



ФІЛОСОФСЬКО-ПЕДАГОГІЧНІ АСПЕКТИ ДУХОВНОГО РОЗВИТКУ УКРАЇНСЬКОГО СУСПІЛЬСТВА: ТЕОРІЯ ТА ПРАКТИКА

УДК 177:004.8

doi: 10.20998/2227-6890.2025.1.01

Р.І. ГОРБАНЬ, Я.В. ТАРАРОЄВ

ЕТИКА ШТУЧНОГО ІНТЕЛЕКТУ: ІСТОРІЯ СТАНОВЛЕННЯ ТА СУЧАСНИЙ СТАН

У статті досліджено ключові етапи становлення та розвитку етичного дискурсу штучного інтелекту, починаючи з кібернетики, через формування комп'ютерної етики в середині ХХ ст. та створення перших етичних кодексів до сучасних викликів, пов'язаних із генеративним штучним інтелектом; проаналізовано становлення ключових етичних питань (відповідальності, автономії та суспільного впливу штучного інтелекту; значну увагу приділено філософській критиці технологічного оптимізму, питанню формалізації етики, а також труднощам інтеграції моральних норм в алгоритмічні системи, висвітлені в сучасних дискусіях; у роботі висвітлено розрив між технічними підходами до розробки «етичного штучного інтелекту» та глибоким філософським осмисленням природи інтелекту та моралі в цифрову епоху.

Ключові слова: штучний інтелект, комп'ютерна етика, філософія техніки, редукціонізм, технологічний оптимізм.

R.I. HORBAN, J.V. TARAROYEV

ETHICS OF ARTIFICIAL INTELLIGENCE: HISTORY OF FORMATION AND CURRENT STATE

The article explores the key stages of the formation and development of the ethical discourse of artificial intelligence, starting with cybernetics, through the formation of computer ethics in the mid-20th century and the creation of the first ethical codes to the modern challenges associated with generative artificial intelligence; the emergence of key ethical issues (responsibility, autonomy, and social impact of artificial intelligence) is analyzed; significant attention is paid to the philosophical critique of technological optimism, the issue of formalizing ethics, as well as the difficulties of integrating moral norms into algorithmic systems highlighted in contemporary discussions; the article highlights the gap between technical approaches to developing "ethical artificial intelligence" and a deep philosophical understanding of the nature of intelligence and morality in the digital age.

Key words: ethics of artificial intelligence, computer ethics, philosophy of technology, reductionism, technological optimism.

Постановка проблеми. Сучасне суспільство характеризується активним розвитком та використанням цифрових технологій. Апогеєм останніх наразі вважається так званий ШІ – алгоритмічні системи, що здатні до обробки та аналізу великих обсягів даних з подальшим прийняттям автономних рішень. ШІ пройшов доволі тернистий шлях від абстрактної концепції до технології, яка глибоко інтегрується в різні сфери життя суспільства, активно впливаючи на його соціальні, економічні та політичні процеси. Традиційний підхід до етики ШІ часто обмежується суто технічними та юридичними аспектами, залишаючи поза увагою її глибші філософські виміри. Це призводить до проблеми недостатньої рефлексії фундаментальних питань суб'єктності,

морального вибору та впливу технологій на саму природу людського буття.

Аналіз останніх досліджень і публікацій. Останні дослідження в цій галузі демонструють зростання уваги до цієї проблематики. Стаття «AI Ethics: A Long History and a Recent Burst of Attention» простежує перші згадки терміну «етика штучного інтелекту», досліджує історію етичних роздумів про ШІ, та аналізує динаміку зростання наукових публікацій у цій сфері [1]. У роботі виявлено, що починаючи з 2010-х кількість наукових робіт цій сфері зростає експоненційно.

