

<https://doi.org/10.26565/2226-0994-2025-73-12>

УДК: 17.022:004.8

Шеїн Ілля

ФІЛОСОФСЬКИЙ АНАЛІЗ СПІВВІДНОШЕННЯ ЛЮДСЬКОЇ ТА МАШИННОЇ МОРАЛІ В КОНТЕКСТІ ЕТИЧНИХ ВИКЛИКІВ ІШТУЧНОГО ІНТЕЛЕКТУ

У статті здійснено комплексний аналіз проблеми співвідношення людської та машинної моралі в контексті стрімкого розвитку й поширення технологій штучного інтелекту (ШІ). Розглянуто етичні виклики, що постають перед сучасним суспільством у зв'язку з появою алгоритмів, здатних ухвалювати рішення, які мають моральні наслідки, й окреслено потенційну роль машин як моральних агентів.

Здійснено порівняльний огляд провідних закордонних концепцій машинної етики, що охоплюють класичні роздуми щодо тотального характеру технологій і сучасні дослідження, де увагу зосереджено на питанні морального статусу штучного інтелекту, машинний інтенціональності, ризиках самовдосконалення алгоритмів і необхідності запровадження механізмів етичного регулювання. Наголошено на неможливості машинної моралі через відсутність у засобів ШІ суб'єктивних переживань.

Висвітлено особливості українського академічного дискурсу, що поєднує нормативно-прикладний підхід до етики ШІ з урахуванням фундаментальних філософських категорій, зокрема суб'єктності, свідомості та відповідальності. Здійснений аналіз дає можливість дійти висновків, що вітчизняні науковці зосереджені як на нормативній та інституційній етиці, зокрема практичних аспектах застосування ШІ, розробленні етичних стандартів, інституційних механізмів контролю й підвищенні етичної грамотності користувачів, так і екзистенційних викликах.

Здійснено спробу окреслити перспективи інтегрування класичних і сучасних філософських підходів до машинної етики із практичними вимогами національного регулювання ШІ, визначити теоретичні орієнтири й актуальні завдання для подальших досліджень у галузі машинної етики.

Обґрунтовано значущість створення національної моделі етики ШІ, визначено й схарактеризовано основні виміри. Описано практичний аспект філософських основ національної моделі ШІ, визначено завдання та наголошено на важливості поєднати міжнародний досвід з локальними культурно-етичними особливостями.

Ключові слова: *штучний інтелект (ШІ), машинна етика, мораль, філософія техніки.*

Стрімкий розвиток засобів штучного інтелекту (далі – ШІ) не тільки змінює технологічний ландшафт сучасності, але й ставить перед традиційною філософською думкою низку безпрецедентних викликів. Якщо раніше роздуми щодо машин, здатних мислити, належали до царини наукової фантастики, то зараз суспільство стикається з інтелектуальними алгоритмами, що можуть ухвалювати рішення, які мають прямі моральні наслідки для людства. На перетині філософії, етики й інформатики формується поняття машинної етики. Через це співвідношення людської та машинної моралі у призмі етичних викликів постає однією з ключових проблем сучасного філософського дискурсу.

Актуальність цього дослідження зумовлена низкою чинників. По-перше, поширення технологій ШІ у галузях освіти, науки, державного управління, економіки й культури призводить до безпрецедентних ситуацій, коли машини починають ухвалювати рішення, що мають етичні наслідки. По-друге, у сучасній філософській науці відсутня єдність щодо питання про можливість і межі машинної моралі: одні дослідники вбачають у ШІ загрозу гуманістичним цінностям, інші трактують його як інструмент розширення пізнавальних можливостей людини. По-третє, на тлі глобальних ініціатив зі створення етичних стандартів використання ШІ українська наука стикається з потребою розробити

власну нормативно-філософську модель, що врахує і національну специфіку, і міжнародний досвід.

Наукова новизна роботи полягає у спробі здійснити комплексний аналіз проблеми співвідношення людської та машинної моралі у призмі етичних викликів ІІІ через порівняння провідних закордонних і вітчизняних філософських концепцій. У статті розглядається еволюція філософського осмислення ІІІ й окреслюються перспективні напрями дослідження, зокрема потенціал створення національної моделі етики.

Мета розвідки полягає в тому, щоб виявити філософські підвалини й сучасні напрями осмислення визначеної проблеми та визначити особливості національного підходу щодо етичного регулювання. Відповідно до поставленої мети, дослідження має такі *завдання*: розглянути провідні західні філософські ідеї, що сформували основу актуального розуміння техніки, визначити основні напрями розвитку машинної етики, окреслити специфіку українського академічного підходу до питання етичного використання ІІІ й виявити теоретичні орієнтири для подальших наукових розвідок проблематики.

