

# ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

УДК 1:001.8

Ю. Д. ГЕНСЦЬКИЙ<sup>1\*</sup>

<sup>1\*</sup>Національний університет біоресурсів і природокористування України (Київ), ел. пошта [yuraletter@gmail.com](mailto:yuraletter@gmail.com), ORCID: 0000-0003-0858-6862

## ФІЛОСОФСЬКО-АНТРОПОЛОГІЧНЕ ЗНАЧЕННЯ РОЗВИТКУ ШТУЧНО СТОВРЕНИХ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ

**Мета.** Осмислення філософського-антропологічного значення розвитку штучно створених інтелектуальних систем вимагає проведення аналізу соціально-антропологічного змісту проблеми розвитку міжмашинної взаємодії в контексті медіафілософського пр�ксису, антропологічного змісту природи інтелекту, розгляду специфіки концепту штучно створених інтелектуальних систем в середовищі M2M, розвитку соціокогнітивних практик міжмашинної взаємодії, їх соціогуманітарного потенціалу. **Методологія дослідження.** Реалізація мета вбачається в використанні науково-теоретичної бази медіафілософії, філософської антропології, медіафілософського підходу до аналізу суспільства, науки та техніки, використанні публікацій з обраної тематики дослідження. **Наукова новизна.** Розглянуто концепт штучно створених інтелектуальних систем в аспекті соціогуманітарного потенціалу їх становлення та розвитку в середовищі M2M. Проаналізована проблема машинного навчання як технологія трансформації M2M. Окреслено антропологічні загрози розвитку штучно створених інтелектуальних систем. **Висновки.** З точки зору глобальних ризиків, одна з найбільш критичних обставин у зв'язку з штучно створеною інтелектуальною системою, це те, що вона може посилити свій інтелект надзвичайно швидко. Очевидна причина підозрювати таку можливість - це рекурсивне самополіпшення. Така система стає розумнішою, в тому числі розумнішою щодо написання внутрішньої когнітивної функції, тобто здатна переписати свою існуючу когнітивну функцію, щоб вона працювала краще. Це зробиць такі системи ще розумнішими, в тому числі розумнішим щодо завдання перероблення себе. Успіх штучного інтелекту може стати початком кінця людської раси. Майже будь-яка технологія, потрапляючи в зловмисні руки, виявляє потенційні можливості для заподіяння шкоди, але коли йдеться про штучно створені інтелектуальні системи, виникає нова проблема, пов'язана з тим, що ці зловмисні руки можуть належати самій технології.

*Ключові слова:* штучно створена інтелектуальна система, міжмашинне середовище, штучний інтелект, нейронна мережа, машинне навчання

### Актуальність теми дослідження

Проблема соціально-антропологічного функціоналу високих технологій давно перестала бути суто інженерною, все більше проникаючи в усі галузі суспільних наук. Взаємодіючи між собою на різних рівнях, такі технології утворюють специфічне середовище, яке в інженерії позначається як мережа міжмашинної взаємодії. Специфічність цього середовища в добу Постмодерну пояснюється тим, що воно створюється смарт-технологіями, які на відміну від звичайних здатні комунікувати без втручання людини. Парадокс полягає в тому, що технологія створена людиною (та уподібнюючись їй) сама створює середовище в якому людина виявляється не потрібною. Створене когнітивними смарт-приладами середовище повністю заповнює всі сфери життя, а людство настільки активно взаємодіє з ними, починаючи від ментального і закінчуючи фізичним рівнем,

що буквально зростається з «розумними речами». Кількість взаємодіючих між собою смарт-речей перевищує кількість людей, а з кожним новим технологічним поступом коли вони стають все більше самостійнішими, роль людини в функціонуванні міжмашинного середовища (що позначається аббревіатурою M2M) зменшується, середовищу міжлюдського спілкування відходить другорядна роль. Розробки філософської антропології відстають від розвитку M2M на десяток років і як правило стосуються лише рефлексії інформаційної парадигми в розробках штучного інтелекту. Залишається поза увагою когнітивна складова, як основний вектор розвитку інтелектуальних систем. Отже, актуальність аналізу соціогуманітарного функціоналу штучно створених інтелектуальних систем залежить конвєргенційний потенціал філософської антропології та медіафілософії.

Займаючись проблемами конструювання

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

штучного інтелекту, ми наближаємось до кращого розуміння людського інтелекту. Потрібно уважно відношення до деталей і тонкощів інтелектуальних процесів, бо інакше, якщо щось не врахували, то їх машинна імітація просто не працюватиме. спрощення в стислих до Створення формальних структур і моделей штучного інтелекту полегшує розуміння роботи природного інтелекту. Крім того, дослідження в області інтелектуальних систем в якомусь сенсі звільняють від ідеологічних догм, якими іноді навантажені дискусії про природу свідомості і мислення, оскільки для створення технічних пристроїв придатні будь-які ідеї і підходи, якщо вони дають плідні результати [6].

*Аналіз останніх досліджень і публікацій.* Проблема розвитку штучного інтелекту за подібністю до нейронної взаємодії мозку з використанням технологій навчання аналізується в статті «How to Make a Mind» Реймонда Курцвейла. Етичний аспект впровадження та розвитку штучного інтелекту розглянутий в роботах Лучано Флориді «Artificial Agents and Their Moral Nature» та Рональда Сандлера «Ethics and Emerging Technologies».

Формулювання **цілей дослідження**. 1) Проаналізувати соціально-антропологічний зміст проблеми розвитку міжмашинної взаємодії в контексті медіафілософського пр�ксису; 2) Розглянути специфіку концепту штучно створених інтелектуальних систем в середовищі M2M; 3) Виокремити та осмислити рефлексію розвитку соціокогнітивних практик міжмашинної взаємодії, їх соціогуманітарний потенціал.

