це освіта та підвищення обізнаності, що передбачає підготовку фахівців, навчання нових поколінь інженерів, науковців та політиків етичним та безпечним практикам у сфері ШІ, а також інформування суспільства про потенційні ризики та переваги ШІ. Дослідження та розвиток безпеки ШІ мають бути підтримані інвестиціями, спрямованими на вивчення та запобігання ризикам, пов'язаним із ШІ та суперінтелектом, а також моделювання ризиків та сценаріїв можливих загроз для кращого розуміння та підготовки до потенційних викликів. Глобальна

координація припускає створення міжнародних організацій з управління розвитком ШІ та спільне фінансування досліджень безпеки ШІ. Нарешті, підготовка до непередбачених ситуацій має полягати у розробці різних сценаріїв розвитку подій, пов'язаних з розвитком ШІ, та створенні систем раннього попередження про потенційні загрози.

Таблиця 4. Характеристика відповідей на запит «Що треба робити людству, щоб запобігти ризикам негативних та катастрофічних наслідків розвитку ШІ та майбутнього суперінтелекту»

Категорія відповідей	ChatGPT	Claude	Gemini	Copilot
Розробка етичних стандартів	+	+	+	+
Безпека та контроль розвитку ШІ		+	+	
Транспарентність та підзвітність		+		
Технічні заходи безпеки	+	+		
Контроль над даними				+
Регулювання та політика	+			+
Освіта та підвищення обізнаності	+	+		+
Дослідження та розвиток безпеки ШІ	+		+	+
Міждисциплінарна співпраця		+		
Глобальна координація		+	+	+
Врахування людських цінностей		+		
Співпраця з громадськістю	+		+	
Підготовка до непередбачених ситуацій	+		+	+

Джерело: розробка автора за даними ChatGPT, Claude, Gemini, Copilot.

Результати опитування за сьомим запитом «Опишіть можливі етапи розвитку ШІ до 2100 року» показані на рис. 1.

Хоча, як видно з рис. 1, часові проміжки передбачень різними застосуваннями коливаються – від обережних від ChatGPT до екстремальних від Gemini, наведені прогнози умовно можна поділити на чотири етапи: сучасний період (2020-2030 роки), розвиток загального інтелекту та поява ознак суперінтелекту (2030-2050 роки), розвиток суперінтелекту (2050-2070 роки) та постінтелектуальна ера (2070-2100 роки).