Джерела сучасної проблеми співвідношення людської та машинної моралі лежать у класичному філософському осмисленні техніки, яке висвітлювало сутність технологій та їх впливу на людське буття задовго до виникнення й поширення ІІІ. Однією з найбільш впливових у цьому дискурсі є позиція представника німецького екзистенціалізму М. Гайдеггера. Його провідною думкою є переосмислення техніки як явища: інструментальний погляд (сприйняття техніки як засобу досягнення цілей) еволюціонує до особливого способу розкриття буття, згідно з яким людина розглядає все суще як готові до використання ресурси [Heidegger, 2000, S. 7]. Концепція М. Гайдеггера закладає фундамент для критики машинної моралі, адже інструмент, який перетворює етичні дилеми на алгоритмічний розрахунок ресурсів, як-от життів, ризиків, користі тощо, за своєю суттю є аморальним, тому що ним керує виключно раціональна експлуатація.

Концепція М. Гайдеггера зазнала глибшого осмислення в роботах французького філософа Ж. Еллюля. У своїй праці «Технологічне суспільство» науковець висунув тезу, що техніка є тотальною, автономною й універсальною силою, якій підпорядковані всі сфери людського життя, зокрема мораль. Ця сила, зі свого боку, діє згідно з власною неухильною логікою абсолютної ефективності [Ellul, 1964, с. 33]. Можна дійти висновку, що філософ трактував техніку як соціологічний феномен, що є сукупністю раціональних і ефективних у будь-якій сфері людської діяльності методів. Слід звернути увагу, що стандартизація людської поведінки є не менш важливим застереженням Ж. Еллюля: на його думку, техніка не може залежати від ірраціональних і непередбачуваних факторів, якими є людина та свідомий вибір. Для усунення небажаної мінливості техніка спроможна призвести до того, що людина втрачатиме здатність до вибору, сприйняття моралі й щастя стане симульованим або запрограмованим, а його основною функцією стане забезпечення інтеграції до технічних процесів.

Дослідження М. Гайдеггера та Ж. Еллюля зосереджені на метафізичному й соціологічному тоталітаризмі техніки. Проте не можна не звернутися до англійського математика А. Тюрінга, який змістив філософський акцент до сфери теорій обчислень, поставивши фундаментальне запитання: «Чи може машина мислити?» [Turing, 1950, р. 433]. Учений запровадив концепцію, згідно з котрою машину можна вважати інтелектуальною, якщо вона може імітувати людську поведінку в письмовому діалозі таким чином, що умовний співбесідник нездатний розрізнити відповіді машини та людини. Отже, проблема внутрішнього інтелекту й свідомості опинилася в площині зовнішньої функціональності.

Прагнення А. Тюрінга розв'язати проблему машинного мислення в суто технічному дискурсі породило чимало питань, що й зараз потребують глибокого філософського осмислення. Наприклад, чи може система, що тільки симулює розуміння, бути наділеною справжньою мораллю? Якщо моральний вибір та відповідальність

людини певною мірою пов'язані зі здатністю відчувати страждання, тоді чи може машина, що не зазнає таких страждань, бути моральним об'єктом? Не менш важливо з'ясувати, чи може дія машини бути морально значущою, якщо машина нездатна усвідомлювати сутність моральних концепцій, як-от справедливість, гідність тощо, а тільки маніпулює ними як символами.

Традиційні філософські підходи заклали підґрунтя для сучасного осмислення морального статусу технологій. Вони окреслили подальші дослідження щодо інтелектуальних можливостей машин і пошуку етичних меж їх використання. Відповідно, сучасна етика ІІІ прагне подолати розрив між інструментальним і гуманістичним вимірами технологій та запропонувати моделі відповідального використання ІІІ.

Зокрема сучасний дослідник Д. Андерсон вивчає поняття «машинна інтенціональність» і пропонує розглянути, яким чином складники машинного розуму здатні породжувати наміри. Філософ наголошує на важливості з'ясувати моральний статус машини, тобто чи може вона відігравати роль об'єкта етичної уваги. Зі свого боку, встановлення морального статусу умовної машини є кроком до розуміння, чи може вона бути повноцінним інтелектом [Anderson, 2012, р. 331]. Концепція вченого змінює традиційний підхід до оцінювання діяльності ІІІ. Замість першочергового аналізу когнітивних здібностей і здатності мати наміри запропоновано зворотну логіку: первинним фактором має стати визначення наявності чи відсутності морального статусу. Іншими словами, акцент із функціонального визначення ІІІ зміщується до антропологічно-етичного виміру, оскільки інтенціональність не може бути розглянута окремо від морального контексту.