### Виклад основного матеріалу

Творці технологій завжди орієнтувалися на своє розуміння того, що таке людський інтелект і як влаштована розумна поведінка, щоб відтворити інтелект, реалізований природою на біологічному субстраті, на субстраті комп'ютерної технології. Показово, що сама ідея про можливість такого відтворення, вписується в дуалістичну традицію європейської думки. У свою чергу, розуміння природи інтелекту мінялося значною мірою під впливом технологічної думки. Адже порівняння того, що робить людина, з тим, що може зробити розумна машина, навіть у випадку якщо машина не справляється з людською формою інтелектуальної діяльності, кінець кінцем конструює людину в термінах машини [10].

Кібернетичні витоки часто застосовуються в антропології, про те, що людина це істота, що переробляє інформацію, здатна до збереження і накопичення, і що культура в кінцевому результаті може бути розглянута як семіотична макросистема, функція якої полягає в поза генетичній передачі комплексу інформації від одного покоління до іншого, спираючись на культурні «коди» або мови, до яких відноситься, разом з природною мовою і інший семіотичні системи. Таким чином, вивчення культури зводиться до опису і аналізу цих культурних кодів, що свого роду програмами, що регулюють людську поведінку [8].

У сучасному світі штучні інтелектуальні системи використовуються буквально у всіх сферах життєдіяльності людини і кардинальним чином трансформують умови її існування. Спостерігається стійке збільшення масштабів застосування штучних інтелектуальних систем в повсякденному житті.

Повсякденність, як вимір антропосфери, де значною мірою формується людина, її характер і світогляд, піддається різного роду трансформаціям, спричинених активним застосуванням цифрових технологій. Перетворюючи структуру повсякденного життя, технології неминуче змінюють тим самим і саму людину. В даний час техніка виступає найбільш активною силою, під впливом якої повсякденність динамічно видозмінюється. Активне застосування систем з елементами штучного інтелекту істотно перетворює повсякденну реальність і формує особливий тип цифрового світосприйняття людини. Сучасна повсякденність складається з продуктів технічної цивілізації, від яких неможливо абстрагуватися, у зв'язку з чим технології стають провідним фактором формування ціннісних орієнтирів людини. У зв'язку з цим виникає необхідність філософсько-антропологічного осмислення наслідків застосування нових технологій, які непомітно входять у життя людства і стають його невід'ємною частиною.

Наслідки поширення штучних інтелектуальних систем носять дуальний характер для людини. Засоби масової інформації та реклами, стимулюючи необмежений розвиток матеріальних потреб, активізують широке застосування різних штучно створених інтелектуальних систем в побуті. Техніка, оснащена елементами штучного інтелекту, що є засобом підвищення

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

рівня комфорту і безпеки, стає характерною рисою повсякденного побуту. При цьому вже не кошти використовуються для зручності людини, а навпаки, людина стає заручником цих коштів. Досягнення комфорту в повсякденному житті, культивування надмірних потреб, орієнтація на споживання здатні привести до «розмивання» людської екзистенції і деградації особистості, до формування індивіда, що представляє собою автомат, нездатний до творчої креативної активності, діяльність якого спрямована тільки на виконання закладеної в нього екзопрограми.

Нові технології дають державі і корпораціям інструменти прямого та непрямого спостереження і контролю над особистістю для здійснення управління соціумом, що тягне за собою позбавлення індивіда свободи волі і перетворення його на слухняного споживача, позбавленої критичного мислення цифровимірної людини. Проблема тотального стеження, в якій опиняється сучасна людина на рівні повсякденного життя стала реальністю вже сьогодні.

Виникає питання про перспективи адаптації людини до непередбачуваної нової реальності, яка раніше завжди здавалася їй простою, зрозумілою і знайомою. Сучасна цифрова реальність з гігантським технопарком штучно створених інтелектуальних систем - мало досліджена реальність, яка обумовлює значну некеріваність зростаючих динамічних змін соціуму, культури і самої людини. Створюючи системи з елементами штучного інтелекту, людина віддає їм право прийняття рішень в різних областях діяльності, тим самим потрапляючи в середовище, яке управляється вже не нею, а машинами. В результаті формується штучна інтелектуальна повсякденність, що володіє новими властивостями в порівнянні з повсякденністю попередніх часів. Ця нова реальність знаменує становлення пост людини, трансформацію антропосфери в «цифрову антропосферу».

Питання, чи здатна машина мислити, має довгу і суперечливу історію. В філософсько-антропологічному аспекті воно тісно пов'язане з двома школами: дуалістичною і матеріалістичною. З точки зору дуалізму, мислення не є матеріальним і тому розум не можна пояснити тільки за допомогою фізичних понять. З іншого боку, матеріалізм свідчить, що розум можна пояснити фізично, таким чином, говорячи про можливість існування розуму,

створеного штучно. Історія штучного інтелекту як нового наукового напрямку починається в середині XX століття. До цього часу вже було сформовано безліч передумов його зародження: серед філософів давно йшли суперечки про природу людини і процес пізнання світу, нейрофізіологи і психологи розробили ряд теорій щодо роботи людського мозку і мислення, економісти і математики задавалися питаннями оптимальних розрахунків і представлення знань про світ у формалізованому вигляді; зародився фундамент математичної теорії обчислень - теорії алгоритмів - і були створені перші комп'ютери.

