



ХАМІТОВ

Назіп Віленович —

член-кореспондент НАН України, доктор філософських наук, професор, провідний науковий співробітник Інституту філософії імені Г.С. Сковороди НАН України, завідувач відділу досліджень наукових і освітянських методологій та практик Центру гуманітарної освіти НАН України

ВЗАЄМОДІЯ ЛЮДИНИ І ШТУЧНОГО ІНТЕЛЕКТУ: МЕТОДОЛОГІЯ ОСМИСЛЕННЯ ВИКЛИКІВ І ПЕРСПЕКТИВ

*Проблема штучного інтелекту гранично актуалізується під час нинішньої екзистенціальної війни в Україні. В ній штучний інтелект не лише стає засобом посилення ефективності на полі бою, а й використовується в когнітивному та світоглядному вимірах війни, трансформуючи ідентичність, цінності та смисли. При цьому штучний інтелект системно впливає на людину і в мирному просторі та часі, зокрема в науці та освіті, для яких виклик чи не найгостріший — ідеться про чат-боти зі штучним інтелектом, здатні генерувати не лише тези й статті, а й дисертації і монографії, спокушаючи студентів і аспірантів. У статті запропоновано нову методологію осмислення проблеми — **метаантропологію штучного інтелекту**. В її координатах акцент зроблено на взаємодії людини і штучного інтелекту та еволюції цієї взаємодії, що може привести до появи суб'єкта сильного штучного інтелекту, який дорівнює людському і буде не лише ефективним, а й безпечним, більше того — сприятиме гуманістичному розвитку особистості та суспільства. Запропоновано дві парадигми відповіді на деструктивні наслідки використання штучного інтелекту силами самого штучного інтелекту.*

Ключові слова: людина, штучний інтелект, сильний штучний інтелект, суб'єкт природного інтелекту, суб'єкт штучного інтелекту, світогляд, методологія, метаантропологія, метаантропологія штучного інтелекту.

Багато придумано для того, щоб не думати.

Карел Чапек

Найважливішим для розуміння конструкції машин є пізнання істот, що їх збудували.

Станіслав Лем

Проблема не в тому, що штучний інтелект може нас ненавидіти, а в тому, що йому може бути байдуже до нас.

Сем Альтман

Постановка проблеми. Штучний інтелект — це дар, випробування і спокуса для носія природного інтелекту. Він породжує безліч викликів. Штучний інтелект може стати каталізатором еволюції носія природного інтелекту, а може завести його в

глухий кут еволюції. Штучний інтелект — це чарівне дзеркало, в якому відображаються як вади, так і досягнення людини.

Проблема штучного інтелекту гранично актуалізується в часи війни; це яскраво і трагічно засвідчує нинішня екзистенціальна війна в Україні. У ній штучний інтелект постає феноменом, який робить ефективними й самостійними вбивчі пристрої, і є виражена тенденція до розширеного й поглибленого його використання. Не слід забувати, що війна в Україні — це не лише війна на полі бою з дронами, ракетами і протидією їм, це й *когнітивна та світоглядна війна* — війна за волю до перемоги, за ідентичність, цінності та смисли. На жаль, штучний інтелект як інструмент посилення руйнівного впливу на людину все більше використовують у всіх вимірах сучасної війни.

Однак сьогодні штучний інтелект не меншою мірою впливає на людину і в мирному просторі та часі. Це майже всі сфери буття сучасної людини: політика, економіка, соціальна сфера, мистецтво, наука, освіта. Щодо останніх двох сфер — науки й освіти — ситуація є чи не найгострішою: йдеться передусім про чат-боти зі штучним інтелектом, які здатні генерувати тези й статті, а платні версії — цілі дисертації та монографії. Це спокушає аспірантів і навіть молодих вчених *без достатньої волі до пізнання і творчості* відходити від зусиль щодо створення текстів, не кажучи вже про їхні переклади.

Ситуація проста і трагічна: згенеровані штучним інтелектом тексти видають за власні. І якщо раніше була проблема, яку я колись назвав *гугл-мисленням*, — проблема копіювання інформації, знайденої пошуковою системою, з подальшим включенням її у свої тексти, то тепер до неї додається проблема *чат-бот-мислення* — проблема присвоєння завершеного тексту, створеного штучним інтелектом. Мислення тепер зводиться лише до запиту (низки уточнюючих запитів) і подальшого редагування матеріалу, яке з часом також може бути віддано штучному інтелекту.

Це виклик не лише майбутньому науки і освіти, а й *самій природі людини як мислячої істоти*. Такий виклик підсилюється глобальними

соціальними й економічними тенденціями. У сучасному світі «...власникам транснаціональних корпорацій стає вигідніше вкладати гроші не в освіту своїх працівників, а в штучний інтелект. Це стає більш рентабельним і дозволяє швидше отримати економічний прибуток» [1, с. 12]. Маємо тривожну тенденцію сучасності: людський капітал деактуалізується і замінюється *III-капіталом*.

Отже, взаємодія людини і штучного інтелекту є актуальною не лише під час екзистенціальної війни в Україні, а й у процесі майбутнього повоєнного відновлення нашої країни. Водночас це глобальна загальнолюдська проблема, яка в найближчі десятиліття буде лише посилюватися.

* * *

У суспільній свідомості сьогодні маємо два ключові міфи щодо взаємодії людини і штучного інтелекту. Перший, більш поширений — **апокаліптичний**: штучний інтелект є руйнівним для людини і людства, він спочатку забере роботу в мільйонів, а може, й мільярдів людей, сприятиме їхній когнітивній деградації, а потім знищить їх. Другий міф — **гіпероптимістичний**: штучний інтелект вирішить усі проблеми людства. Йдеться про подолання екологічної кризи, участь носіїв штучного інтелекту у війнах замість людей, заміну людей на шкідливих виробництвах, створення андронів, здатних ретельно та професійно турбуватися про літніх людей, до того ж позбавляючи їх від самотності.