Стаття «An overview of artificial intelligence ethics» пропонує систематизацію сучасних підходів до етики ШІ та нову класифікацію її проблем [6]. Основна увага приділяється чотирьом аспектам:

1) ідентифікація етичних ризиків, 2) огляд наявних етичних принципів і керівництв, 3) методи розв'язання етичних проблем, 4) оцінювання етичності ШІ. Аналіз понад 140 документів, розроблених урядами, компаніями та науковими установами, показав, що універсального підходу до регулювання ШІ досі не існує.

Дисертація Кука «Essays on Machine Ethics» досліджує проблеми розробки етичного ШІ [4]. Автор пропонує концепцію «етичного ШІ з обмеженими цілями» як баланс між безпекою та автономністю, відстоюючи переваги top-down підходу (впровадження чітких етичних норм) над самонавчанням ШІ. Також розглядаються ризики, коли «розумний» ШІ може робити метанормативні висновки, що суперечать людським цінностям.

Стаття «Ethical AI is not about AI» ставить під сумнів саму можливість інтеграції етики в ШІ, стверджуючи, що етика як людський феномен, заснований на моральному судженні, рефлексії та соціальній взаємодії, принципово відрізняється від алгоритмічної природи ШІ [7]. Автори наголошують на онтологічній та епістемологічній несумісності цих сфер, вказуючи, що етику не можна звести до формальної системи правил, оскільки вона контекстуально обумовлена та змінюється разом із суспільством. Вони критикують редукацію етики до технічного питання та наголошують, що проблема «етичного ШІ» полягає не в самих алгоритмах, а в соціальному регулюванні їхнього застосування, що робить її радше суспільною, ніж технічною проблемою.

Метою статті є аналіз історичних етапів становлення етики штучного інтелекту, простеження еволюції ключових філософських концепцій та критичних підходів до проблеми автономності та відповідальності технологій, а також окреслення сучасних напрямів досліджень у цій галузі.

Виклад основного матеріалу. Етика ШІ як сфера досліджень з'явилася відносно нещодавно. Її проблематика є прямим продовженням комп'ютерної етики, яка бере свій початок в середині ХХ століття, коли на світ з'явилися перші комп'ютери [12].

Говорячи про сам термін «штучний інтелект» (Artificial Intelligence), його ввів Джон Маккарті у 1956 році під час Дартмутської конференції [8]. Саме тоді було сформульовано ШІ як науку про створення розумних машин і комп'ютерних програм, що можуть виконувати завдання, які традиційно вимагають людського інтелекту. Ця подія започаткувала новий напрям досліджень у сфері комп'ютерних наук, когнітивістики та робототехніки.

Зародження кібернетики та початок етичного дискурсу. Паралельно із появою перших комп'ютерів на початку 1940-х років в Сполучених Штатах Америки професор Массачусетського технологічного інституту (MIT) Норберт Вінер працював над військовим проектом для армії США, в результаті якого він з колегами започаткував нову галузь науки, яка отримала назву «кібернетика» [2].

У своїй книзі «Cybernetics or Control and Communication in the Animal and the Machine» він розглядає автономні системи крізь призму військової

науки [15]. Водночас ця праця не обмежується лише технічними аспектами – у ній присутні гострі етичні занепокоєння, які відображають післявоєнну атмосферу тривоги та усвідомлення наукової відповідальності. Саме цей історичний аспект є важливим для розуміння етичних міркувань Вінера.

Усвідомлюючи, що технологічний прогрес може призвести до катастрофічних наслідків, він критикує неконтрольоване використання наукових відкриттів. На відміну від технократів-оптимістів, він вбачає у цьому загрозу, що потребує етичного осмислення, і прямо попереджає, що зловживання кібернетикою є не просто можливим, а майже неминучим. Цей песимізм Вінера демонструє його усвідомлення класичної проблеми відповідальності науковця. Його рефлексія перегукується з традицією критичної думки про технології, яку можна простежити ще від Платона та його міфу про Тамуса [16, 274d–275b].