Нашою реальністю стає те, що штучно створена інтелектуальна система активно взаємодіє із зовнішнім середовищем, тобто сприймає вплив навколишнього середовища і відповідним чином реагує на них. Як результат впливу середовища, система по тим чи іншим правилам формує «відповідь» (реакцію) на даний вплив. Характер активної взаємодії з середовищем, реалізованої системою, визначається властивостями даної системи і ресурсами, якими вона володіє. У процесі формування реакції на той чи інший вплив середовища система може як спиратися на свої ресурси, так і використовувати якісь інформаційні джерела зовнішнього характеру. Чим менше система залежить від інформаційних джерел і керуючих команд, тим вище ступінь її автономності. Штучно створені інтелектуальні системи - це в певному сенсі самодостатні системи, на які може покладатися вирішення певного комплексу прикладних задач. Штучні інтелектуальні системи націлені на те, щоб максимально самостійно вирішувати поставлені перед ними завдання. Структурно роботу штучних інтелектуальних систем в деякому середовищі можна розділити на три сфери: 1. Сприйняття ситуації, яка є поєднанням стану навколишнього середовища і внутрішнього стану (сенсорна сфера). 2. Формування «відповіді» (реакції) на поточну або прогнозовану ситуацію (зміна стану системи в її фазовому просторі, реконфігурація, самонавчання, самоорганізація). 3. Реалізація сформованої реакції на поточну або прогнозовану ситуацію. Штучні інтелектуальні системи здатні: досягати поставлених цілей в високодинамічному середовищі зі значним числом різномірних невизначеностей в ній; коригувати поставлені цілі; формувати нові цілі і комплек-

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

си цілей, виходячи із закладених в систему установок (мотивацій); отримувати нові знання, накопичувати досвід вирішення завдань, модифікувати свою поведінку (реакції на зміну ситуації) на основі отриманих знань і накопиченого досвіду, в тому числі навчитися вирішенню завдань, не передбачених початковим проектом системи; утворювати «спільноти» штучно створених інтелектуальних систем в середовищі М2М, спрямованих на взаємодію їх членів при вирішенні деякої загальної задачі; здійснювати самовідтворення із залученням автономних ресурсів, можливо, зі змінами в «генотипі» системи (для підтримки процесів еволюції у спільнотах штучно створених інтелектуальних систем). При такому підході поведінка таких систем наближається до людської. Можливість емуляції чи симуляції інтелектуальними системами соціальної комунікації становить складну антропологічну проблему докорінної трансформації антропосфери.

Можливості нових машин, реалізованих на штучно створених інтелектуальних системах, в плані швидкості обчислень виявилися більше людських, тому у науковому світі постало питання: які межі можливостей комп'ютерів і чи зможуть машини досягнути рівня розвитку людського мозку? У 1950 році один з піонерів в області обчислювальної техніки, англійський учений Алан Тюрінг, написав статтю під назвою «Чи може машина мислити?» [4], в якій описує процедуру, за допомогою якої можна буде визначити момент, коли машина зрівняється в плані розуму з людиною, що отримала назву тесту Тюрінга. Стандартна інтерпретація цього тесту звучить наступним чином: «Людина взаємодіє з одним комп'ютером і однією людиною. На підставі відповідей на питання людина повинна визначити, з ким вона розмовляє: з людиною чи комп'ютерною програмою. Завдання комп'ютерної програми - ввести людину в оману, змусивши зробити невірний вибір» [4].

Проблема розуміння штучного інтелекту ускладнюється ще й тим, що у філософії не вирішено питання про природу людського інтелекту. Комп'ютерне моделювання інтелекту зіткнулося з рядом труднощів, які показують недостатні евристичні можливості сучасної філософії і психології свідомості, у розумінні інтелекту.

Те розрізнення між інтелектом, розсудком і

розумом, яке приймало різні форми в історії філософії виявляло не лише різні сегменти раціональної свідомості, але і дозволяло вказувати на ті сегменти, які забезпечують продуктивний пізнавальний синтез в різних формах, розкривати за межами відтворюваного в актах мислення те, що залишається невідтворюваним залишком, творчим началом, що відноситься або до інтуїції, або до поза дискурсивних актів.

Зараз встановлено, що елементи мислення не є винятковою прерогативою людини, а є у тварин і проявляються у них в різних формах. Найбільш віддаленою від здібностей інших живих організмів властивістю ментальної інформаційної системи людини є довільне мислення в поняттях - рефлексія, властива і буденному, і теоретичному рівню пізнання [6].

Немає точного критерію досягнення комп'ютерами «розумності», хоча на зорі штучного інтелекту було запропоновано ряд гіпотез, наприклад, тест Тюрінга або гіпотеза Ньюелла - Саймона.

Штучні системи стають «розумними», «інтелектуальними» не тому, що в них є органи чуттів і розум. «Розумні» штучні системи здатні отримувати потрібну їм інформацію, накопичувати й переробляти її, обмінюватися нею між собою. І саме тому, що поведінка «розумних» речей зовні нагадує розумну поведінку людей, творці інформатики характеризують такі речі метафорами «інтелектуальний», «розумний» і говорять про властивий їм обчислювальний інтелект, про розсіяну розумність, про інтелектуалізацію фізичних, хімічних, біологічних речей [2].

Незважаючи на наявність безлічі підходів як до розуміння задач штучного інтелекту, так і створенню когнітивних інформаційних систем М2М, можна виділити два основних підходи до розробки штучно створених інтелектуальних систем [13]:

- Спадний, семіотичний - створення експертних систем, баз знань і систем логічного висновку, що імітують високорівневі психічні процеси: мислення, міркування, мова, емоції, творчість і т. д.;

- Висхідний, біологічний - вивчення нейронних мереж і еволюційних обчислень, що моделюють інтелектуальну поведінку на основі біологічних елементів, а також створення відповідних обчислювальних систем, таких як нейрокомп'ютер або біокомп'ютер на базі

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

мемристорів та нейристорів.

Вважаю, що є доцільним формування гібридного підходу, при якому експертні правила умовиводів можуть генеруватися нейронними мережами, а породжуючі правила отримують за допомогою статистичного навчання. Прихильники цього підходу вважають, що гібридні інформаційні системи будуть значно сильнішими, ніж сума різних концепцій окремо.

Серед великої кількості високотехнологічних продуктів, штучно створені інтелектуальні системи в мережі M2M володіють певними особливостями, що відрізняють їх від звичайних комп'ютерних систем:

- Вони здатні вирішувати завдання не лише за заданим алгоритмом, але й за побудовою завдання, навіть якщо воно не чітко сформульовано;

- Вони здатні створювати нові алгоритми і навчатись, вдосконалюючи тим самим свої можливості.