У цьому контексті доцільно розглянути концепти щодо ризиків формування машинної етики. Сучасний науковець Р. Ямпольський наголошує, що спроби дозволити машинам ухвалювати етичні рішення є докорінно хибними, оскільки неможливо довести, чи може бути система ІІІ безпечною в умовах самовдосконалення [Yampolskiy, 2013, р. 390]. Учений акцентує на важливості мінімізувати ризики ще на етапі розроблення засобів ІІІ. Йдеться про запровадження таких обмежень, які були б первинною інженерною вимогою, залишалися сталим активним механізмом протягом тривалого часу та звели вірогідність завдання шкоди потенційному користувачеві до нульової. Водночас досягнення етичної досконалості ІІІ не є предметом фокуса [Ibid., р. 391]. Іншою тезою Р. Ямпольського є критичне осмислення сучасного філософського трактування царини машинної етики: на думку науковця, наявні провідні дослідження є здебільшого суто теоретичними та не сприяють практичному розв'язанню питання безпеки.

Значну увагу філософському осмисленню феномена ІІІ приділяє М. Боден. Позиція дослідниці виходить із тези, що ІІІ – не тільки інструмент, але й потужне джерело для набуття нових знань [Boden, 2018]. Учена наголошує, що ІІІ дедалі більше охоплює різні сфери людської діяльності, тому недоцільно обмежуватися суто технічним осмисленням цього явища. Натомість не менш важливим питанням постає критична дилема: чи справді здатна машина мислити, бути свідомою й проявляти креативність?

Якщо припустити, що сутність людської моралі нерозривно пов'язана зі свідомістю, творчістю та суб'єктивним досвідом, філософський аспект ІІІ стає критичним: думка М. Боден спонукає до потреби глибокого розуміння, чи може ІІІ бути моральним суб'єктом. Не менш значущим постає і аспект безпеки: якщо Р. Ямпольський наголошує на необхідності певних обмежень використання технологій ІІІ через їхню непередбачуваність, М. Боден прагне пояснити джерело цієї непередбачуваності.

Можна констатувати, що висновки науковця не тільки акцентують на морально-етичному осмисленні використання ІІІ, але і відкривають ширше поле для досліджень. Нагальною проблемою залишається те, яким чином філософські ідеї та концепти можуть бути вбудовані в інженерні протоколи безпеки й етики технологій ІІІ. Варто зауважити, що М. Боден, зосереджуючись на поміркованих формах ІІІ, не враховує радикальнішого

сценарію розвитку технологій у контексті самовдосконалення та гіпотетичної появи «суперінтелекту», про який згадує Р. Ямпольський.

Філософську проблему етики ІІІ порушує британський філософ Б. Шталь. Науковець зосереджується на важливості етичних аспектів розвитку й застосування ІІІ та наголошує, що технологія має ризики, пов'язані з упередженнями та порушенням приватності [Stahl, 2021, p. 19]. Зокрема, він слушно згадує важливість широкої репрезентативності інформаційних масивів, що застосовують для навчання, перевірки й тестування засобів ІІІ, але справедливо висловлює сумнів, чи можуть ці вимоги до якості бути повною мірою виконані за умови використання реальних [Ibid., 2021, p. 17].

На думку науковця, міжнародні й національні структури відіграють особливо важливу роль у запровадженні правил, що забезпечують безпеку застосування технологій ІІІ. Окрім цього, Б. Шталь наголошує на принципі відкритих алгоритмів, звітності та відповідальності для запобігання зловживань та ухваленні неправильних рішень [Ibid., 2021, p. 17]. Можна стверджувати, що концепція Б. Штала зосереджена на технічному аспекті розв'язання морально-етичних проблем. Учений порушує філософську за своєю суттю проблему, але запропонованому розв'язанню бракує глибинного аналізу сутнісних питань, зокрема питання свідомості та суб'єктності машини й природи морального вибору, яка розмежовує алгоритмічну калькуляцію та унікальність людської моралі, що ставить нові виклики перед сучасною наукою.

Суттєвий внесок у дослідження етики техніки й зокрема ІІІ зробив італійський філософ А. Флоріді. Учений наголошує на критичній важливості розроблення загальносвітових етичних меж використання ІІІ з метою уникнення негативних наслідків [Floridi, 2019, p. 6]. А. Флоріді слушно зауважує, що існує велика кількість ініціатив щодо етики ІІІ, серед яких важливо виробити єдиний підхід. Важливим нюансом постає культурне різноманіття, оскільки моральний світогляд і культурні перспективи у різних частинах світу відрізняються. Філософ наголошує, що єдина система етичних принципів може позитивно вплинути на розроблення стандартів щодо використання ІІІ [Ibid., 2019, p. 14].

Цінність висновків А. Флоріді полягає в конструктивній відповіді на тези щодо тоталітаризму й небезпеки технологій ІІІ. Слушною є думка про те, що суспільні інститути можуть і повинні забезпечити регулювання ІІІ, розробивши міжкультурні етичні стандарти. Ця концепція формує чимало перспективних, наразі неповною мірою розглянутих напрямів дослідження.