Вінер порівнює вплив цифрових технологій на інтелектуальну діяльність із тим, як паровий двигун змінив фізичну працю під час Промислової революції, підкреслюючи, що технології не лише розширюють можливості, а й змінюють саму природу людського буття. Подібно до платонівського міфу про Тевта і Тамуса, де письмо могло послабити здатність до істинного знання, автоматизація, на думку Вінера, ставить під питання статус людини як суб'єкта праці та мислення [15, с. 41].

У цьому контексті доречно згадати й неолітична революція, яка започаткувала перехід від привласнювального господарства до виробничого [3]. Технології, народжені цим процесом, такі як землеробство, тваринництво й іригація, так само докорінно змінили спосіб життя людини, як і згодом промислові машини, а сьогодні – цифрові системи, що визначають нові межі людської діяльності.

Спроби Вінера попередити профспілки про потенційні зміни у форматі праці є частиною ширшого мотиву: суспільство часто є неготовим до осмислення радикальних змін, які несуть технології. Тут можна провести аналогію з середньовічною Європою, яка стикнулася з друкарським верстатом Гутенберга або з першими промисловими машинами, що призвели до протестного руху луддитів.

Цікавою є аналогія Вінера між технологіями та магією, через яку він показує, що проблема використання автономних машин має глибоке коріння в культурних сюжетах. Звертаючись до таких творів, як «Учень чаклуна» Гете, «Рибак і джин» з «Тисячі й однієї ночі» та «Мавпяча лапа» Джейкобса, він показує, що люди за своєю природою схильні покладатися на зовнішні сили (чи то магія, чи технології), очікуючи, при цьому, що ті зрозуміють їхні бажання так, як й вони самі. Однак, що магічні сутності, що автономні системи сприймають людські бажання буквально, що й веде до непередбачуваних наслідків. На цих прикладах Вінер показує, що фундаментальні етичні дилеми залишаються незмінними, змінюється лише їхній контекст. Крім того, це відображає один з напрямів філософії техніки: чи може людина контролювати свої творіння, чи вони, зрештою, почнуть визначати її власну долю?

У своїй наступній книзі «The human use of human beings: Cybernetics and society», Вінер продовжив свою розвідку з питання етичного використання технологій, зокрема машин, і їхнього потенційного впливу на суспільство [14]. Однією з центральних тем книги є питання відповідальності при ухваленні рішень. Він наголошує, що навіть якщо машини ухвалюють рішення швидше й точніше, люди повинні зберігати контроль, оскільки алгоритмічні рішення можуть не відповідати людським моральним орієнтирам. Це порушує питання автономії та морального вибору: чи не втрачає людина суб'єктність, коли передає контроль машинам? Ця думка перегукується з критикою технологічного оптимізму, зокрема у працях Гайдеггера, який застерігав від пасивного прийняття технологічних рішень.

Вінер також розглядає концепцію «машин для управління» (*machines à gouverner*), де державний апарат може стати єдиним «гравцем» у соціальних і політичних процесах [14, с. 207-209]. Це перегукується з «Левіафаном» Гоббса – абсолютною владою, яка забезпечує порядок, протиставлений хаосу природного стану [5]. Проте, на відміну від Гоббса, у якого влада базується на страху перед хаосом, у Вінера влада спирається на тотальний контроль інформації та наукові методи управління. В якомусь сенсі, це більш технологічна та алгоритмічна версія Левіафана.

Ще однією фундаментальною проблемою взаємодії людини з машиною, яку влучно передбачив Вінер є сліпа довіра до технологій і фокус, в першу чергу, на їхній ефективності, а не мотивах та принципах, які лежать в їхній основі [14, с. 212]. Тут Вінер продовжує свою тезу про покладання на зовнішні сили та критику технологічного оптимізму, який проявляється в тенденції до пошуку “ефективних рішень” всупереч критичному мисленню, що може призвести до втрати людиною суб'єктності у взаємодії з технологіями.