Ці характеристики є визначальними в умовах вирішення таких завдань, які характеризуються наявністю великих масивів інформації Big Data, її неоднорідністю, неповнотою та розпорошеністю.

Більшість дослідників вважають наявність власної внутрішньої моделі світу у технічних систем передумовою їх "інтелектуальності". Формування такої моделі пов'язано з подоланням синтаксичної однобічності системи, тобто з тим, що символи або та їх частина, якою оперує система, інтерпретовані, мають семантику. Гіпотеза про необхідність внутрішньої моделі світу є дійсною в разі висунення критеріїв людської інтелектуальності до штучної, але у зв'язку з недостатнім розумінням природи і сутності інтелекту є дуже хиткою основою.

Характеризуючи особливості інтелектуальних систем, Л. Т. Кузін визначає такі їх властивості: 1) наявність в них власної внутрішньої моделі зовнішнього світу; ця модель забезпечує індивідуальність, відносну самостійність системи в оцінці ситуації, можливість семантичної і прагматичної інтерпретації запитів до системи; 2) здатність поповнення наявних знань; 3) здатність до дедуктивного висновку, тобто до генерації інформації, яка в явному вигляді не міститься в системі; ця якість дозволяє системі конструювати інформаційну структуру з новою семантикою і практичною спрямованістю; 4)

вміння оперувати в ситуаціях, пов'язаних з різними аспектами нечіткості, включаючи "розуміння" природної мови; 5) здатність до діалогового взаємодії з людиною; 6) здатність до адаптації [2].

На питання, чи всі перераховані умови є обов'язковими, необхідні для визнання системи інтелектуальною, вчені відповідають по-різному. У реальних дослідженнях, як правило, визнається абсолютно необхідним наявність внутрішньої моделі зовнішнього світу, і при цьому вважається достатнім виконання хоча б однієї з перерахованих вище умов.

Недоліком існуючих штучно створених інтелектуальних систем є те, що інформація в них передається односторонньо, на відміну від головного мозку людини, де одночасно функціонують мільйони нейронів, які забезпечують паралельну обробку великих масивів різномірної інформації. Саме тому в основу роботи вчених лягло прагнення створити прототип системи штучного інтелекту з допомогою нанотехнології, що дозволить машині працювати подібно клітинам головного мозку - одночасно зчитуючи та обробляючи інформацію в різних напрямках. Іншими словами, робот зможе приймати рішення в незнайомій ситуації, яка не описана в його програмі, на основі того, що він пережив в минулому, використовуючи знання, отримані в процесі навчання, за аналогією до життя людини [7].

Спробою реалізувати функціонал мозку в M2M став принцип створення штучних нейронних мереж, що симулював би роботу нейронів. Формуються вони з елементів, які відтворюють елементарні функції біологічного нейрона. Штучні нейронні мережі відтворюють певні властивості, які притаманні мозку людини. Вони навчаються на основі досвіду, узагальнюють свій досвід, здатні виділяти головне з інформації, що надходить.

Для моделей, побудованих за типом нейронних мереж людського мозку на мемристорах, характерно легке розпаралелювання алгоритмів і висока продуктивність. З людським мозком їх зближує важлива властивість, яка відсутня в простих електронних машинах: нейронні мережі працюють навіть за умови неповної інформації про навколишнє середовище, тобто, як і людина, вони можуть відповідати не тільки "так" або "ні", але і "не знаю точно, але скоріше так". Нейронним ме-

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

режам сьогодні під силу розпізнавання сигналів, мови, зображень, пошук даних, фінансове прогнозування, шифрування даних.

Нейромережний підхід використовується у великій кількості завдань - для кластеризації інформації з Інтернету, для імітації та моделювання складно влаштованого людського мозку, для розпізнавання образів і ін. Зараз продовжується вдосконалення методів синхронної роботи нейронних мереж на паралельних пристроях.

Нейронних мережі здатні до самонавчання, самоналаштування, гнучкості конфігурації, високої ефективності. Серед найбільш відомих сьогодні нейронних мереж виділяють мережі Хопфілда, нейронні мережі зі зворотним поширенням похибки і самоорганізовані карти. Здатність машини до самонавчання стане ключовим кроком в трансформації М2М. Бази даних та пошукові системи М2М можуть стати своєрідною енциклопедією для смарт-технологій. Аналізуючи пошукові запити, пошукові агенти вже здатні навчатися і підлаштовувати результати відповідно до потреб.

Предметом інтересу дослідників штучного інтелекту стають також питання про знання, які не пов'язані безпосередньо з технологічними питаннями і є досить автономними щодо розробок інтелектуальних систем. Це насамперед належить до питання «що є знання?». Автономний процес формування баз знань дає можливість відмовитися від дачі відповіді на це питання, адже загальноприйнятої відповіді не існує.

Термін «знання» придбав у цифровій реальності специфічний зміст, який Д.А.Поспелов характеризував таким чином: під знаннями розуміється форма подання інформації ЕОМ, якій притаманні такі особливості, як: а) внутрішня інтерпретованість (коли кожна інформаційна одиниця повинна мати унікальне ім'я, за яким система знаходить її, а також відповідає на запити, в яких це ім'я згадано); б) структурованість (включеність одних інформаційних одиниць до складу інших); в) зв'язність (можливість завдання тимчасових, просторових чи іншого роду відносин); г) семантична метрика (можливість завдання відносин, що характеризують ситуаційну близькість); д) активність (виконання програм ініціюється поточним станом інформаційної бази). Саме ці характеристики відрізняють знання в штучно створеній

інтелектуальній системі від даних - тобто вони визначають ту межу, за якою дані перетворюються в знання, а бази даних - в бази знань. Разом з тим такі риси, як внутрішня інтерпретованість, структурованість, зв'язність, семантична метрика і активність притаманні будь-яким більш-менш великим блокам людських знань і в цьому сенсі знання в комп'ютерній системі можна розглядати як модель того чи іншого фрагмента людського знання.