Зокрема, доцільно проаналізувати, яким чином абстрактні моральні принципи можуть бути перетворені на технічні вимоги й алгоритмічні протоколи. Не менш важливою проблемою є розроблення юридичних і технічних інструментів для визначення відповідальності у випадках завдання засобами ІІІ моральної шкоди. Водночас постає дилема щодо того, хто саме має бути провідною ланкою в аспекті відповідальності: розробник, користувач чи безпосередньо алгоритм.

З іншого боку, розроблення універсальних норм використання ІІІ породжує проблему: якою мірою етика ІІІ має перспективу перетворитися на принцип всесвітнього управління? А. Флоріді згадує, що моральні цінності глибоко вкорінені в культуру, проте масове підпорядкування глобальним нормам несе ризик нав'язування однієї, домінантної морально-етичної парадигми іншим. Отже, є підстави вважати, що концепт етики ІІІ постає однією з найскладніших геополітичних і правових проблем сучасності.

У результаті аналізу позицій учених в контексті можливості машинної моралі як такої переконливим видається аргумент щодо неможливості машини відчувати страждання. Це дає підстави стверджувати, що мораль у своєму антропологічному сенсі неможлива, коли йдеться про машинний вимір, адже, скажімо, категорію свідомості та суб'єктного переживання неможливо відтворити за допомогою алгоритмів. Проте рішення ІІІ можуть формувати морально-етичні наслідки для ситуацій у реальному житті, що ставить перед розробниками необхідність створити нормативно-прикладні регулятори.

Аналіз ключових закордонних філософських позицій щодо етичних концептів ІІІ окреслює суперечливий ландшафт проблеми машинної моралі. Втім, для повноти картини й з огляду на актуальність проблеми не менш важливо звернутися до вітчизняної наукової думки. Українські дослідження в цій галузі мають свою специфіку, оскільки певною мірою зосереджені на особливостях морально-етичних норм практичного застосування ІІІ в різних сферах людської діяльності. Вочевидь такий акцент зумовлений потребою інтегрувати нову технологію в умовах національної специфіки з урахуванням наявних правових, соціальних і економічних структур.

Системні спроби осмислення технологій в українському академічному дискурсі спостерігаємо ще з початку 2000-х років. Уже тоді В. Штанько констатує факт появи нового культурного простору, «основним матеріальним носієм текстів якої є не писемність, а “екранність” – “екранна культура”» [Штанько, 2002, с. 20]. Дослідниця розглядає явище «віртуальна реальність» і вдається до питання, як може свідомість адаптуватися до перебування в кіберпросторі. Водночас вона порушує морально-етичний вимір проблеми в контексті потенційних маніпуляцій і зауважує, що «занурення людини у світ віртуальної реальності дозволяє нав'язувати їй певні способи існування» [Ibid., 2002, с. 27].

Водночас О. Ральчук, оцінюючи добу інформаційного суспільства, окреслив потенційні ризики виникнення ІІІ та зауважив, що «сучасна дійсність виявляє <...> прагнення створити вже не тільки штучний інтелект, а й саму штучну людину замість людини природної» [Ральчук, 2003]. Окрім цього, ще задовго до масового поширення технологій ІІІ В. Штанько та Т. Бордюгова зазначили, що важливо з'ясувати, хто є морально відповідальним за рішення, які ухвалюють інформаційно-комп'ютерні системи [Штанько, 2012, с. 151]. Водночас автори зауважують, що «вивчення етичних аспектів може й має сприяти подоланню проблем та знаходженню усталених рішень» [Ibid., 2012, с. 152].

В умовах масового використання ІІІ це питання набуває принципово іншого масштабу. Можна констатувати, що філософське осмислення екзистенційних і морально-етичних ризиків ІІІ 2000-х років значною мірою передбачило реальність, з якою стикнулося сучасне суспільство. Проте наразі це дослідження набуває ще більшої актуальності. Зокрема Ю. Мельник, С. Тодорова та Г. Шевченко окреслюють проблеми філософського визначення природи ІІІ й наголошують, що особливу увагу доцільно приділити можливості втрати контролю над його розвитком [Мельник, 2025, с. 7]. Дослідження І. Дятлової, О. Патлайчук і О. Ступак акцентують на питанні відповідальності у випадках, коли система ІІІ ухвалює рішення без людського втручання й наголошують на проблемі алгоритмічного упередження ІІІ під час навчання на основі історичних даних [Дятлова, 2025, с. 70]. Н. Поліщук, Г. Поперечна й К. Ніколенко досліджують етичні проблеми взаємин людини й ІІІ у метавсесвіті, що пов'язані з розмиттям меж реальності, маніпулятивністю алгоритмів і, як наслідок, кризою цифрової ідентичності [Поліщук, 2025].