Таким чином, роздуми Вінера, сформульовані пів століття тому, залишаються актуальними й сьогодні, коли ШІ починає виходити з абстрактної площини в реальну.

Перші кіберзлочини та створення кодексу професійної етики АСМ. Наступні важливі напрацювання в цій сфері в середині 1960-х пов'язані з першими кіберзлочинами. Так, тут варто згадати інженера-дослідника з МІТ Дона Паркера, який запропонував Асоціації обчислювальної техніки прийняти кодекс етики для своїх членів [2]. Його стаття опублікована в 1968 році, присвячена аналізу етичних викликів у сфері інформаційних технологій через призму людського професіоналізму [10]. У ній автор досліджує концептуальні засади етики інформаційних технологій, історичний контекст подібних ініціатив, їхнє значення для професійного середовища й обґрунтовує необхідність розробки професійного етичного кодексу для фахівців у цій галузі. Паркер виходить з припущення, що інформаційні технології формують нову професійну спільноту, яка, як і будь-яка інша сфера людської діяльності, має потребу в чітко визначених етичних принципах. Запровадження першого етичного кодексу у сфері ІТ можна інтерпретувати як певну спробу

формалізувати моральні обов'язки інженерів перед суспільством.

Крім того, Паркер акцентує на тому, що порушення етики у сфері ІТ пов'язане не стільки з самими технологіями, скільки з людським фактором. Так, його згадка про перший задокументований випадок комп'ютерного злочину є прикладом того, що технологія не є автономним суб'єктом моралі – відповідальність лежить на людині.

Паркер також одним з перших розглядає проблему конфіденційності в цифровому просторі. Якщо у 1968 р. питання полягало в доступі до окремо взятих баз даних та потенційних маніпуляцій з цими даними, то у XXI ст., масштаб цієї проблеми зріс експоненційно. У сучасних умовах проблеми, підняті Паркером, переростають у питання цифрової нерівності, контролю за даними та балансування між безпекою й особистими правами.

Критика машинного інтелекту. Наступною значущою постаттю в цей період є науковець з МІТ Джозеф Вайзенбаум, який створив комп'ютерну програму під назвою ELIZA, яка імітувала розмову психотерапевта з пацієнтом [2]. Його проєкт викликав бурхливу реакцію серед наукової спільноти, яка була здебільшого зосереджена на протиставленні інтелектуальних здібностей людини та комп'ютера. На фоні такої реакції, Вайзенбаум написав книгу «Computer power and human reason: From judgment to calculation» [13].

Центральною темою книги є критика порівняння людського розуму та обчислювальних машин. Вайзенбаум виступає проти ідеї, що інтелект можна звести до алгоритмічного процесу, а отже, що його можна відтворити у цифровій формі. Його аргумент виходить за межі технічних міркувань і має виразний філософський характер: людина мислить не як машина, оскільки її пізнання невіддільне від тілесного досвіду, інтуїції, моральних рефлексій та соціальної взаємодії. Він підкреслює, що емпіричне відтворення явища не є тотожним до відтворення його сутності: комп'ютери можуть демонструвати мовну компетентність, проте їхній спосіб обробки інформації принципово відрізняється від людського. Це співзвучно, зокрема, з аргументом Джона Сьорля у «Китайській кімнаті», який доводив, що розпізнавання символів не рівнозначне справжньому розумінню [11]. Таким чином, Вайзенбаум вказує на ключовий розрив між симуляцією розуму та справжнім розумінням: перше може бути досягнуте шляхом статистичних або символічних обчислень, але друге вимагає суб'єктивного досвіду, якого машини позбавлені.

Його критика концепції IQ є частиною ширшої дискусії про редукціонізм у визначенні інтелекту. Вайзенбаум зазначає, що спроба кількісно виміряти інтелект спрощує його природу, що, своєю чергою, веде до хибного уявлення про можливість створення машини, розумнішої за людину. Це є класичним прикладом категоріальної помилки, адже оскільки людський розум не є лише набором обчислюваних операцій, то жодне збільшення обчислювальної потужності не зробить машину по-справжньому розумною. Він показує, що ототожнення людського розуму з комп'ютерною обробкою інформації є не лише

онтологічно хибним, а й небезпечним з погляду соціальних та етичних наслідків.