Спробуємо оминати ці крайності. Але передусім визнаємо, що штучний інтелект сьогодні створює глибинні виклики людству і людяності. І апокаліптичний міф не випадково є таким поширеним.

Підходи до проблеми у філософській та науковій літературі

Подолання міфів в осмисленні викликів і перспектив штучного інтелекту потребує створення сучасної філософської концепції штучного інтелекту й обґрунтування філософської мето-

дології її розгортання, яка з необхідністю повинна бути міждисциплінарною, — лише тоді вона здатна буде розвинути зазначену концепцію у *міждисциплінарну теорію штучного інтелекту*. Таке завдання в науковій літературі виконується лише частково, що, власне, й зумовило появу цієї статті.

Водночас слід відзначити дослідження окремих авторів, які заклали підґрунтя, і тепер можна спиратися на їхні методологічні надбання й результати.

У глобалізованому вимірі рівень становлення світу, який В.І. Вернадський визначив концептом «ноосфера», все більше стає сферою штучного інтелекту, точніше — сферою взаємодії людей і носіїв штучного інтелекту. Саме тому дослідження відносин людини і штучного інтелекту слід починати з нового осмислення творчої спадщини В.І. Вернадського, який є не лише видатним вченим, а й видатним філософом. Крім того, варто заглибитися у творчу спадщину Г.С. Сковороди, який трактує людину як мікрокосм, що має в собі як зовнішній, так і внутрішній виміри.

Розгляд питання відносин людини і штучного інтелекту, осмислення людини як творця останнього потребує усвідомлення проблеми людини, її природи, духовної і душевної еволюції. Це зумовлює звернення до класиків філософської антропології — М. Шелера, Г. Плеснера, А. Гелена, напрацювань київської світоглядно-антропологічної школи, започаткованої В.І. Шинкаруком, творчості І.В. Бичка, М.О. Булатова, В.П. Іванова, В.Г. Кременя, В.Г. Табачковського та ін.

Осмислення людини як унікальної істоти у Всесвіті, яка здатна до концептуалізації, потребує звернення до праць С.Б. Кримського, М.В. Поповича, В.А. Рижка та ін. Дослідження наявного стану і перспектив відносин людини і штучного інтелекту неможливе без осягнення сучасної наукової картини світу, осмислення соціокультурних чинників наукової творчості, етосу науки, міждисциплінарних зв'язків природничого, соціального і гуманітарного пізнання, що зумовлює необхідність звернення до робіт Л.В. Губерського, І.С. Добронравової,

В.П. Драпогуза, І.В. Єршової-Бабенко, О.Н. Кубальського, В.І. Кузнецова, О.С. Онищенко, С.І. Пирожкова, О.О. Рафальського, Л.В. Рижко, Л.О. Шашкової та ін. Аналітика відносин людини і штучного інтелекту в соціумі в його історичному розвитку та сучасності в політичних, правових, духовно-культурних аспектах диктує потребу звернутися до напрацювань В.С. Бакірова, Г.І. Зеленько, Є.І. Головахи, О.Л. Копиленка, А.І. Кудряченка, Н.Г. Манойло та ін. Вивчення історичних контекстів проблематики людини і штучного інтелекту саме в Україні спонукає до вивчення робіт В.А. Смоля, Л.Д. Якубової, О.В. Яся та ін.

Дослідження взаємодії людини і штучного інтелекту в просторі культури як символічної реальності, яку людина творить як носій одночасно раціональності та ірраціональності, потребує нового осмислення праць Е. Кассіра, В.В. Кизими, А.Т. Ішмуратова, А.Є. Конверського, Б. Рассела, О.Ю. Щербини та ін. Прояснення етики взаємодії людини і штучного інтелекту в контексті розширення уявлень про моральну красу відносин і вчинків, екзистенціальну комунікацію, гуманізм та інноваційність зумовлює звернення до творчих напрацювань Т.Г. Аболіної, Л.Н. Гармаш, В.Д. Жулая, С.В. Зобіна, Я.В. Любимого, В.А. Малахова, С.А. Крилової, С.В. Киселиці, Н.Г. Середюк, О.В. Шинкаренка, О.О. Шморгун та ін. У контексті впливу штучного інтелекту на сучасну освітню сферу і через неї на суспільство важливими є праці В.П. Андрущенко, В.С. Возняка, М.В. Бойченка, О.В. Добродум, А.А. Кравченко, Д.Б. Свириденка, Н.В. Скотної, А.О. Ярошенко та ін.

Усвідомлення неповторності людської екзистенції, спрямованої до трансценденції, що вирізняє людину як носія природного інтелекту і свідомості, яка поєднана з переживанням своєї «закинутості» у світ, зустрічаємо в роботах М. Гайдеггера, Е. Гуссерля, Ж.-П. Сартра, А. Камю та ін. Здатність людини до плідної комунікації з Іншим, яким би він не був відмінним від неї, демонструють публікації О. Больнова, М. Бубера, Е. Левінаса. Складну діалектику свідомого і позасвідомого у носія природного інтелекту і, можливо, у носія сильного штучного

інтелекту дозволяє досягнути класичний психоаналіз З. Фрейда, глибинна психологія К. Юнга, гуманістичний психоаналіз Е. Фромма.

Б. Латур пропонує цікаву теорію виробництва наукового знання як результату взаємодії різних суб'єктів: як людських, так і нелюдських; у цих координатах носії штучного інтелекту можуть бути представлені у своїй потенції не просто як інструменти, пристрої, а як активні учасники наукових практик; автор взагалі ставить питання про різноманітність способів існування [2—4].