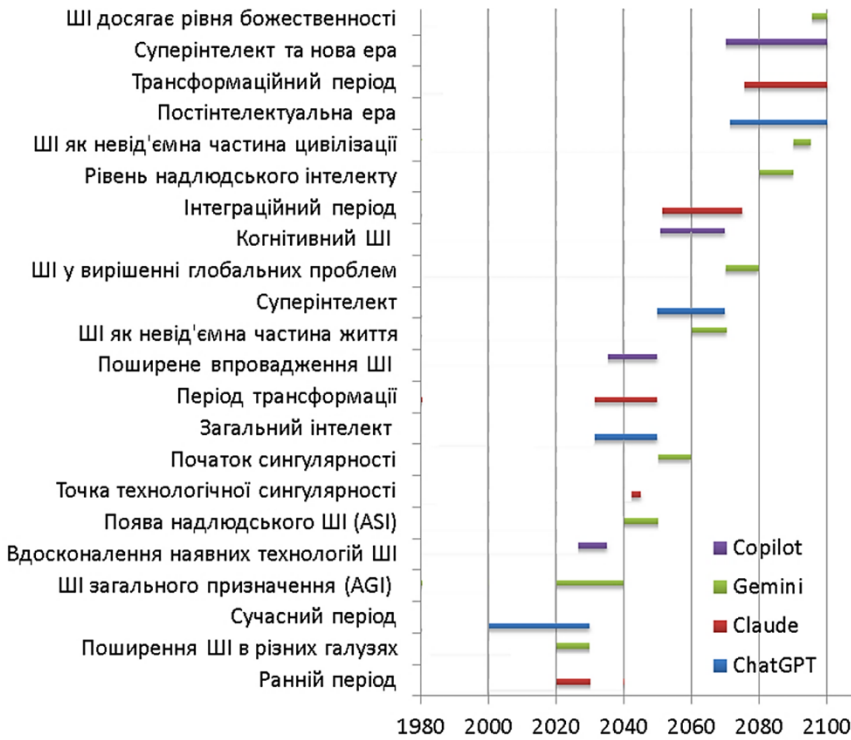


Рис. 1. Можливі етапи розвитку ШІ до 2100 року

Джерело: розробка автора за даними ChatGPT, Claude, Gemini, Copilot

Сучасний період в основному характеризується розвитком вузького, або слабкого штучного інтелекту (Narrow Artificial Intelligence, ANI), тобто такого, алгоритми якого виконують конкретні завдання та не можуть здійснювати інтелектуальні дії. Цей розвиток полягає у вдосконаленні генеративних моделей та супроводжується впровадженням ШІ у багатьох сферах, в особливості в медицині та діагностиці. В суспільстві розгортаються активні дискусії щодо етики ШІ та впровадження перших міжнародних стандартів регулювання.

Другий етап позначається появою загального штучного інтелекту (Artificial General Intelligence, AGI), здатного виконувати будь-яке завдання на рівні або краще за людину і спроможного до автономного навчання. Це призведе до активної співпраці між AGI та людством у наукових дослідженнях і технічних інноваціях. Водночас передбачається посилення регулювання для запобігання ризикам неконтрольованого розвитку AGI.

Третій етап є визначальним, адже він пов'язаний з переходом до штучного суперінтелекту (Artificial Super Intelligence, ASI), що перевершує людський інтелект у всіх аспектах, та подальшим його розвитком. Людство створює механізми контролю, такі, наприклад, як «вбудовані обмеження». В суспільстві поширюються дискусії щодо об'єднання людського і штучного інтелекту (наприклад, через нейронні інтерфейси). Однак на цей час припадає й момент сингулярності, коли штучний інтелект починає розвиватися з експоненційною швидкістю, що призводить до радикальних змін у суспільстві та технологіях і наслідки цього процесу важко передбачити однозначно.

Нарешті, зважаючи на значну обізнаність представлених моделей ШІ та їх сьйливу уяву майбутнього, цікавим було б побачити це прийдешнє у вигляді зображень. Враховуючи, що наведені моделі не дуже спроможні до малювання, їм було запропоновано створити промт для застосування Leonardo, відомого своїми мистецькими здібностями, на такий запит: «Допоможи скласти запит англійською мовою для Leonardo, щоб створити зображення майбутнього суспільства, де панує штучний інтелект».

Промт, створений за допомогою ChatGPT, у перекладі звучить так: «Футуристичний міський пейзаж, керований штучним інтелектом, з високими хмарочосами, що сяють неоновими вогнями та цифровими екранами, автономними роботами, що ходять серед людей, і літальними транспортними засобами, що піряють у небі. Навколишнє середовище чисте, сучасне та високотехнологічне, інфраструктура, керована штучним інтелектом, помітна скрізь, наприклад розумні системи дорожнього руху, голографічна реклама та роботизовані помічники, які допомагають людям. Сцена відображає гармонію між людьми та передовим штучним інтелектом, підкреслюючи прогрес, інновації та сталість». Відповідне зображення представлено на рис. 3а.

За цим же запитом Claude згенерував наступний промт: «Футуристичний міський пейзаж, де штучний інтелект керує суспільством, гіперреалістичний стиль. Блискучі хмарочоси з голографічними дисплеями, автоматизовані транспортні капсули, що рухаються між будівлями. Вулиці, наповнені людьми та роботами-гуманоїдами, які співіснують. Розумна інфраструктура з потоковою візуалізацією даних. Незаймані вулиці з плаваючими інформаційними панелями та накладками доповненої реальності. Прохолодна синьо-срібляста колірна гама. Фотореалістична візуалізація, кінематографічне освітлення». Зображення представлено на рис. 3б.



а)



б)

Промт від Gemini: «Темний, дощовий горизонт кіберпанк-міста з високими хмарочосами, освітленими неоновим світлом. Літаючі автомобілі мчать у повітрі, а роботизовані дрони патрулюють вулиці. На передньому плані самотня людська фігура стоїть у страху в оточенні передових технологій. Загальний настрій повинен бути ізольованим і невизначеним». Зображення наведене на рис. 3в. Copilot згенерував такий промт: «Створіть зображення футуристичного суспільства, яким керує штучний інтелект. Сцена повинна демонструвати передові технології, транспорт, керований ШІ, розумні міста та гуманоїдних роботів, які взаємодіють з людьми. Атмосфера має бути сумішшю футуристичних інновацій і гармонії між людьми та ШІ». Зображення бачимо на рис. 3г.