Дослідження розвитку М2М в напрямку створення штучного інтелекту перетворилось на ключову стратегію транснаціональних ІТ-компаній. Такі технічні гіганти, як Google, Facebook, Apple, Baidu і Microsoft, борються за залучення провідних умів у цій галузі. Відродження досліджень штучного інтелекту відбувається, здебільшого, через ряд революційних досягнень, пов'язаних з явищем, відомим як «глибоке навчання».

Одна з теоретичних властивостей нейронних мереж - здатність до навчання. Нейронні мережі не програмуються у звичному розумінні цього слова, вони навчаються. Можливість навчання - одна з головних переваг нейронних мереж перед традиційними алгоритмами. Технічно навчання полягає в знаходженні коефіцієнтів зв'язків між нейронами. У процесі навчання нейронна мережа здатна виявляти складні залежності між вхідними даними і вихідними, а також виконувати узагальнення. Це означає, що в разі успішного навчання мережа зможе повернути вірний результат на підставі даних, які були відсутні в навчальній вибірці, а також неповних та або «зашумлених», частково перекручених даних.

Хоча машинне навчання набуло домінуючу позицію у сфері штучного інтелекту, у нього все ж є недоліки. По-перше, воно займає дуже багато часу. По-друге, машинне навчання все ж не може бути істинним мірилом комп'ютерного інтелекту, воно використовує винахідливість людини і його абстрактні поняття, що дозволяють машині вчитися.

На відміну від машинного навчання, глибоке навчання в більшості випадків проходить неконтрольовано. Так, для нього необхідно створити великі нейронні мережі, які дозволяють комп'ютеру самостійно вчитися і «думати» без необхідності в безпосередньому втручанні людини. Глибоке навчання - набір алгоритмів машинного навчання, які намагаються моделю-

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

вати високорівневі абстракції в даних, використовуючи архітектури, що складаються з безлічі нелінійних трансформацій.

Глибоке навчання зовсім не схоже на комп'ютерну програму, вважає психолог і фахівець в сфері штучного інтелекту Гері Маркус. Як правило, комп'ютерний код пишеться відповідно з дуже суворими логічними етапами. На думку Маркуса у глибокому навчанні ми бачимо щось зовсім інше. У ньому немає безлічі інструкцій, які свідчать: якщо це істинно, то роби те.

Замість лінійної логіки глибоке навчання засноване на теоріях про те, як діє людський мозок. Програма складається з переплетених шарів взаємопов'язаних вузлів. Вона навчається шляхом зміни поєднань з'єднань між вузлами після кожного нового досвіду.

Глибоке навчання продемонструвало потенціал як основа для програмного забезпечення, здатного працювати на емоціях чи події, описані в тексті (навіть якщо вони не виражені явно), розпізнавати предмети на фотографіях і робити складні передбачення про можливе майбутнє поведінці людини.

Найбільш поширеними архітектурами, що дозволяють реалізувати "глибоке навчання" є:

- Deep Belief Network (DBN);
- Autoencoder;
- Convolutional neural network (CNN).

З перерахованих вище архітектур нейронних мереж найкращі результати в задачах ідентифікації об'єктів на статичному зображенні та їх класифікації продемонструвала архітектура CNN.

Ідея CNN або згортальної нейронної мережі полягає в чергуванні згортальних шарів і використання на виході повноз'язного шару нейронів, які в цілому утворюють ансамбль спеціалізованих нейромереж.

Рей Курцвейл, винахідник, що стоїть біля витоків таких напрямків комп'ютерної науки, як розпізнавання мови і символів, синтез музики, віртуальна реальність і штучний інтелект, у своєму інтерв'ю журналу «Computer world» виклав свій погляд на майбутнє штучного інтелекту. Згідно з його точки зору, викладеної також в його роботі «Сингулярність поруч: коли людство вийде за межі біології» [11]. Курцвейл пророкує, що врешті-решт людський і комп'ютерний інтелекти зіллються і не відрізнятимуться. Обсяг даних про мозок лю-

дини, які збирають вчені, з кожним роком збільшується майже вдвічі; у міру отримання все нових даних про конкретні ділянки мозку, люди досить скоро і швидко зможуть створювати докладні математичні моделі цих ділянок. За найскромнішими підрахунками, до кінця 2020-х років ми отримаємо дуже точну, детальну емуляцію роботи всіх ділянок мозку. Десять квадрильйонів (тобто 10<sup>16</sup>) операцій в секунду - цілком достатньо для того, щоб емулювати всі ділянки мозку людини. Люди навчаться об'єднувати переваги людського інтелекту, зокрема здатність розпізнавати образи, з тими можливостями, в яких, як уже очевидно, машини нас перевершують. Більше того, до кінця 2040х років один кубічний дюйм мікросхеми на нанотрубках буде в 100 мільйонів разів більш потужним, ніж мозок людини. Що стосується програмного забезпечення, то в 3030х роках машини зможуть звертатися до власного вихідного коду й удосконалювати його в рамках все більш прискореного циклу проектування. Тому врешті-решт ці системи стануть значно більш інтелектуальними, ніж люди, і будуть поєднувати в собі переваги біологічного та небіологічного інтелекту. До кінця 2020х років нанороботи (тобто машини розміром, порівняним з молекулою, що володіють функціями руху, обробки і передачі інформації, виконання запрограмованих команд, а також здатні до створення своїх копій, тобто самовідтворення) будуть володіти серйозними обчислювальними, комунікативними і роботизованими можливостями. Наприклад, нанороботичні білі кров'яні тільця зможуть завантажувати програмне забезпечення для конкретного патогенного мікроорганізму в тілі людини і руйнувати його буквально за кілька секунд, при тому що наші біологічні кров'яні тільця витрачають на це кілька годин. Нанороботи, спрямовані в мозок, дозволять людині значно розширити можливості свого інтелекту, ми зможемо вийти за межі біологічних меж і замінити наявне у нас «людське тіло версії 1.0» на кардинально оновлену версію 2.0, тим самим радикально збільшити тривалість життя. Небіологічних частина буде збільшуватися приблизно в 1000 разів за десять років, і біологічна частина в кінцевому підсумку стане дуже незначною. У ХХІІ столітті, за прогнозом Курцвейла, ми будемо використовувати можливості матерії та енергії на Землі і навколо неї

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

для підтримки обчислювальних процесів, і інтелектуальні обчислення почнуть розповсюджуватися по решті частини всесвіту. Однак при конструюванні нових машин ми завжди повинні усвідомлювати, які можливі наслідки від їх застосування [12].