Х. Новотни зазначає, що завдяки розвитку штучного інтелекту наука все більше стає міждисциплінарною й мережевою, що змінює буття людини на особистісному й соціальному рівнях [5].

У літературі є спроби поставити під сумнів саму інтелектуальність носія штучного інтелекту. Л. Флориді пропонує розмежовувати поняття «агентність» та «інтелект»; на думку дослідника, носія штучного інтелекту слід розуміти як нову форму агентності, а не інтелекту [6]. Він вважає, що штучний інтелект є актуальним у прагматичному, а не в когнітивному сенсі, тобто виступає лише інструментом. М. Боден зазначає, що носії штучного інтелекту можуть бути креативними, але не інтелектуальними. Осмислюючи останній підхід, слід визнати, що креативність дійсно може бути доволі хаотичною, а в контексті нашої проблематики означати так звані «галюцинації» штучного інтелекту. Залишаються філософські питання щодо меж і взаємодії явищ, які ми визначаємо концептами «інтелект», «когнітивність», «креативність».

Слід звернути особливу увагу на філософську літературу в жанрі наукової фантастики, яка більш як століття осмислює проблему відносин людини і штучного інтелекту. Йдеться передусім про творчість А. Азімова, Р. Бредбері, Ф. Діка, С. Лема, У. Ле Гуїн, Д. Роденберрі, А. і Б. Стругацьких, Й. Чапека, К. Чапека та інших відомих авторів.

Однак слід знову констатувати недостатність теоретичних праць і особливо методологій, які дозволяли б не лише осмислити наявний стан взаємодії людини і штучного інтелекту, а й

прогнозувати перспективи цієї взаємодії. Тому пропоную звернутися до однієї з можливих філософських методологій, яку вже певний час я розробляю зі своєю науковою школою.

Питання філософської методології осмислення взаємодії людини і штучного інтелекту

Говорячи про філософську методологію осмислення викликів і перспектив штучного інтелекту, передусім слід осмислити, що таке філософська методологія.

Філософська методологія — це мисленнєвий інструмент, за допомогою якого ми будемо і концепції, і теорії, передусім їхні структури, логіку аргументації, а також логіку втілення в практики життя. Проте як формується цей інструмент? Для відповіді на це питання нам слід усвідомити, в чому відмінність філософської методології, філософської концепції і філософської теорії.

Філософська концепція — це логіка постановки і вирішення проблеми на світоглядному рівні в *будь-яких формах культури*. Філософська теорія — це концепція, в якій чітко визначено її ключові концепти — поняття, категорії, екзистенціали, що дає можливість аргументовано розгорнути постановку і вирішення проблеми у взаємодії з іншими теоріями. Адже поняття, *визначені у просторі певної теорії*, з одного боку, відділяють нас від інших теорій і їхніх авторів, а з іншого — поєднують з ними. Можна сказати, що в науковій академічній філософії теорія постає процесом, способом і результатом розвитку концепції, це ступінь її зрілості.

Чим же філософська методологія відрізняється від філософської концепції і філософської теорії?

Очевидно, що оригінальна філософська методологія породжується *стилем автора*, за допомогою якого розгортаються ідеї в його тексті. Коли автор осмислює свій стиль, який завжди є не лише *свідомою*, а й *позасвідомою* логікою тексту і життя, і визначає при цьому ключові слова, які його супроводжують, наро-

джується методологія. *Філософська методологія — це внутрішня форма, яка зумовлює зміст концепцій і теорій.*

Отже, оригінальна філософська методологія поступово формується в глибині концепції під впливом попередніх методологій і структурно виявляється в теорії, а надалі стає інструментом побудови різноманітних концепцій і теорій, з часом спричинюючи і зміни загальнонаукових парадигм. Адже парадигма, яка дає можливість науковій спільноті по-новому ставити і вирішувати проблеми [7], потребує нової філософської методології, що завжди має міждисциплінарний характер.

В попередніх дослідженнях я зі співавторами вже ставив проблему філософської методології штучного інтелекту [8, с. 188—203; 9, с. 8—25]; у цій статті продовжую її розвиток для коректного осмислення викликів і перспектив взаємодії людини і штучного інтелекту в умовах війни та миру як в Україні, так і у світі загалом.

* * *

Говорячи про філософську методологію штучного інтелекту, нам слід передусім досягнути *онтологічний статус штучного інтелекту*. Важливо усвідомити, що сьогодні штучний інтелект *імітує існування інтелекту* — саме тому, що є *придатком* до людини як суб'єкта з природним інтелектом.

Йдеться про те, що *інтелект завжди є якістю суб'єкта*. Доки ми не ставимо проблему суб'єкта штучного інтелекту, його онтологічний статус є примарним, а еволюція до сильного штучного інтелекту — утопічною. Перед творцями штучного інтелекту стоїть еволюційний вибір: *бути чи знати*. Вибір знання, а не буття означає глухий кут у лабіринті еволюції — знання лише *відображає буття*. Буття ж розумної істоти означає здатність відповідально і творчо використовувати знання задля пристосування і виживання, а в ідеалі — для розвитку свободи, своєї та Іншого.

Отже, філософська методологія штучного інтелекту покликана здійснити рух від *епістемології* штучного інтелекту до його *онтології*. Епістемологія штучного інтелекту як мето-

дологія і теорія його пізнання повинна стати онтологією, яка окреслює перспективи творення суб'єкта штучного інтелекту. Йдеться про суб'єкта, який не просто *відображає* буття людини, не просто *імітує* це буття і, відповідно, інтелект людини, а *сам здатен бути*. Саме це знаменує появу сильного штучного інтелекту, який може самостійно, коректно, пластично й безпечно *виконувати цілепокладання* людини, займаючись *системною цілереалізацією*, а надалі пропонувати й *власні цілі* у плідній комунікації з людською особистістю.