Окрему увагу Вайзенбаум приділяє питанню чи можуть машини мати власні цілі, чи ж вони лише виконують те, що закладене в них людиною? Він зазначає, що суспільство часто делегує їм відповідальність, немовби визнаючи за ними певну автономність. Це підіймає дилему штучної суб'єктності, яка залишається актуальною в сучасній етиці ШІ, особливо у контексті розробки останнього.

Виокремлення комп'ютерної етики як галузі досліджень. Говорячи про сам термін «комп'ютерна етика», варто відзначити, що формально він не був присутній ні у роботах Вінера, ні у роботах Паркера. Цей термін вперше ввів американський викладач медичної етики Волтер Манер у 1976 році, досліджуючи вплив комп'ютерів на медичну етику. Він визначив комп'ютерну етику як галузь прикладної етики, що вивчає етичні проблеми, «загострені, трансформовані або створені комп'ютерними технологіями» [2].

У 1978 році його ідеями зацікавився філософ Террел Уорд Байнам, та головний редактор філософського журналу «Метафілософія», який згодом присвятив випуск свого журналу темі комп'ютерної етики. У межах цієї ініціативи він оголосив конкурс есе, переможцем якого став Джеймс Мур зі своєю працею «What is Computer Ethics» [9]. Ця робота, написана у 1985 році стала однією з фундаментальних праць у галузі комп'ютерної етики, заклавши її теоретичні основи як окремої галузі досліджень та сформулювавши ключові проблеми, які досі залишаються актуальними.

Мур вбачає появу комп'ютерної етики в тому, що нові технології створюють політичні та концептуальні вакууми, оскільки на момент їхнього впровадження часто відсутні чіткі норми регулювання та моральні орієнтири. Він також розмежує загальні етичні питання та ті, що породжені саме комп'ютерними технологіями.

Мур трактує появу комп'ютерів як технологічну революцію. Серед усіх проаналізованих ним аргументів (новизна, швидке поширення та доступність, зростання продуктивності), головну унікальність комп'ютерів він вбачає в їхній логічній гнучкості, яка забезпечує універсальність їхнього застосування, подібно до того, як парова енергія свого часу стала рушієм промислової революції: «Оскільки логіка може бути застосована всюди, потенційні можливості застосування комп'ютерних технологій здаються безмежними. Комп'ютер – це найближче до універсального інструменту, що ми маємо. Дійсно, межі комп'ютерів – це значною мірою межі нашої власної творчості [9, с. 269].» Він критикує редукаціоністське уявлення про комп'ютери як «калькулятори» та наголошує, що їхній потенціал залежить від людської творчості. Усвідомлення цієї особливості є ключовим для формування ефективної політики використання комп'ютерних технологій.

Мур, подібно до Вінера, порівнює комп'ютерну революцію з промисловою, але з трансформаційного аспекту. Він виділяє дві стадії технічної революції: впровадження (коли нові технології лише

створюються та вдосконалюються) та проникнення (коли вони стають вже невіддільною частиною соціальних інституцій). На його думку, комп'ютери змінюють не лише ефективність, а й саму природу явищ: замість «Як добре вони рахують голоси?» постає питання «Що таке чесні вибори?», замість «Як вони допомагають у роботі?» – «Що таке робота?», а в освіті – «Що є освітою в цифрову епоху?»

Ще одним важливим питанням, яке порушує Мур є фактор невидимості. Він розрізняє три її типи: невидимі складні обчислення. Перше охоплює приховані форми неетичного використання, як-от: кіберзлочини та шпигунство. Друге вказує на закладені в кодї ціннісні припущення, що впливають на поведінку користувачів, наприклад, алгоритмічну упередженість. Третє стосується обчислень, які людина, в силу своїх природних обмежень, не здатна осмислити та перевірити в повному обсязі, що породжує питання довіри до машинного аналізу.