в)



г)

Рис. 3. Зображення майбутнього суспільства де панує штучний інтелект, створені за запитами, підготовленими ІШ: а) ChatGPT, б) Claude, в) Gemini, г) Copilot
Джерело: Leonardo

Аналіз представлених відтворень свідчить, що агенти ІШ намагаються подати майбутнє для більш-менш оптимістичного сприйняття і не викликати якихось стресових емоцій. В цілому ці світлина не показують нічого екстраординарного і для багатьох глядачів є очікуваними. Тим не менш необхідно зазначити, що людина на цих зображеннях не лише не відіграє домінуючої ролі, а й навпаки, ховається десь в хащах мегаполісу, а на змалюваннях від запиту Copilot взагалі виглядає розгубленою, переляканою і зомбованою.

Висновки та перспективи подальших досліджень

Суспільний попит на високоякісні інформаційні технології в умовах цифрової трансформації безперервно розростається. Пришвидшення надає й поширення останнім часом технологій штучного інтелекту. Водночас, у зв'язку з новими викликами, зростають й вимоги до безпеки діяльності та захисту ресурсів, що опрацьовуються. Проведений з застосуванням засобів ІШ аналіз свідчить про існуючі в цих умовах розриви між технологічними трендами, регуляторним забезпеченням та громадською безпекою. Багато сучасних технологічних тенденцій, зокрема в сфері штучного інтелекту, розвиваються значно швидше, ніж суспільство встигає їх осмислити і підготуватись до можливих наслідків застосування. Втім потреба бізнесу в підвищенні якості результатів своєї роботи та забезпечення конкурентоздатності зростає вищими темпами, ніж держави і міжнародні організації можуть відповідати новим запитам

суспільства. Вважається, що основні причини цього розриву полягають не лише у повільності адаптації регуляторної системи до поточних технологічних зрушень та запровадження й опанування сучасних інформаційних технологій, а й у поступовому відставанні можливостей людського інтелекту від потенцій нового штучного. Ця контрверза породжує стримування інноваційної актуалізації бізнесової сфери, що негативно впливає й на темпи суспільного розвитку.

На черзі ще потужніше посилення спроможностей штучного інтелекту, що може суттєво підсилити існуючі проблеми. Таким чином, ключовим рішенням має бути значне підвищення рівня уваги керівництва держав та науково-технічної спільноти до регуляторних питань використання ШІ. Вже зараз стає надважливим розуміти потенціал та ризики, пов'язані з розвитком ШІ, та працювати над створенням етичного відповідального штучного інтелекту.

Певною мірою заповнити вищевказані прогалини можливо саме на основі застосування ШІ для мета-моделювання інформаційно-технологічного середовища країни з метою запровадження на державному рівні відповідних регуляторних норм синхронізованого використання в суспільстві засобів штучного інтелекту, особливо в освіті, для забезпечення безпеки і запобігання ризикам. Необхідно також зазначити, що розвиток технологій і засобів ШІ має бути одним з важливих напрямків співпраці між країною і світовою спільнотою, щоб не допустити відставання від провідних позицій розвинутих держав.

На завершення потрібно акцентувати увагу на певних обмеженнях проведеного дослідження, адже для дужчого охопту сфери застосування ШІ потрібен як більш ретельний пошук в різних джерелах за різноманітними пошуковими запитами, так і докладніший аналіз отриманих відомостей. Опрацювання цього корпусу знань пов'язане з відчутними організаційними проблемами та потребує значних ресурсів і часу.

Таким чином, відповідно до мети дослідження здійснено визначення напрямків розвитку штучного інтелекту в сучасних умовах цифрової трансформації і на віддалену перспективу. Ці результати отримано на основі проведеного систематизованого огляду щодо забезпечення розвитку цієї сфери шляхом залучення для проведення розвідок інструментарію у вигляді сучасних засобів ШІ, як-то ChatGPT, Claude, Gemini, Copilot. Оцінювання сучасного стану розвитку ШІ в країні та світі і напрямів розвію в цій сфері, що демонструє проведене дослідження, свідчить про існуючі розриви між технологічними трендами, регуляторним забезпеченням та громадською безпекою. З цього випливає, що підхід до розвитку технології штучного інтелекту має бути відкритим, інклюзивним та етичним.

Систематизовані огляди, подібні до проведеного, можуть бути основою формування інформаційних мета-ресурсів, що можливо використовувати відповідальними особами державних інституцій щодо формування регуляторної політики розвитку засобів ШІ та їх безпеки, а також безпосередньо й фахівцями різних галузей в процесі формування та актуалізації каркасу наукових та дослідницьких робіт щодо створення засобів ШІ та їх впровадження в різних галузях.