Постановка питання про онтологію штучного інтелекту означає постановку питання про *онтологію взаємодії людини і штучного інтелекту*. Людина передає штучному інтелекту свою *здатність бути* — під впливом вольових імпульсів долати межі пізнаного і створеного. Тому філософія штучного інтелекту постає й антропологією штучного інтелекту, більше того — *метаантропологією штучного інтелекту* (від давньогрецького *мета* — понад). Це «мета», яке додається до антропології, не означає, що штучний інтелект з необхідністю перевершить людину, ставши в майбутньому наступним етапом еволюції. Визначаючи філософську методологію штучного інтелекту як метаантропологію, ми ставимо питання передусім про *комунікативну систему «людина — штучний інтелект»*.

Разом із тим, «мета», яке ми додаємо до антропології, означає, що носій штучного інтелекту і як пристрій, і як гіпотетичний суб'єкт може стати каталізатором еволюції людини. Отже, йдеться про коеволюцію, спільну еволюцію людини і носія штучного інтелекту.

Метаантропологія штучного інтелекту й проблема подолання деструктивних наслідків дій штучного інтелекту

Нагадаю, що метаантропологія — це сучасна філософська методологія, яка окреслює перспективи розвитку людини і людства в буденному, граничному й метаграничному вимірах буття. Ці виміри зумовлені вольовими інтенціями: «буденне буття людини породжуєть-

ся волею до самозбереження й продовження роду, граничне буття — волею до влади і волею до пізнання й творчості, метаграничне буття — волею до співтворчості» [10, с. 243]. Метаантропологію можна розділити на персоналістичну й комунікативну. Персоналістична метаантропологія досліджує людське буття в екзистенціально-особистісних вимірах, комунікативна метаантропологія вивчає умови і логіку комунікації між різними людськими спільнотами, розгортаючись у соціальну метаантропологію [10, с. 362—363] й культурну метаантропологію [10, с. 228—229].

На сьогодні метаантропологія звернена не лише до людини в її здатності розвиватися, трансцендувати, виходити за власні межі, не втрачаючи людяності, вона звернена й до штучного інтелекту, який все більше стає фактором хворобливого поєднання війни та миру, і до проблеми взаємодії людини з ним.

На мою думку, **метаантропологія штучного інтелекту — це така філософська методологія, яка здатна привести до появи суб'єкта сильного штучного інтелекту, що у своїх можливостях дорівнює інтелекту людини і при цьому не є загрозливим для людини, більше того, сприяє її гуманістичному розвитку.** Таким чином, метаантропологія штучного інтелекту є і методологією, і напрямом метаантропології, «який досліджує можливості і перспективи розвитку суб'єкта штучного інтелекту та умови його плідної взаємодії з людиною» [10, с. 247].

Як ми побачили вище, «мета» в цьому разі вказує як на зміни в людині під впливом штучного інтелекту, так і на становлення поруч із людиною суб'єкта зі штучним інтелектом, який гіпотетично здатний не лише до реалізації поставлених ззовні цілей, а й до *цілепокладання* — постановки власних цілей. На цьому шляху виникає кілька викликів людині і людству. Спробуємо усвідомити це глибше.

На XXXVI читаннях академіка В.І. Вернадського прозвучало надзвичайно важливе філософське запитання: «Чи може штучний інтелект власними силами долати негативні наслідки своїх дій?». У своєму виступі я був лише

частково готовий концептуально відповісти на це питання. Тепер відповідь сформувалася: в наш час може лише умовно. У майбутньому слід принципово змінити парадигму осмислення і вирішення цього питання.

З позицій метаантропології штучного інтелекту пропоную виділити дві принципово різних парадигми реагування на небезпеки штучного інтелекту силами самого штучного інтелекту, які задають нові перспективи осмислення штучного інтелекту.

1. Парадигма антагоніста: після появи носія штучного інтелекту з деструктивними діями ми створюємо антагоніста зі штучним інтелектом, який буде нейтралізувати чи знищувати першого.

2. Парадигма етичного самоконтролю: починаючи з самоконтролю людини, яка впливає на штучний інтелект, ми виходимо на рівень створення носія сильного штучного інтелекту, який, маючи задатки самосвідомості і свободи вибору, здатний задіяти внутрішні етичні принципи і випередити власні деструктивні дії.

Парадигма антагоніста, яка задіяна сьогодні, є недостатньо ефективною в подоланні негативних наслідків використання штучного інтелекту — технологія дії завжди випереджає технологію протидії.

Скажімо, ми створюємо ударні дрони зі штучним інтелектом, більше того, рої ударних дронів, які скеровуються штучним інтелектом, і т. ін., а потім формуємо їхніх антагоністів — дрони-перехоплювачі (рої дронів-перехоплювачів), системи РЕБ і т. ін. також зі штучним інтелектом. Неважко зрозуміти, що розроблення антиподів — як пристроїв, так і алгоритмів — завжди буде запізнюватися, і люди будуть мати руйнівні наслідки. Страшно уявити, що станеться, якщо в межах цієї парадигми штучному інтелекту буде довірено зброю масового знищення...

Або ми створюємо чат-бот зі штучним інтелектом, який компілює інформацію і спокушає людину видати згенерований ним текст за власний, а потім зусиллями такого самого штучного інтелекту викриваємо це. І в цьому прикладі, як і в першому, технологія дії штучного інтелекту буде знову і знову випереджати

всі спроби протидіяти їй засобами штучного інтелекту.