У 1985 р. Мур писав, що людство входить у другу стадію комп'ютерної революції. Станом на першу чверть ХХІ ст., можна стверджувати, що друга стадія вже реалізована, що підтверджується трансформацією багатьох сфер людської діяльності під впливом комп'ютерних технологій. Якщо у 1985 р. комп'ютери сприймалися переважно як інструменти для оптимізації наявних завдань, то сучасний етап розвитку цифрових технологій змінює саму природу соціальних, економічних та політичних процесів. Відповідно, змінюються і питання, які постають у дискурсі: замість запитів щодо ефективності виконання завдань постає вже питання про саму природу і цінність тих чи інших видів діяльності в умовах цифрової трансформації та інтеграції ШІ.

Сучасний стан. Повертаючись до сьогодення, варто зауважити, що, якщо у ХХ – на початку ХХІ ст. дослідження мали переважно теоретичний характер, то починаючи з другої декади ХХІ ст. з розвитком генеративного ШІ акцент суттєво змістився на практичні аспекти. Етика ШІ перестала бути абстрактною проблемою майбутнього і перетворилася на нагальну сферу досліджень, що вимагає як теоретичних розробок, так і практичних рішень. Відповідно, однією з ключових особливостей сучасного етапу досліджень є зміщення акценту з теоретико-філософських роздумів до більш прикладного аналізу вже наявних проблем.

Висновки. Підсумовуючи, на основі опрацьованої літератури, можна зробити декілька висновків. По-перше, дослідження в галузі етики ШІ можна розділити на дві категорії: дослідження створення етичного ШІ та дослідження впливу ШІ на суспільство. По-друге, більшість сучасних досліджень зосереджені на практичних аспектах технології. Як наслідок, можна помітити тенденцію, що праці на цю тему написані переважно з погляду комп'ютерних наук (акцент на створенні «етичного ШІ» – алгоритмів, що мінімізують упередження, забезпечують прозорість, відповідальність тощо) і права (акцент на юридичній регуляції ШІ). Ця прогалина в академічному дискурсі відкриває

перспективи для подальших досліджень, зокрема в напрямі філософсько-етичного осмислення впливу інтеграції ШІ на соціальні та культурні процеси.

Список джерел інформації

1. Borenstein, J., Grodzinsky, F.S., Howard, A., Miller, K.W., & Wolf, M.J. (2021). AI ethics: A long history and a recent burst of attention. *Computer*, 54(1), 96 – 102. DOI <https://doi.org/10.1109/mc.2020.3034950>
2. Bynum, T.W. (2000). *A very short history of computer ethics*. The Research Center on Values in Emerging Science and Technology. Retrieved from <https://rcvest.southernct.edu/a-very-short-history-of-computer-ethics/>
3. Childe, V.G. (1936). *Man makes himself*. Watts.
4. Cook, T.B. (2023). *Essays on Machine Ethics*. The Ohio State University.
5. Hobbes, T. (1651). *Leviathan, or, The matter, forme, & power of a common-wealth ecclesiasticall and civill*. Printed for Andrew Ckooke [i.e. Crooke], at the Green Dragon in St. Pauls Church-yard.
6. Huang, C., Zhang, Z., Mao B., & Yao X. (2022). An Overview of Artificial Intelligence Ethics. *IEEE Transactions on Artificial Intelligence*, 1-21. DOI <https://doi.org/10.1109/taai.2022.3194503>
7. Johnson, D.G., & Verdicchio, M. (2023). Ethical AI is not about AI. *Communications of the ACM*, 66(2), 32 – 34. DOI <https://doi.org/10.1145/3576932>
8. McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4), 12 – 12.
9. Moor, J.H. (1985). What is Computer Ethics? *Metaphilosophy*, 16(4), 266 – 275. DOI <https://doi.org/10.1111/j.1467-9973.1985.tb00173.x>
10. Parker, D.B. (1968). Rules of ethics in information processing. *Communications of the ACM*, 11(3), 198 – 201. DOI <https://doi.org/10.1145/362929.362987>
11. Searle, J.R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417 – 424. DOI <https://doi.org/10.1017/s0140525x00005756>
12. The Z3. Retrieved from https://web.archive.org/web/20060211035420/http://irb.cs.tu-berlin.de/~zuse/Konrad_Zuse/en/Rechner_Z3.html
13. Weizenbaum, J. (1976). *Computer power and human reason: from judgment to calculation*. Freeman.
14. Wiener, N. (1950). *The human use of human beings: Cybernetics and society*. Houghton Mifflin.
15. Wiener, N. (2019). *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press.
16. Платон (2008). Діалоги. Фоліо.