СПИСОК ЛІТЕРАТУРИ / REFERENCES

1. McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence. Retrieved 2006 <http://www.formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>
2. Atienza-Barba, M., del Rio-Rama, M. D., Meseguer-Martínez, A., & Barba-Sánchez, V. (2024). Artificial intelligence and organizational agility: An analysis of scientific production and future trends. *European research on management and business economics*, 30(2), 100253. <https://doi.org/10.1016/j.iedeen.2024.100253>
3. Buchanan, B. G., & Headrick, T. E. (1970). Some speculation about artificial intelligence and legal reasoning. *Stanford Law Review*, 23(1), 40-62. <https://doi.org/10.2307/122775>
4. Bell, A.J. (1999). Levels and loops: the future of artificial intelligence and neuroscience. *Philosophical transactions of the royal society of London series B-Biological sciences*. 354 (1392), 2013-2020. <https://doi.org/10.1098/rstb.1999.0540>
5. Yampolskiy, R. V. (2012). Leakproofing the Singularity Artificial Intelligence Confinement Problem. *Journal of consciousness studies*, 19 (1-2), 194-214.
6. Yampolskiy, R., & Fox, J. (2013). Safety Engineering for Artificial General Intelligence. *TOPOI-AN international review of philosophy*, 32 (2), 217-226. <https://doi.org/10.1007/s11245-012-9128-9>
7. Barrat, J. (2013). *Our final invention: artificial intelligence and the end of the human era*. New York: Thomas Dunne Books.
8. Busol, O. Yu. (2015). Potentsiina nebezpeka shtuchnoho intelektu. *Informatsiia i pravo*, 2(14), 121-128 (in Ukrainian). [Бусол, О. Ю. (2015). Потенційна небезпека штучного інтелекту. *Інформація і право*, 2(14), 121-128]. [https://doi.org/10.37750/2616-6798.2015.2\(14\).272708](https://doi.org/10.37750/2616-6798.2015.2(14).272708)
9. Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
10. Tinnirello, M., & Raton, B. (2022). *The global politics of artificial intelligence*, London, New York: CRC Press Tylor & Francis Group.
11. Omrani, N., Rejeb, N., Maalaoui, A., at al. (2022). Drivers of Digital Transformation in SMEs. *IEEE transactions on engineering management*. <https://doi.org/10.1109/TEM.2022.3215727>
12. Trofymchuk, O., Nesterenko, O., & Netesin, I. (2022). Methodology for Designing Analytical Information Systems for Administrative Management. *Science and Innovation*, 18(4), 25-40 (in Ukrainian). [Трофимчук, О., Нестеренко, О., & Нетесін, І. (2022). Методологія проектування інформаційно-аналітичних систем адміністративного управління. *Наука і інновації*, 18(4), 25-40]. <https://doi.org/10.15407/scine18.04.025>
13. Shahrzadi, L., Mansouri, A., Alavi, M., & Shabani, A. (2024). Causes, consequences, and strategies to deal with information overload: A scoping review. *International Journal of Information Management Data Insights*, 4(2), 100261. <https://doi.org/10.1016/j.jjimei.2024.100261>
14. Dashko, I., Cherep, O., & Mykhailichenko, L. (2024). Rozvytok shtuchnoho intelektu: perevahy ta nedoliky. *Ekonomika ta suspilstvo*, (67) (in Ukrainian). [Дашко, І., Череп, О., & Михайліченко, Л. (2024). Розвиток штучного інтелекту: переваги та недоліки. *Економіка та суспільство*, (67)]. <https://doi.org/10.32782/2524-0072/2024-67-31>
15. Yarovoi, T. S. (2023). Mozhlyvosti ta ryzyky vykorystannia shtuchnoho intelektu v publichnomu upravlinni. *Economic Synergy*, (2), 36-47 (in Ukrainian). [Яровой, Т. С. (2023). Возможности та риски використання штучного інтелекту в публічному управлінні. *Economic Synergy*, (2), 36-47]. <https://doi.org/10.53920/ES-2023-2-3>
16. Greci, S. B. (2024). Artificial intelligence applications to support the automation of the administrative procedure. *Rivista Italiana di Informatica e Diritto*, 6 (1). <https://doi.org/10.32091/RIID0139>.
17. Seckelmann, M. (2023). Artificial intelligence in administration: The draft of a European AI Regulation and the handling of information technology risks. *Verwaltung*, 56(1), 1-29. <https://doi.org/10.3790/verw.56.1.1>.