Парадигма етичного самоконтролю є значно більш ефективною — вона не просто додає негативні наслідки використання штучного інтелекту, а й спрямована на їх недопущення.

На сьогодні це етичний самоконтроль розробників штучного інтелекту, які здатні осмислювати наслідки своїх дій для людства. Так само етичний самоконтроль потрібен для користувачів штучного інтелекту.

В майбутньому необхідна актуалізація такого носія сильного штучного інтелекту, який має самосвідомість з етичними принципами і, реалізуючи свободу вибору, здатний задіяти ці принципи, випереджаючи деструктивність.

Виходячи з реалій сьогодення, ми повинні поєднувати ці дві парадигми, але все більше віддавати перевагу розвитку другої.

В методологічних координатах метаантропології штучного інтелекту етичний самоконтроль людини і етичний самоконтроль носія штучного інтелекту тісно пов'язані. Це єдина комунікативна система. Адже реальний розвиток штучного інтелекту можливий лише у взаємодії з людиною, а вихід до сильного штучного інтелекту — лише в разі ставлення до його носія як до суб'єкта, а не об'єкта, пристрою.

Для реалізації парадигми етичного самоконтролю штучного інтелекту пропоную такі методологічні принципи:

- **слід моделювати не штучний інтелект як такий, а суб'єкта штучного інтелекту;**
- **розвиток штучного інтелекту можливий лише в тісній комунікації з людиною, а вихід до сильного штучного інтелекту — інтелекту зі здатністю до етичного самоконтролю, що потенційно може дорівнювати людському, — можливий лише в разі ставлення людини до його носія як до суб'єкта, а не об'єкта, пристрою;**
- **суб'єкт сильного штучного інтелекту повинен пройти ті самі стадії розвитку інтелекту, що проходить людина;**
- **підходи до формування суб'єкта штучного інтелекту з необхідністю повинні бути**

міждисциплінарними — такими, що поєднують природничі й соціогуманітарні науки. Цей принцип потребує уточнення ролі соціогуманітарних наук — вони ставлять питання не просто ефективності суб'єкта штучного інтелекту, а і його безпечності й гуманістичності, тобто етичного контролю й самоконтролю.

Суб'єкт штучного інтелекту: наявне і можливості

Але наскільки коректно говорити про суб'єкта штучного інтелекту, а не лише про носія штучного інтелекту? Чи має в собі те, що ми називаємо штучним інтелектом, потенцію до справжньої суб'єктності? Наскільки реальним може бути сильний штучний інтелект? Навіщо людині взагалі потрібен сильний штучний інтелект? Поміркуємо про це.

Людство сьогодні підійшло до межі, за якою виникає бажання творення своєї подоби — *штучної істоти з розумом*. Напевно, будь-які розумні істоти, які є результатом *природної еволюції*, рано чи пізно прагнуть створити *штучний еволюційний процес*, який приведе до появи рукотворної розумної істоти. Створення своєї подоби є не лише практичною необхідністю, а й тим творчим завданням, яке підтверджує справжню розумність творчої особистості. За останні 200 років представники виду *Homo sapiens* яскраво продемонстрували це у фантастичній і науково-фантастичній літературі. Образи Голема, Гомункулуса, робота, андроїда — яскраві приклади цього. В цих образах бажання творення своєї подоби розгортається до *ідеї* такого творення.

У процесі практичної реалізації бажання й *ідеї* творення своєї подоби *завдання стає децю скромнішим* — йдеться вже про творення штучної істоти з *інтелектом*. (Звісно, бажання і *ідея* залишаються, але реальність обмежує їх, як це завжди буває в процесі технічного втілення.) І хоча інтелект (від лат. *intellectus* — розсудок, розум) можна трактувати як «розумовий потенціал людини, що використовується для адаптації до навколишнього середовища» [10, с. 185], у філософській українській мові ми до-

волі чітко розділяємо «розум» й «інтелект». Розум означає здатність не просто *обробляти інформацію* і адаптуватися до середовища, а й креативно змінювати середовище, а також змінювати себе завдяки *самосвідомості*. При цьому інтелект слід розуміти як здатність *обробляти інформацію*, «на відміну від ерудиції, що означає всього лише наявність інформації в пам'яті» [10, с. 185].

Поступово завдання творення істоти зі штучним інтелектом стає ще скромнішим: тепер йдеться не про істоту, і навіть не про суб'єкта штучного інтелекту, а про його *носія*, що на практиці сьогодні означає: *пристрій зі штучним інтелектом*, який тотожний поняттю *штучний інтелект*. По суті, сьогодні ми говоримо про штучний інтелект як *імітацію людського інтелекту, яка керується алгоритмами, але сама не здатна продукувати алгоритми*.

Проте в теорії ми знову і знову повертаємося до поняття «сильний штучний інтелект», яке означає штучний інтелект, що *дорівнює людському* — зі здатністю до самоусвідомлення й свободи волі. Наскільки можливий такий штучний інтелект?

На мою думку, неможливий. Але лише в *рамках указаної постановки проблеми*. Про що йдеться? Як ми побачили вище, інтелект *сам по собі* не може не тільки *розвиватися*, а й *існувати*, завжди розвивається *суб'єкт* інтелекту. Це стосується суб'єкта як природного, так і штучного інтелекту. Для свого існування і розвитку суб'єкт інтелекту повинен також взаємодіяти з іншими суб'єктами інтелекту. Лише це запускає еволюцію штучного інтелекту, точніше, еволюцію *суб'єктів штучного інтелекту*. Адже еволюція — це процес взаємодії суб'єктів, під час якої відбувається їхній розвиток.