References (transliterated)

1. Borenstein, J., Grodzinsky, F.S., Howard, A., Miller, K.W., & Wolf, M.J. (2021). AI ethics: A long history and a recent burst of attention. *Computer*, 54(1), 96 – 102. DOI <https://doi.org/10.1109/mc.2020.3034950>
2. Bynum, T.W. (2000). *A very short history of computer ethics*. The Research Center on Values in Emerging Science and Technology. Retrieved from <https://rcvest.southernct.edu/a-very-short-history-of-computer-ethics/>
3. Childe, V.G. (1936). *Man makes himself*. Watts.
4. Cook, T.B. (2023). *Essays on Machine Ethics*. The Ohio State University.
5. Hobbes, T. (1651). *Leviathan, or, The matter, forme, & power of a common-wealth ecclesiasticall and civill*. Printed for Andrew Ckooke [i.e. Crooke], at the Green Dragon in St. Pauls Church-yard.
6. Huang, C., Zhang, Z., Mao B., & Yao X. (2022). An Overview of Artificial Intelligence Ethics. *IEEE Transactions on Artificial Intelligence*, 1-21. DOI <https://doi.org/10.1109/taai.2022.3194503>
7. Johnson, D.G., & Verdicchio, M. (2023). Ethical AI is not about AI. *Communications of the ACM*, 66(2), 32 – 34. DOI <https://doi.org/10.1145/3576932>
8. McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4), 12 – 12.
9. Moor, J.H. (1985). What is Computer Ethics? *Metaphilosophy*, 16(4), 266 – 275. DOI <https://doi.org/10.1111/j.1467-9973.1985.tb00173.x>
10. Parker, D.B. (1968). Rules of ethics in information processing. *Communications of the ACM*, 11(3), 198 – 201. DOI <https://doi.org/10.1145/362929.362987>
11. Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417 – 424. DOI <https://doi.org/10.1017/s0140525x00005756>
12. The Z3. Retrieved from https://web.archive.org/web/20060211035420/http://irb.cs.tu-berlin.de/~zuse/Konrad_Zuse/en/Rechner_Z3.html
13. Weizenbaum, J. (1976). *Computer power and human reason: from judgment to calculation*. Freeman.
14. Wiener, N. (1950). *The human use of human beings: Cybernetics and society*. Houghton Mifflin.
15. Wiener, N. (2019). *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press.
16. Платон (2008). Діалоги. Фоліо.

Надійшла (received) 26.03.2025

Відомості про авторів / About the Authors

Горбань Радомір Ігорович (Radomyr Horban) – Національний технічний університет «Харківський політехнічний інститут», аспірант кафедри філософії, Харків, Україна; ORCID: <https://orcid.org/0009-0002-5660-2038>

Тараросє Яків Володимирович (Jakov Tararouev) – Національний технічний університет «Харківський політехнічний інститут», доктор філософських наук, професор, професор кафедри філософії; Харків, Україна; ORCID: <https://orcid.org/0000-0002-0734-0804>