В даний час активне дослідження нейронних мереж і глибинного навчання поступово переміщується з академічних інститутів та організацій в сторону корпорацій, що працюють в галузі інформаційних технологій. Тому є кілька причин. По-перше, корпорації розуміють, що Інтернет, а разом з ним і суспільство змінюються, і необхідно передбачити, бажано швидше за інших, що ж буде потрібно суспільству завтра. Розвиток технологій, пов'язаних з нейронними мережами, які демонструють порівнянне з людським якість вирішення низки завдань, таких як класифікація зображень чи розуміння семантики текстів, виглядає дуже перспективним з цієї точки зору. По-друге, дослідження глибинного навчання вимагає значних обчислювальних ресурсів, яких часто не вистачає в академічному середовищі, але зазвичай в достатній кількості у ІТ-компаній.

Амбітні проекти ІТ-компаній, претендуючи на розширення когнітивного потенціалу машин, створюють передумови для все більшої автономізації M2M. Важко спрогнозувати які будуть наслідки створення і існування штучного інтелекту в новому середовищі M2M. Зрозуміло тільки те, що створений людиною цифровий Голлем згодом зможе перевершити людські можливості.

У своїй книзі *Computer Power and Human Reason* Вейценбаум, автор програми Eliza, вказав на деякі потенційні загрози, з якими стикається суспільство у зв'язку з розвитком штучного інтелекту. Одним з найважливіших доводів Вейценбаума є те, що в результаті досліджень в галузі штучного інтелекту здається вже не такою неймовірною ідеєю про те, що люди являють собою окремих вид автоматів, а ця ідея призводить до втрати самостійності або навіть людяності [15].

«Корисність» штучного інтелекту можна розуміти як «розвантаження» людини: машині довіряється усе, що підзвітно лівій півкулі, а людина, що співпрацює з машиною, доповнює у роботі правою півкулі [3]. Якщо розцінювати це як вид «симбіозу», то, ймовірно, цим

можна пояснювати і відмічене нині вивільнення ірраціональності, що сьогодні привертає особливу увагу.

Екзистенціалізм поставив закономірне питання: чи не стоїть людина на порозі такої зміни довкілля, яка виявиться вища за її адаптаційні можливості. За М.Гайдеггером найзловісніший гість сучасності - бездумність. Свою думку М.Гайдеггер аргументує тим, що завдяки досягненням людського інтелекту пізнання усього доступно так швидко і так дешево, що в наступну мить так само швидко забувається. Людина рятується втечею від мислення, а така втеча і є основа для бездумності. Подібно до того, як осліпнути ми можемо тому, що були зрячими, так само ми можемо стати бездумними лише тому, що в самій основі свого буття людина має здатність до мислення.

Сучасна людина заперечує свою втечу від мислення, посилаючись на те, що об'єми наукових досліджень сьогодні значно перевершують минулі віки. М.Гайдеггер стверджує, що наукові дослідження пов'язані з приватним видом мислення, витрачаються на який дуже вигідно і корисно. Захоплюючись обчислюючим мисленням, людина рятується втечею від осмислюючого роздуму, мотивуючи це тим, що осмислюючий роздум парить над дійсністю, втратив ґрунт, не в силах допомогти справитись з повсякденними справами, є марним у вирішенні практичних питань [6].

### Висновки і перспективи подальших розвідок

Перший винахід надінтелектуальної машини стане останнім винаходом, який коли-небудь зможе зробити людина. Цьому "вибуху інтелектуальності" Вернор Віндж дав іншу назву - технологічна перевага. За його прогнозами, протягом тридцяти років люди отримають технологічні засоби для створення надлюдського інтелекту. Незабаром після цього ера людей закінчиться.

Стівен Хокінг зазначив, що примітивні форми штучного інтелекту, вже існуючі на сьогоднішній день, довели свою корисність, але він побоюється, що людство створить щось таке, що перевершить свого творця. Такий розум візьме ініціативу на себе і стане сам себе вдосконалювати з усе зростаючою швидкістю. Можливості людей обмежені занадто повільною еволюцією, ми не зможемо змагатися зі швидкістю машин і програємо [14].



## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

З точки зору глобальних ризиків, одна з найбільш критичних обставин у зв'язку з штучно створеною інтелектуальною системою, це те, що вона може посилити свій інтелект надзвичайно швидко. Очевидна причина підозрювати таку можливість - це рекурсивне самополіпшення. Така система стає розумнішою, в тому числі розумнішою щодо написання внутрішньої когнітивної функції, тобто здатна переписати свою існуючу когнітивну функцію, щоб вона працювала краще. Це зробить такі системи ще розумнішими, в тому числі розумнішим щодо завдання перероблення себе [8].

Успіх штучного інтелекту може стати початком кінця людської раси. Майже будь-яка технологія, потрапляючи в зловмисні руки, виявляє потенційні можливості для заподіяння шкоди, але коли йдеться про штучно створені інтелектуальні системи, виникає нова проблема, пов'язана з тим, що ці зловмисні руки можуть належати самій технології [9].

### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Індустрія наукових знань: вплив на соціокультурну сферу. Монографія / В. С. Лук'янець, О. М. Кравченко, О. Я. Мороз [та інші] – Київ : УкрСІЧ, 2015. – 407 с.
2. Кузин, Л. Т. Основы кибернетики: В 2-х т. Т. 2. Основы кибернетических моделей / Л. Т. Кузин – Москва : Энергия, 1979. – 584 с.
3. Ротенберг, В. С. Внутрішня мова і динамізм поетичного мислення / В. С. Ротенберг // Філософські науки. – 1991. – № 6. – С. 157–164.
4. Тьюринг А. Могут ли машины мыслить? / А. Тьюринг // С прил. ст. Дж. фон Неймана «Общая и логическая теория автоматов». – 2004. – 67 с.
5. Утехин, И. Взаимодействие с «умными вещами»: введение в проблематику / И. Утехин // Антропологический форум 17. – 2012. – С. 134–156.
6. Философия искусственного интеллекта // Материалы Всероссийской междисциплинарной конференции. – 2005. – Москва : ИФ РАН, 2005. – 400 с.
7. Швирков, О. І. Проблема штучного інтелекту і людиновимірність штучних інтелектуальних систем : дис... канд. філос. наук: 09.00.09 / О. І. Швирков. — Житомир, 2006. – 170 с.
8. Юдковски, Е. Искусственный интеллект как позитивный и негативный фактор глобального риска [Электронный ресурс] / Е. Юдковски. – Режим доступа:

На відміну від людини, на думку вчених, машини, що наділені штучним інтелектом, матимуть більше можливостей для отримання інформації. Так само як і людина вони можуть отримувати знання за допомогою слухових, зорових та сенсорних аналізаторів, а також з електронних носіїв (наприклад, Інтернет). Можливості пам'яті робота не залежать від настрою, умов середовища перебування та особистих можливостей, відповідно інтелектуальна машина здатна запам'ятовувати більше інформації та ефективніше виділяти якісну.

Перспективним напрямком продовження даного дослідження вбачається проведення належної соціогуманітарної експертизи потенційних загроз розвитку штучно створених інтелектуальних систем, етичних та аксіологічних передумов екзистенційної готовності людини до трансформацій життєсвіту під впливом штучно створених інтелектуальних систем.

<http://www.scribd.com/doc/6250354>. – Назва з екрана. – Перевірено 26.11.2015.

9. Floridi, L. Artificial Agents and Their Moral Nature / L. Floridi // Philosophy of Engineering and Technology. – 2014. – Vol. 17. – P. 185–212. doi: 10.5860/choice.37-0072.
10. Hayles, K. A. How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics / K. A. Hayles. – Chicago : University of Chicago Press, 1999. – 364 p. doi: 10.5860/choice.37-0072.
11. Kurzweil, R. The singularity is near: When humans transcend biology / R. Kurzweil. – New York : Penguin, 2006. – 672 p. doi: 10.2307/20031996.
12. Kurzweil, R. How to Make a Mind / R. Kurzweil // Futurist – 2013. – Vol. 47. – №. 2. – P. 14–17.
13. Turing, A. The essential Turing: seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life, plus the secrets of Enigma / A. Turing, Copeland B. J. – Oxford : Oxford University Press, 2004. – 600 p.
14. Stephen Hawking : Transcendence looks at the implications of artificial intelligence-but are we taking seriously enough? [Електронний ресурс] / S. Hawking, S. Russell, M. Tegmark, F. Wilczek // The Independent. – 2014. – Режим доступу : <http://goo.gl/Me91ha>. – Назва з екрана. – Перевірено 26.11.2015.
15. Weizenbaum, J. Computer power and human reason: From judgment to calculation / J. Weizenbaum. – Oxford : W. H. Freeman & Co, 1976. – 300 p. doi: 10.2307/3103715.

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

Ю. Д. ГЕНСИЦКИЙ<sup>1\*</sup>

<sup>1\*</sup>Национальный университет биоресурсов и природопользования Украины (Киев), эл. почта [yuraletter@gmail.com](mailto:yuraletter@gmail.com), ORCID: 0000-0003-0858-6862

### ФИЛОСОФСКО-АНТРОПОЛОГИЧЕСКОЕ ЗНАЧЕНИЕ РАЗВИТИЯ ИСКУССТВЕННО СОЗДАННЫХ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

**Цель.** Осмысление философско-антропологического значение развития искусственно созданных интеллектуальных систем требует проведения анализа социально-антропологического содержания проблемы развития межмашинного взаимодействия в контексте медиафилософского праксиса, антропологического содержания природы интеллекта, рассмотрения специфики концепта искусственно созданных интеллектуальных систем в среде М2М, развития социокогнитивных практик межмашинного взаимодействия, их социогуманитарного потенциала. **Методология исследования.** Реализация цели видится в использовании научно-теоретической базы медиафилософии, философской антропологии, медиафилософского подхода к пониманию общества, науки и техники, использовании публикаций по выбранной тематике исследования. **Научная новизна.** Рассмотрены концепт искусственно созданных интеллектуальных систем в аспекте социогуманитарного потенциала их становления и развития в среде М2М. Проанализирована проблема машинного обучения как технология трансформации М2М. Очерчены антропологические угрозы развития искусственно созданных интеллектуальных систем. **Выводы.** С точки зрения глобальных рисков, одно из самых критических обстоятельств в связи с искусственно созданной интеллектуальной системой, это то, что она может усилить свой интеллект чрезвычайно быстро. Очевидная причина подозревать такую возможность - это рекурсивное самоулучшение. Такая система становится умнее, в том числе умнее по написанию внутренней когнитивной функции, то есть способна переписать свою существующую когнитивную функцию, чтобы она работала лучше. Это делает такие системы еще умнее, в том числе умнее в аспекте переработки себя. Успех искусственного интеллекта может стать началом конца человеческой расы. Почти любая технология, попадая в злонамеренные руки, выявляет потенциальные возможности для причинения вреда, но когда речь идет об искусственно созданной интеллектуальной системе, возникает новая проблема, связанная с тем, что эти злонамеренные руки могут принадлежать самой технологии.