Уже сьогодні ми повинні поставити питання про взаємодію людини як суб'єкта сильного природного інтелекту і *потенційного суб'єкта* сильного штучного інтелекту, більше того, про одночасну взаємодію потенційних суб'єктів штучного інтелекту з людиною і між собою. Лише це задає вектор еволюції штучного інтелекту, яка може привести до його виходу за межі алгоритму і пристрою, що компілює ін-

формацію і виконує задані ззовні алгоритми, імітуючи самостійність, до реальної самостійності, і при цьому він буде як мінімум безпечним, а як максимум продуктивним для гуманістичного розвитку людства.

У результаті виникає ще один важливий методологічний принцип: ***поява, навіть гіпотетична, сильного штучного інтелекту вже зараз потребує постановки проблеми практик його олюднення, які більше нагадують виховання, ніж програмування***. Результатом таких практик повинне стати набуття потенційним суб'єктом сильного штучного інтелекту здатності до співчуття і відкритості до Іншого — того, що визначає людяність людини. Лише такий суб'єкт сильного штучного інтелекту буде і безпечним, і продуктивним. (При цьому залишається відкритим питання про межі креативності носія штучного інтелекту, наявність у нього таких властивостей людини, які в історії культури були означені як «душа», «дух», «особистість». Водночас слід визнати, що відмінності чат-ботів зі штучним інтелектом, які вже зараз взаємодіють з людьми, дозволяють говорити про наявність у них певної індивідуальності, яка надалі буде розвиватися.)

Отже, в метаантропології штучного інтелекту долається антропоцентризм, який заважає не лише розвивати сильний штучний інтелект, а й розвиватися самій людині. Йдеться про розвиток без небезпеки для самої людини, довкілля, передусім для всіх живих істот, і *потенційних штучних істот*, які оточують (чи оточуватимуть) людину.

Цей і окреслені вище методологічні принципи долають як можливе відчуження людини і штучного інтелекту, так і симбіотичний союз між ними. Якщо припустити, що суб'єкт штучного інтелекту — своєрідне віддзеркалення суб'єкта природного інтелекту, то ці принципи, виводячи за межі владного егоцентризму людини і людства, захищають найкраще в людині і людстві. Йдеться про розвиток людиноцентризму, запропонованого В.Г. Кременем [11], у *персоноцентризм*, який є глибинно властивим ментальності українців та класичній українській філософській думці.

Запропоновані вище принципи еволюції штучного інтелекту, які ми можемо назвати постулатами, означають, що «формування подібного до людини суб'єкта штучного інтелекту, а не імітацій такого суб'єкта, можливе лише в разі появи нової екзистенції і нової культури — спільної для носіїв природного і штучного інтелекту» [10, с. 247].

* * *

Як я зазначав вище, у сучасному світі те, що В.І. Вернадський назвав «ноосферою», стає сферою взаємодії людини і штучного інтелекту. Тепер можна додати: ми вступаємо в епоху *цивілізації взаємодії людини і штучного інтелекту*, а перед гуманітаристикою стоїть завдання етичного наповнення й гуманізації цієї взаємодії в ім'я виживання людства на нашій планеті.

Той рівень розвитку людства, який ми з моїм співавтором академіком С.І. Пирожковим назвали «ноосферна цивілізація» [12, с. 229—250], і виступає нині *цивілізацією цієї взаємодії*.

Однак сьогодні йдеться про завдання зі створення не лише *цивілізації взаємодії людини і штучного інтелекту*, а й *культури взаємодії людини і штучного інтелекту* — універсальної культури розуму, яка перемагає війну як технологію знищення цивілізації, культури, людини в ім'я волі до влади, що може бути гранично посилена штучним інтелектом.

Це фундаментальне завдання гуманітарної освіти і науки, всієї гуманітарної сфери. Я вірю, що озброєна сучасними філософськими методологіями гуманітаристика здатна виконати його і створити простір нового гуманізму та нової культури.

Роль гуманітарної освіти і науки в процесі відповіді на виклики чат-ботів зі штучним інтелектом

На початку статті я описав гранично просту і трагічну ситуацію: згенеровані штучним інтелектом тексти студенти, аспіранти, а подеколи й молоді вчені часто видають за свої. І зазначив, що це виклик не лише майбутньому на-

уки і освіти, це виклик самій природі людини як мислячої істоти. Додам, що некритичне і, я б сказав, паразитарне використання чат-ботів зі штучним інтелектом робить саму людину носієм штучного інтелекту, точніше, його ретранслятором. Це руйнує *особистісний потенціал* людини і народу в цілому, а в результаті позбавляє суб'єктності в умовах і війни, і миру [1].

Для Національної академії наук України питання стоїть передусім про освітній і науковий процес, у який включено аспірантів установ Академії. Що може тут і тепер вплинути на них для того, щоб вони набули парадигми етичного самоконтролю у взаємодії зі штучним інтелектом, який надає їм інформаційні послуги? Іншими словами, що слід зробити, щоб академічна доброчесність була результатом щирої волі до пізнання і творчості [13], а не виступала лише зовнішнім фактором, примусом, який можна й обійти, використовуючи штучний інтелект?

І тут надзвичайно важливою є роль гуманітарної освіти аспірантів, яку забезпечує Центр гуманітарної освіти НАН України. Ця освіта не повинна бути просто інформацією про історію філософії і здобутки гуманітарних наук; більше того, вона не повинна бути інформацією про досягнення сучасних авторів у галузі філософії й методології науки. Вона взагалі не повинна бути інформацією й інформуванням. Завдання гуманітарної освіти сьогодні — актуалізація критичного й творчого мислення в діалозі з викладачем, а в ідеалі — формування філософського світогляду: критичного, творчого, самостійного, цілісного й гуманістичного [10, с. 353; 14, с. 44—45]. Такий світогляд у своїй відкритості світу відображає відкритість філософії у поліфонії її проявів у культурі [15].