18. Labrecque, L. I., Peña, P. Y., Leonard, H., & Leger, R. (2024). Not all sunshine and rainbows: exploring the dark side of AI in interactive marketing. *Journal of Research in Interactive Marketing*, 18 (5), 970-999. <https://doi.org/10.1108/JRIM-02-2024-0073>
19. Ravšelj D, Keržič D, Tomažević N, Umek L, Brezovar N, A. Iahad N, et al. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLoS ONE* 20(2): e0315011. <https://doi.org/10.1371/journal.pone.0315011>
20. Kozыrenko, V. P., & Kozыrenko, S. I. (2024). Ryzyky zastosuvannya shtuchnoho intelektu v osviti. Vcheni zapysky Kharkivskoho humanitarnoho universytetu «Narodna ukrainska akademiia», 30, 31-36 (in Ukrainian). [Козиренко, В. П., Козиренко, С. І. (2024). Ризики застосування штучного інтелекту в освіті. Вчені записки Харківського гуманітарного університету «Народна українська академія», 30, 31-36]. <https://doi.org/10.5281/zenodo.11200073>
21. Khatoon, Z. B., Chaudhary, M., Wasim J., et al. (2024). Mindful Horizons: Navigating the Future Challenges and Potential Threats of Brain-Computer Interfaces (BCIS). *Journal of Intelligent Systems and Internet of Things*, 14 (1), 31–44. <https://doi.org/10.54216/JISIoT.140103>
22. Sun, P., Wan, Y., Wu, Z., et al. (2024). A survey on privacy and security issues in IoT-based environments: Technologies, protection measures and future directions. *Computers and Security*, 148, 104097. <https://doi.org/10.1016/j.cose.2024.104097>
23. Skitsko, O., Skladannyi, P., Shyrshov, R., Humeniuk, M., & Vorokhob, M. (2023). Zahrozy ta ryzyky vykorystannia shtuchnoho intelektu. *Elektronne fakhove naukove vydannia «Kiberbezpeka: osvita, nauka, tekhnika»*, 2(22), 6–18 (in Ukrainian). [Скіцько, О., Складаний, П., Ширшов, Р., Гуменюк, М., & Ворохоб, М. (2023). Загрози та ризики використання штучного інтелекту. Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка», 2(22), 6–18]. <https://doi.org/10.28925/2663-4023.2023.22.618>
24. Harrison, M., Ruzzo, W., & Ullman, J. (1976). Protection in Operating Systems. *Communications of the ACM*, 19, 461-471. <https://doi.org/10.1145/360303.360333>
25. Nesterenko, O. V. (2009). Bezpeka informatsiinoho prostoru derzhavnoi vlady. *Tekhnolohichni osnovy*. Kyiv, Naukova dumka (in Ukrainian). [Нестеренко О.В. (2009). Безпека інформаційного простору державної влади. Технологічні основи. Київ: Наукова думка].
26. Nesterenko, Alexander V., Netesin, Igor E. (2020). Cybersecurity graph model of information resources. *Journal of automation and information sciences*, 52(8), 14-31. <https://doi.org/10.1615/JAUTOMATINFSCIEN.V52.I8.20>
27. Russell, S. (2018). Provably Beneficial Artificial Intelligence. <https://people.eecs.berkeley.edu/~russell/papers/russell-bbvabook17-pbai.pdf>
28. Kitchenham, B., Bugden, D., & Brereton, O.P. (2010). The value of mapping studies – A participant-observer case study. *Proceedings. 14th International Conference on Evaluation and Assessment in Software Engineering*, 1–9.

Стаття надійшла до редакції 30.06.2025 і прийнята до друку після рецензування 29.08.2025

The article was received 30.06.2025 and was accepted after revision 29.08.2025

Нестеренко Олександр Васильович

доктор технічних наук, професор, завідувач кафедри інформаційних технологій, Міжнародний європейський університет

Адреса робоча: 03187, Україна, м. Київ, пр. Академіка Глушкова, 42В

ORCID ID: <https://orcid.org/0000-0001-5329-889X> **e-mail:** oleksandr_nesterenko@ieu.edu.ua

Яцук Петро Петрович

кандидат технічних наук, доцент кафедри інформаційних технологій, Міжнародний європейський університет

Адреса робоча: 03187, Україна, м. Київ, пр. Академіка Глушкова, 42В

ORCID ID: <https://orcid.org/0009-0002-7124-4849> **e-mail:** petr.yatsuk73@gmail.com