Окрім цього, можна спостерігати, що сучасний філософський дискурс певною мірою набуває інституційного й прикладного характеру. Українські дослідники зазначають, що «існує пробіл між теоретичними дослідженнями філософів і практичною діяльністю розробників у галузі штучного інтелекту» [Добровольська, 2019, с. 15] та пояснюють це необхідністю зосередитись на розв'язанні актуальних проблем. Через це існує сталий інтерес до морально-етичних норм використання ІІІ в конкретних галузях.

Наприклад, дослідження О. Пінчук та І. Малицької стосуються етичних викликів ІІІ в царині науки. Вчені зауважують, що розвиток ІІІ вимагає правової бази, а редакційні політики наукових видань мають адаптуватися до впровадження етичних стандартів і академічної доброчесності, що пов'язані з особливостями рецензування [Пінчук, 2024, с. 192]. Потреба в освітніх заходах щодо доброчесного використання ІІІ, як зауважують автори, свідчить про необхідність підвищення етичної свідомості користувачів.

Цю думку підтримує Ю. Трач, яка зазначає, що «необхідно систематично докласти зусиль для підвищення грамотності, обізнаності та освіти щодо етичних наслідків застосування технологій ІІІ» [Трач, 2024, с. 306]. Увагу питанню етики ІІІ в царині наукових досліджень приділяє і Л. Хоружа, наголошуючи, що «актуалізується важливість відповідального використання ІІІ, що пов'язано з розробкою нових стандартів та норм академічної етики» [Хоружа, 2025, с. 10]. Можна зробити висновок, що інституційна етика використання ІІІ набуває дедалі більшої ваги.

Ю. Карпенко акцентує на етиці ІІІ в галузі публічного управління, що порушує морально-етичну дилему: чи можна трактувати ІІІ як спосіб організувати людську діяльність. Думка науковця щодо розроблення єдиних етичних норм щодо використання новітніх цифрових технологій має безпосередні паралелі з позицією закордонних учених [Карпенко, 2019, с. 96].

В. Воронкова й В. Нікітенко детально розглядають філософію ІІІ в економічному вимірі й окреслюють так звані «хвилі розвитку» [Воронкова, 2022, с. 247]. На основі дослідження вчених можна дійти висновку щодо актуальності створення наднаціональних інститутів контролю над ІІІ для запобігання системній шкоді економіці й суспільству, хоча питання щодо об'єктивності та прозорості запровадження таких механізмів залишається відкритим.

Аналізуючи український філософський дискурс, можна констатувати, що вітчизняні науковці приділяють суттєву увагу нормативній та інституційній етиці. Паралельно є і сталий інтерес до екзистенційних ризиків. Втім, оскільки світова наукова спільнота приділяє дедалі більше уваги регулюванню технологій, ще одним перспективним напрямом є філософський аспект розроблення національної моделі етики ІІІ, котра була б здатна поєднати як міжнародний досвід, так і локальні особливості.

Логічно припустити, що поняття «національна модель етики ІІІ» можна окреслити в законодавчому, технологічному й дослідницькому вимірах, де філософський аспект відіграв би одну з провідних ролей. Зокрема законодавчий рівень може визначити моральні принципи використання ІІІ в державній сфері. Поміж іншого, вони стосуються правничої галузі й питання, наскільки етично застосовувати ІІІ під час ухвалення юридичних рішень.

Технологічний рівень має на меті застосувати морально-етичні норми безпосередньо в алгоритмах ІІІ. Водночас зміст таких норм і їх відповідність національним особливостям є предметом подальших досліджень. Не менш важливим складником моделі є освітній аспект, який стосується навчання щодо етичних ризиків, застосування та обмежень ІІІ. Зі свого боку, дослідницький вимір передбачає міждисциплінарну співпрацю, під час якої філософське розуміння й межі моральних норм є важливим фактором для розроблення етичних алгоритмів. Не менш суттєвими питаннями є взаємодія людини та ІІІ в ситуаціях, що мають моральну неоднозначність, та вплив рішень ІІІ на остаточне ухвалення рішень як у віртуальному, так і реальному середовищі.

У практичному контексті філософські підвалини національної моделі етики ІІІ полягають у систематизуванні й концептуальному окресленні морально-етичних принципів, які відтворюють українські культурні, соціальні та правові норми з урахуванням міжнародного досвіду. Окрім цього, важливу роль відіграє посилення методологічної бази, що стосується моделі аналізу моральних дилем і конфліктів, що можуть виникати під час роботи людини з засобами ІІІ.