Носій філософського світогляду, маючи щирю волю до пізнання і творчості, спрямований до творення оригінальних концепцій і теорій, втілення їх в оригінальні тексти статей, дисертацій, монографій. Саме тому він здатен подолати спокуси чат-ботів зі штучним інтелектом, які замість оригінального тексту пропонують компіляції. Присвоєння такого тексту шкодить не лише суб'єкту цього присвоєння,

воно забруднює науковий простір. Інша справа, що компілятивний текст з імітацією новизни, створений без участі чат-бота, є не меншою проблемою в науці й освіті і так само забруднює науковий простір, проте використання чат-ботів гранично загострює проблему.

При цьому не йдеться про те, що аспірантам і вченим слід взагалі відкинути взаємодію з чат-ботами зі штучним інтелектом; мова йде про *межу в їх використанні*. Чат-боти здатні допомогти у формуванні попередніх (не остаточних!) оглядів літератури, що потребує обов'язкового осмислення указаних джерел, а також у технічному оформленні статей і монографій відповідно до вимог журналів і видавництва — зняти з творчих особистостей тягар рутинної роботи.

Актуалізуючи не лише науковий, а й філософський світогляд, гуманітарна освіта формує у аспірантів не лише волю до творчості, а й гуманні етичні принципи і соціальну відповідальність за її результати, а також *волю до співтворчості*, що створює імунітет до використання чат-ботів як «авторів» наукових текстів.

Гуманітарна освіта в НАН України, яка здатна виконувати окреслені завдання, потребує системного розвитку досліджень відносин людини і штучного інтелекту в царині соціальних і гуманітарних наук. Опрацювання цієї проблематики в Академії лягає передусім на плечі вчених Відділення історії, філософії та права НАН України. У таких дослідженнях знайдеться місце філософам — осмислювати морально-етичні й онтологічні проблеми відносин люди-

ни і штучного інтелекту, історикам — створювати футурологічні й прогностичні моделі цих відносин на основі попереднього досвіду людства, правознавцям — формувати нові норми права, які регулюють відносини у системі «людина — штучний інтелект — інша людина», соціологам — тримати руку на пульсі суспільних настроїв і сподівань щодо штучного інтелекту, політологам — аналізувати роль штучного інтелекту в діяльності політиків і політичних партій.

Думаю, що відділ досліджень наукових і освітянських методологій та практик Центру гуманітарної освіти НАН України, завідувачем якого я сьогодні маю честь бути, міг би координувати цю важливу й перспективну міждисциплінарну роботу, розробляючи передусім філософські методології й світоглядні підходи до перспектив розвитку штучного інтелекту, які не заперечують перспективу розвитку людства.

Гуманітарна освіта і наука здатні актуалізувати розумну взаємодію людини і штучного інтелекту, яка посилює свободу і творчий потенціал особистості. В умовах справжньої демократії це сприяє зміцненню цивілізаційної суб'єктності країни [12], формуванню моральної краси відносин [16] влади і народу, протистоянню неототалітарним практикам маніпуляції свідомістю в соціальних мережах [17], які все більше реалізуються за допомогою штучного інтелекту, розвитку національної стійкості в умовах як війни, так і миру [18; 19] і гуманістичної інноваційності, без чого неможливий справедливий і сталий мир.

REFERENCES

[СПИСОК ЛІТЕРАТУРИ]