*Ключевые слова:* искусственно созданная интеллектуальная система, межмашинная среда, искусственный интеллект, нейронная сеть, машинное обучение

YU. D. GENSITSKIY<sup>1\*</sup>

<sup>1\*</sup>National University of Life and Environmental Sciences of Ukraine (Kyiv), e-mail [yuraletter@gmail.com](mailto:yuraletter@gmail.com), ORCID: 0000-0003-0858-6862

### PHILOSOPHICAL AND ANTHROPOLOGICAL IMPORTANCE OF DEVELOPMENT OF ARTIFICIALLY CREATED INTELLIGENT SYSTEMS

**Purpose.** Understanding the philosophical and anthropological importance of the development the artificial intelligence systems requires the analysis of the socio and anthropological content of intercomputer problems of interaction in the context of media philosophical praxis, anthropological maintenance of intellect nature, considering the specifics of the concept of artificial intelligence systems in the environment of M2M development of socio-cognitive practices of intercomputer interaction of social and humanitarian potential. **Methodology.** The implementation target is seen in the use of scientific and theoretical basis of the media philosophical, philosophical anthropology, the media philosophical approach to understanding society, science and technology, the use of publications on selected topics of research. **Scientific novelty.** The concept of artificial intelligence systems in the aspect of social and humanitarian potential of their formation and development in the environment of M2M was considered. The problems of machine learning as technology transformation M2M were analysed. The anthropological threats to the development of artificially created intelligent systems were defined. **Conclusions.** From the global risks point of view, one of the most critical circumstances due to the artificial intelligent system can strengthen its intelligence very quickly. The obvious reason for suspecting such an opportunity - a recursive self-improvement. Such system becomes smarter, including the intelligent writing of internal cognitive function, that the ability to rewrite their existing cognitive function to make it work better. This will make such systems more

© Генсіцький Ю. Д., 2015

## ФІЛОСОФІЯ НАУКИ І ТЕХНІКИ

intelligent, and smarter in terms of the processing itself. The success of artificial intelligence may be the beginning of the end of the human race. Almost any technology falling into malicious hands reveals the potential for harm, but when it comes to artificial intelligent system, there is a new problem with the fact that these malicious hands may belong to the technology itself.

*Key words:* artificial intelligence system designed, inter-machine environment, artificial intelligence, neural networks, machine learning.

### REFERENCES

1. Lukianets V.S., Kravchenko O.M., Moroz O.Ya. *Industriia naukovykh znan: vplyv na sotsiohumanitarnu sferu* [The industry of scientific knowledge: the impact on the world socioetra]. Kyiv, UkrSICH Publ., 2015. 407 p.
2. Kuzin L.T. *Osnovy kibernetiki: V 2-kh t. T. 2. Osnovy kiberneticheskikh modeley* [Fundamentals of Cybernetics: In 2 vo. Vol.2 . Fundamentals of cybernetic models]. Moscow,:Energiya Publ., 1979. 584 p.
3. Rotenberh V.S. Vnutrishnia mova i dynamizm poetychnoho myslennia [Inner speech and the dynamism of poetic thinking]. *Filosofski nauky – Philosophy sciences*, 1991, no. 6, pp. 157-164.
4. Turing A. *Mogut li mashiny myslit?* [Can the machines think?], 2004. 67 p.
5. Utekhin I. Vzaimodeystviye s «umnymi veshchami»: vvedeniye v problematiku [Interaction with "smart things": introduction to issues]. *Antropologicheskii forum 17* [Anthropological forum 17], 2012, pp. 134-156.
6. Filosofiya iskusstvennogo intellekta [The philosophy of artificial intelligence]. *Materialy Vserossiyskoy mezhdistsiplinarnoy konferentsii* [Materials of all-Russian Interdisciplinary Conf.]. Moscow, IF RAN Publ., 2005. 400 p.
7. Shvyrkov O.I. *Problema shtuchnogo intelektu i liudynovymirnist shtuchnykh intelektualnykh system Dok. Diss.* [The problem of artificial intelligence and luginbill artificial intelligent systems. Doc. Diss.]. Zhytomyr, 2006. 170 p.
8. Yudkovski Ye. *Iskusstvennyy intellekt kak pozitivnyy i negativnyy faktor globalnogo riska* [Artificial intelligence as a positive and negative factor in global risk]. Available at: <http://www.scribd.com/doc/6250354>. (Accessed 26.Nov.2015).
9. Floridi L. Artificial Agents and Their Moral Nature. *Philosophy of Engineering and Technology*, 2014, vol. 17, pp. 185-212. doi: 10.5860/choice.37-0072.
10. Hayles K.A. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago, University of Chicago Press Publ., 1999. 364 p. doi: 10.5860/choice.37-0072.
11. Kurzweil R. *The singularity is near: When humans transcend biology*. New York, Penguin Publ., 2006. 672 p. doi: 10.2307/20031996.
12. Kurzweil R. *How to Make a Mind*. *Futurist*, 2013, vol. 47, no. 2, pp. 14-17.
13. Turing A., Copeland B.J. *The essential Turing: seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life, plus the secrets of Enigma*. Oxford, Oxford University Press Publ., 2004. 600 p.
14. Hawking S., Russell S., Tegmark M., Wilczek F. Stephen Hawking : Transcendence looks at the implications of artificial intelligence-but are we taking seriously enough? *The Independent*, 2014. Available at : <http://goo.gl/Me91ha>. (Accessed 26.Nov.2015).
15. Weizenbaum J. *Computer power and human reason: From judgment to calculation*. Oxford, W. H. Freeman & Co Publ., 1976. 300 p. doi: 10.2307/3103715.

*Стаття рекомендована до публікації д. філ. н., проф. В. В. Хмілем (Україна)*

Надійшла до редколегії 23.11.2015

Прийнята до друку 03.12.2015