Можна підсумувати, що філософське осмислення проблеми людської та машинної моралі в контексті ІІІ сформувало суперечливий світовий консенсус. Кінцевий успіх у розв'язанні цієї проблеми залежить від можливості перетворити наявні філософські застереження на дієві правові інструменти. Для цього перспективним шляхом є розроблення національної моделі етики, що визначатиме нормативні орієнтири та

гармонійно поєднуватиме міжнародні стандарти етичних норм з локальними культурними особливостями.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Воронкова В. Г., Нікітенко В. О. *Філософія цифрової людини і цифрового суспільства: теорія і практика*. Львів-Торунь: Liha-Pres, 2022. 460 с. Електронний ресурс. URL: <https://dspace.znu.edu.ua/xmlui/handle/12345/6591>

Добровольська О. В., Штанько В. І. Філософський аналіз еволюції штучного інтелекту. *Дослідження з історії і філософії науки і техніки*. 2019. Вип. 28, № 1. С. 10–19. <https://doi.org/10.15421/271902>

Дятлова, І., Патлайчук, О., & Ступак, О. Етичні виклики та перспективи філософії штучного інтелекту у ХХІ столітті. *Науково-теоретичний альманах Грани*. 2025. Т. 28. № 2. С. 66–71. <https://doi.org/10.15421/172540>

Карпенко Ю. В. Етичні принципи застосування штучного інтелекту в публічному управлінні. *Вісник Національної академії державного управління при Президентові України. Серія: Державне управління*. 2019. № 4. С. 93–97. Електронний ресурс. URL: http://nbuv.gov.ua/UJRN/vnaddy_2019_4_15

Мельник Ю., Тодорова С., Шевченко Г. Філософське розуміння штучного інтелекту як технології. *Philosophy and Governance*. 2025. № 3 (7). DOI <https://doi.org/10.70651/3041-248X/2025.3.01>

Пінчук О., Малицька І. Відповідальне та етичне використання штучного інтелекту в дослідницькій і публікаційній діяльності. *Інформаційні технології і засоби навчання*. 2024. Т. 100, № 2. С. 45–57. <https://doi.org/10.33407/itlt.v100i2.5676>

Поліщук Н. В., Поперечна Г. А., Ніколенко К. В. Етичні дилеми взаємодії людини та штучного інтелекту у віртуальних просторах метавсесвіту. *Вісник гуманітарних наук*. 2025. № 8. <https://doi.org/10.5281/zenodo.15616278>

Ральчук О. Інформаційне суспільство: між ейфорією спокус та законами універсуму. *Вісник НАН України*. 2003. № 2. С. 36–50. Електронний ресурс. URL: http://nbuv.gov.ua/UJRN/vnanu_2003_2_7

Трач Ю. Дискурс навколо етики штучного інтелекту: особливості формування та інституалізації. *Цифрова платформа: інформаційні технології в соціокультурній сфері*. 2024. Т. 7, № 2. С. 299–310. <https://doi.org/10.31866/2617-796X.7.2.2024.317738>

Хоружа Л. Наукова етика й доброчесність в обіймах штучного інтелекту. *Педагогічна освіта: Теорія і практика. Психологія. Педагогіка*. 2025. Т. 43, № 1. С. 6–12. <https://doi.org/10.28925/2311-2409.2025.431>

Штанько В. І. Інформаційні технології та проблеми трансформації сучасної культури. *Культура України: зб. наук. праць. Мистецтвознавство. Філософія*. Харків : ХДАК, 2002. Вип. 10. 222 с.

Штанько В. І., Бордюгова Т. Г. Інформаційне суспільство: соціально-філософські проблеми становлення: Навч. посібник. Харків: ХНУРЕ, 2012. 172 с.

Anderson D. L. Machine Intentionality, the Moral Status of Machines, and the Composition Problem. *The Philosophy & Theory of Artificial Intelligence* / ed. by V. C. Müller. Berlin: Springer, 2012. P. 312–333. URL: <https://philarchive.org/rec/ANDMIT-3>

Boden M. A. *Artificial Intelligence: A Very Short Introduction*. Oxford: Oxford University Press, 2018. 152 p. URL: <https://academic.oup.com/book/415>

Ellul J. *The Technological Society*. New York: Alfred A. Knopf, 1964. 512 p. URL: <https://voidnetwork.gr/wp-content/uploads/2021/09/The-Technological-Society-Jacques-Ellul.pdf>

Floridi L. Ethics, Governance, and Policies in Artificial Intelligence / ed. by L. Floridi. Cham: Springer, 2019. *Philosophical Studies Series*, Vol. 144. 312 p. Електронний ресурс. URL: <https://content.e-bookshelf.de/media/reading/L-16871700-5ce6df8467.pdf>

Heidegger M. Die Frage nach der Technik. *Gesamtausgabe I. Abteilung: Veröffentlichte Schriften 1910-1976. Band 7. Vorträge und Aufsätze*. Frankfurt am Main: Vittorio Klostermann, 2000. S. 5-36.