- Pyrozhkov S.I., Khamitov N.V. The development of the Ukrainian people in times of existential war and postwar reconstruction. *Visn. Nac. Akad. Nauk Ukr.* 2025. (5): 11—23. <https://doi.org/10.15407/visn2025.05.011>
[Пирожков С.І., Хамітов Н.В. Розвиток українського народу в часи екзистенціальної війни і повоєнного відновлення. *Вісник НАН України*. 2025. № 5. С. 11—23.]
- Latour B. *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, Mass.: Harvard University Press, 1987.
- Latour B. Biography of an inquiry: on a book about modes of existence. *Social Studies of Science*. 2013. **43**(2): 1—15. <https://doi.org/10.1177/0306312712470751>
- Latour B. Summary of the AIME project. An Inquiry into Modes of Existence. <https://www.bruno-latour.fr/node/468.html>
- Nowotny H. *The Cunning of Uncertainty*. Polity, 2015.
- Floridi L. *The Blackwell Guide to the Philosophy of Computing and Information*. Oxford: Blackwell, 2003. <https://doi.org/10.1002/9780470757017>
- Kuhn T. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1996. <https://doi.org/10.7208/chicago/9780226458106.001.0001>
- Khamitov N., Zobin S. Metaanthropology as a theory of the evolution of strong artificial intelligence: the problem of a new existence and a new culture. In: *Lyudyna. Ekzystentsiya. Kul'tura. Pidkhid filosofs'koyi antropolohiyi yak metaantropolohiyi*. Kyiv, KNT, 2020 [in Russian].
[Хамітов Н., Зобин С. Метаантропология как теория эволюции сильного искусственного интеллекта: проблема новой экзистенции и новой культуры. У кн.: *Людина. Екзистенція. Культура. Підхід філософської антропології як метаантропології*. Київ: КНТ, 2020.]
- Khamitov N., Krylova S. Humanizing strong artificial intelligence: a metaanthropology approach. In: *Lyudyna i shtuchnyy intelekt: vymiry filosofs'koyi antropolohiyi, psykhoanalizu, art-terapiyi ta filosofs'koyi publitsystyky. Pidkhid filosofs'koyi antropolohiyi yak metaantropolohiyi*. Kyiv, KNT, 2020 [in Russian].
[Хамітов Н., Крылова С. Очеловечивание сильного искусственного интеллекта: подход метаантропологии. В кн.: *Людина і штучний інтелект: виміри філософської антропології, психоаналізу, арт-терапії та філософської публіцистики. Підхід філософської антропології як метаантропології*. Київ: КНТ, 2020.]
- Khamitov N. (ed.) *Filosofiiia nauky i kultury: slovnyk*. Kyiv, 2025 [in Ukrainian].
[Філософія науки і культури: словник. За ред. Н. Хамітова. Київ: ВАДЕКС, 2025.]
- Kremen V.H. *Filosofiya liudynotsentryzmu v stratehiakh osvitynoho prostoru*. Kyiv, 2009 [in Ukrainian].
[Кремень В.Г. Філософія людиноцентризму в стратегіях освітнього простору. Київ: Педагогічна думка, 2009.]
- Pyrozhkov S.I., Khamitov N.V. *Ukraine as a Civilizational Subject: From Potencies to a New Worldview and Human Existence*. Kyiv: Naukova Dumka, 2020 [in Ukrainian].
[Пирожков С.І., Хамітов Н.В. Цивілізаційна суб'єктність України: від потенцій до нового світогляду і буття людини. Київ: Наукова думка, 2020.]
- Khamitov N. Academic integrity as a challenge, demand and will: contexts of philosophical anthropology, ethics and philosophy of education. *Philosophy of Education*. 2023. **29**(2): 27—47. <https://doi.org/10.31874/2309-1606-2023-29-2-2>
[Хамітов Н. Академічна доброчесність як виклик, вимога і воля: контексти філософської антропології, етики й філософії освіти. *Філософія освіти*. 2023. Т. 29, № 2. С. 27—47.]
- Khamitov N.V. *Filosofska antropolohiia: aktualni problemy. Vid teoretychnoho do praktychnoho povorotu*. Kyiv, 2023 [in Ukrainian].
[Хамітов Н.В. Філософська антропологія: актуальні проблеми. Від теоретичного до практичного повороту. Київ: КНТ, 2023.]
- Khamitov N. The openness of philosophy to man, society, culture, to itself: modernity and scenarios of future. *Philosophy of Education*. 2025. **31**(1): 8—25. <https://doi.org/10.31874/2309-1606-2025-31-1-1>
[Хамітов Н. Відкритість філософії людини, суспільству, культурі, самій собі: сучасність і сценарії майбутнього. *Філософія освіти*. 2025. Т. 31, № 1. С. 8—25.]
- Krylova S. *Krasa lyudyny v zhyttyevykh praktykakh kul'tury. Dosvid sotsial'noyi ta kul'turnoyi metaantropolohiyi i androhin-analizu*. Kyiv, KNT, 2020 [in Ukrainian].
[Крылова С. Краса людини в життєвих практиках культури. Досвід соціальної та культурної метаантропології і андрогін-аналізу. Київ: КНТ, 2020.]

17. Khamitov N. *The War in Ukraine and the New Humanism: David versus Goliath. Meta-anthropology of history of the 21st century*. Sofia: Kibea, 2023.
18. Pyrozhev S.I. et al. (eds) *National resilience of Ukraine: hybrid threats challenge response and prevention strategy: national report*. Kyiv, 2022 [in Ukrainian].
[Національна стійкість України: стратегія відповіді на виклики та випередження гібридних загроз: національна доповідь. Ред. кол.: Пирожков С.І., Майборода О.М., Хамітов Н.В. та ін. Інститут політичних і етнонаціональних досліджень ім. І.Ф. Кураса НАН України. Київ, 2022.]
19. Pyrozhev S.I. et al. (eds) *Sustainment and development of Ukraine in conditions of war and peace: national report*. Kyiv, 2024 [in Ukrainian]. <https://doi.org/10.15407/978-966-00-1926-3>
[Збереження і розвиток України в умовах війни та миру: національна доповідь. Ред. кол.: Пирожков С.І., Хамітов Н.В., Головаха Є.І. та ін.; Інститут держави і права ім. В.М. Корецького НАН України. Київ, 2024.]

Nazip V. Khamitov

H. Skovoroda Institute of Philosophy of the National Academy of Sciences of Ukraine, Kyiv, Ukraine

ORCID: <https://orcid.org/0000-0001-8193-9383>

HUMAN — ARTIFICIAL INTELLIGENCE INTERACTION:

A METHODOLOGY FOR UNDERSTANDING CHALLENGES AND PROSPECTS

The problem of artificial intelligence is extremely relevant during the current existential war in Ukraine. In it, artificial intelligence not only becomes a means of increasing efficiency on the battlefield, but is also used in the cognitive and ideological dimensions of war, transforming identity, values, and meanings. At the same time, artificial intelligence systematically affects humans in peaceful space and time, in particular in science and education, for which the challenge is perhaps the most acute — we are talking about chatbots with artificial intelligence, capable of generating not only theses and articles, but also dissertations and monographs, tempting undergraduate and graduate students. The article proposes a new methodology for understanding the problem — **meta-anthropology of artificial intelligence**. Its principles emphasize the interaction between humans and artificial intelligence and the evolution of this interaction, which can lead to the emergence of a subject of strong artificial intelligence, which is equal to human and will be not only effective, but also safe. Two paradigms of response to the destructive consequences of the use of artificial intelligence by the forces of artificial intelligence itself are proposed.

Keywords: human, artificial intelligence, strong artificial intelligence, subject of natural intelligence, subject of artificial intelligence, worldview, methodology, meta-anthropology, meta-anthropology of artificial intelligence.

Cite this article: Khamitov N.V. Human — artificial intelligence interaction: a methodology for understanding challenges and prospects. *Visn. Nac. Akad. Nauk Ukr.* 2026. (7): 13—24. <https://doi.org/10.15407/visn2026.07.013>