URL: https://monoskop.org/images/2/27/Heidegger_Martin_1953_2000_Die_Frage_nach_der_Technik.pdf

Stahl B. C., Schroeder D., Rodrigues R. *Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenges*. Cham: Springer, 2021. 220 p. URL: <https://philpapers.org/rec/STAEOA-5>, DOI 10.1007/978-3-031-17040-9

Turing A. M. Computing Machinery and Intelligence. *Mind*. 1950. V. 49. P. 433–460. URL: <https://courses.cs.umbc.edu/471/papers/turing.pdf>

Yampolskiy R. V. Artificial Intelligence Safety Engineering: Why Machine Ethics Is a Wrong Approach. *Philosophy and Theory of Artificial Intelligence* / ed. by V. C. Müller. Berlin: Springer, 2013. P. 389–396. URL: <http://cecs.louisville.edu/ry/AIsafety.pdf>, DOI 10.1007/978-3-642-31674-6_29.

Шейн Ілля Олександрович

аспірант кафедри 701

Національний аерокосмічний університет «ХАІ»

м. Харків, вул. Вадима Манька, 17, Україна

E-mail: i.o.shein@khai.edu

ORCID: <https://orcid.org/0000-0002-1850-8529>

Стаття надійшла до редакції 01.08.2025

Стаття рекомендована до друку 25.10.2025

Переглянуто 20.10.2025

Опубліковано 24.11.2025

PHILOSOPHICAL ANALYSIS OF THE RELATIONSHIP BETWEEN HUMAN AND MACHINE MORALITY IN THE CONTEXT OF ETHICAL CHALLENGES OF ARTIFICIAL INTELLIGENCE

Shein Illia O.

Postgraduate student of the Department 701

National Aerospace University “KhAI”

17 Vadym Manko Street, Kharkiv, Ukraine

E-mail: i.o.shein@khai.edu

ORCID: <https://orcid.org/0000-0002-1850-8529>

ABSTRACT

The article deals with a comprehensive analysis of the problem of the correlation between human and machine morality in the context of the rapid development and spread of artificial intelligence technologies. It addresses the ethical challenges facing modern society regarding the emergence of algorithms capable of making decisions that have moral consequences, and outlines the potential role of machines as moral agents.

The article provides a comparative review of leading foreign concepts of machine ethics, covering classical reflections on the total nature of technologies and contemporary research focusing on the artificial intelligence moral status, machine intentionality, the risks of algorithm self-improvement, and the need to introduce ethical regulation mechanisms. There is an emphasis on the impossibility of machine morality due to the lack of subjective experiences in AI.

The paper highlights the peculiarities of Ukrainian academic discourse, which combines a normative-applied approach to artificial intelligence ethics with fundamental philosophical categories, including subjectivity, consciousness, and responsibility. The analysis allows one to conclude that national scientists focus mainly on the practical aspects of the artificial intelligence application, the development of ethical standards, institutional control mechanisms, and the improvement of users' ethical literacy.

There is an attempt to outline the prospects for integrating classical and modern philosophical approaches to machine ethics with the practical requirements of the national regulation of artificial intelligence, to identify theoretical guidelines and current tasks for further research in the field of machine ethics.

The article proves the importance of creating a national model of artificial intelligence ethics, identifies and characterizes its main components. It describes the practical application of the philosophical foundations of the national model of artificial intelligence, defines tasks, and emphasizes the importance of combining international experience with local cultural and ethical characteristics.

Keywords: *artificial intelligence, machine ethics, morality, philosophy of technology.*

REFERENCES

- Voronkova, V. H., & Nikitenko, V. O. (2022). Philosophy of the digital human and digital society: Theory and practice. Lviv-Torun: Liha-Pres. Retrieved from <https://dspace.znu.edu.ua/xmlui/handle/12345/6591> (In Ukrainian)
- Dobrovolska, O. V., & Shtanko, V. I. (2019). Philosophical analysis of artificial intelligence evolution. *Studies in history and philosophy of science and technology*, 28(1), 10–19. <https://doi.org/10.15421/271902> (In Ukrainian)
- Diatlova, I., Patlaichuk, O., & Stupak, O. (2025). Ethical Challenges and Prospects for the Philosophy of Artificial Intelligence in the XXI century. *Scientific and Theoretical Almanac Grani*, 28(2), 66–71. <https://doi.org/10.15421/172540> (In Ukrainian)
- Karpenko, Y. V. (2019). Ethical principles of artificial intelligence application in public administration. Bulletin of the National Academy of Public Administration under the President of Ukraine. Series: Public Administration, (4), 93–97. Retrieved from http://nbuv.gov.ua/UJRN/vnaddy_2019_4_15 (In Ukrainian)
- Melnyk, Y., Todorova, S., & Shevchenko, H. (2025). Philosophical understanding of artificial intelligence as technology. *Philosophy and Governance*, 3(7). <https://doi.org/10.70651/3041-248X/2025.3.01> (In Ukrainian)
- Pinchuk, O., & Malytska, I. (2024). Responsible and ethical use of artificial intelligence in research and publishing activities. *Information Technologies and Learning Tools*, 100(2), 45–57. <https://doi.org/10.33407/itlt.v100i2.5676> (In Ukrainian)
- Polishchuk, N., Poperechna, G., & Nikolenko, K. (2025). Ethical dilemmas of human–artificial intelligence interaction in the virtual spaces of the metaverse. *Bulletin of Humanities*, (8). <https://doi.org/10.5281/zenodo.15616278> (In Ukrainian)
- Ralchuk, O. (2003). Information society: Between the euphoria of temptations and the laws of the universe. National Academy of Sciences of Ukraine, (2), 36–50. Retrieved from http://nbuv.gov.ua/UJRN/vnanu_2003_2_7 (In Ukrainian)
- Trach, Y. (2024). Discourse on artificial intelligence ethics: Features of formation and institutionalization. *Digital Platform: Information Technologies in the Socio-Cultural Sphere*, 7(2), 299–310. <https://doi.org/10.31866/2617-796X.7.2.2024.317738> (In Ukrainian)
- Khoruzha, L. (2025). Scientific ethics and integrity in the embrace of artificial intelligence. *Pedagogical Education: Theory and Practice. Psychology. Pedagogy*, 43(1), 6–12. <https://doi.org/10.28925/2311-2409.2025.431> (In Ukrainian)
- Shtanko, V. I. (2002). Information technologies and the problems of transformation of modern culture. In Culture of Ukraine: Collection of scientific papers. Art studies. Philosophy (Issue 10, 222 p.). Kharkiv: KhSAC. (In Ukrainian)
- Shtanko, V. I., & Bordyugova, T. H. (2012). Information society: Socio-philosophical problems of formation [Textbook]. Kharkiv: KhNURE. (In Ukrainian)
- Anderson, D. L. (2012). Machine intentionality, the moral status of machines, and the composition problem. In V. C. Müller (Ed.), *The philosophy & theory of artificial intelligence* (pp. 312–333). Berlin: Springer. Retrieved from <https://philarchive.org/rec/ANDMIT-3>
- Boden, M. A. (2018). Artificial intelligence: A very short introduction. Oxford: Oxford University Press. Retrieved from <https://academic.oup.com/book/415>
- Ellul, J. (1964). The technological society. New York: Alfred A. Knopf. Retrieved from <https://voidnetwork.gr/wp-content/uploads/2021/09/The-Technological-Society-Jacques-Ellul.pdf>

- Floridi, L. (Ed.). (2019). Ethics, governance, and policies in artificial intelligence (Philosophical Studies Series, Vol. 144, 312 p.). Cham: Springer. Retrieved from <https://content.e-bookshelf.de/media/reading/L-168717005ce6df8467.pdf>
- Heidegger, M. (2000). Vorträge und Aufsätze: Die Frage nach der Technik. Frankfurt am Main: Vittorio Klostermann. Retrieved from https://monoskop.org/images/2/27/Heidegger_Martin_1953_2000_Die_Frage_nach_der_Technik.pdf
- Stahl, B. C., Schroeder, D., & Rodrigues, R. (2021). Ethics of artificial intelligence: Case studies and options for addressing ethical challenges. Cham: Springer. Retrieved from <https://philpapers.org/rec/STAEOA-5>, DOI 10.1007/978-3-031-17040-9
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 49, 433–460. Retrieved from <https://courses.cs.umbc.edu/471/papers/turing.pdf>
- Yampolskiy, R. V. (2013). Artificial intelligence safety engineering: Why machine ethics is a wrong approach. In V. C. Müller (Ed.), *Philosophy and theory of artificial intelligence* (pp. 389–396). Berlin: Springer. Retrieved from <http://cecs.louisville.edu/ry/AIsafety.pdf>, DOI 10.1007/978-3-642-31674-6_29.

The article was received by the editors 01.08.2025

The article was revised 20.10.2025

The article is recommended for printing 25.10.2025

This article published 24.11.2025

Як цитувати: Шеїн, І. (2025). ФІЛОСОФСЬКИЙ АНАЛІЗ СПІВВІДНОШЕННЯ ЛЮДСЬКОЇ ТА МАШИНОЇ МОРАЛІ В КОНТЕКСТІ ЕТИЧНИХ ВИКЛИКІВ ШТУЧНОГО ІНТЕЛЕКТУ. *Вісник Харківського національного університету імені В. Н. Каразіна. Серія «Філософія. Філософські перипетії»*, (73), 141-150. <https://doi.org/10.26565/2226-0994-2025-73-12>

In cites: Shein, I. (2025). PHILOSOPHICAL ANALYSIS OF THE RELATIONSHIP BETWEEN HUMAN AND MACHINE MORALITY IN THE CONTEXT OF ETHICAL CHALLENGES OF ARTIFICIAL INTELLIGENCE. *The Journal of V. N. Karazin Kharkiv National University, Series Philosophy. Philosophical Peripeteias*, (73), 141-150. <https://doi.org/10.26565/2226-0994-2025-73-12> [In Ukrainian]