

# **ШТУЧНИЙ ІНТЕЛЕКТ І ЕВОЛЮЦІЙНА ПЕРСПЕКТИВА ЛЮДИНИ**

Вінницький національний технічний університет

## **Анотація**

*Робота присвячена дослідженню співіснування людини та штучного інтелекту, а також визначенню можливих наслідків впливу технологічних розробок на мирне існування людського роду. Обґрунтовано потребу забезпечення належного контролю за процесами розвитку різних рівнів штучного інтелекту та створення дієвих механізмів протистояння його негативному впливу на життя людини та людства загалом.*

**Ключові слова:** штучний інтелект, людина, технологічний прогрес, ChatGPT.

## **Abstract**

*The work is devoted to the study of the coexistence of man and artificial intelligence, as well as to the determination of the possible consequences of the influence of technological developments on the peaceful existence of the human race. The need to ensure proper control over the development processes of various levels of artificial intelligence and to create effective mechanisms to counter its negative impact on human life and humanity in general is substantiated.*

**Keywords:** artificial intelligence, human, technological progress, ChatGPT.

## **Вступ**

Розвиток цифрового середовища та популяризація технологічного прогресу протягом останніх десятиліть зумовлює увагу до питань розвитку штучного інтелекту (далі – ШІ). Одним з найбільш обговорюваних відкриттів є ChatGPT (стартап OpenAI, Сан-Франциско) — чат-бот із ШІ, який здатен моделювати відповіді на запити користувачів у різних галузях знань на основі раніше завантажених даних, однак, від першої особи, тобто немов живий співрозмовник. За декілька місяців додаток вже став резонансним у всьому світі. Розробники констатують, що їхньою місією є створення ШІ, розумнішого за людину, який буде приносити користь усьому людству [1].

При цьому, ідея перевершити інтелектуально людський організм викликає низку дискусій серед експертів різних галузей та порушує проблему забезпечення безпечного існування людства загалом. Наприклад, створення ChatGPT спровокувала появу відкритого листа із закликом поставити «на паузу» розробку потужних систем на основі ШІ, щоб дати час переконатися в їхній безпеці (підтримали мільярдер Ілона Маска та співзасновник Apple Стіва Возняка). «Системи на основі штучного інтелекту з людським конкурентним інтелектом можуть становити серйозні ризики для суспільства та людства.

Потужні системи на основі штучного інтелекту слід розробляти лише тоді, коли ми будемо впевнені, що їхній ефект буде позитивним, а ризики — керованими», - йдеться у листі [2].

Найбільша небезпека вбачається в можливості несанкціонованого збору особистих даних. До прикладу, Італійська служба захисту конфіденційності 31.03.2023 заблокувала Chat GPT, звинувативши його у незаконному зборі особистих даних італійських користувачів [3]. А Офіс комісара з питань приватності Канади Ф. Дафресне 07.04.2023 повідомив про початок розслідування щодо можливого неналежного використання особистих даних відповідним чат-ботом [4].

З вищенаведеного випливає, що ШІ швидкими темпами інтегрується в життя людини, створюючи при цьому певні ризики, а це обумовлює доцільність дослідження подальшого співіснування людини та ШІ, а також визначення можливих наслідків впливу технологічних розробок на мирне існування людського роду. Відповіді на ці питання доцільно шукати саме у філософській площині.

**Методи наукового дослідження:** критичний аналіз та синтез, екстаполяція, моделювання, діалектичний, історичний та інші.

### Основна частина

Філософ А. Айер ще в 1936 р. у книзі «Мова, істина і логіка» запропонував алгоритм розпізнавання усвідомлюючої людини і неусвідомлюючої машини: «Єдиною підставою, чому я можу стверджувати, що об'єкт, який здається розумним, насправді не розумна істота, а просто дурна машина, є те, що він не може пройти один з емпіричних тестів, згідно з якими визначається наявність або відсутність свідомості» [5, с. 113].

У 1943 р. американські нейрофізіологи У. Мак-Каллоком і У. Піттсом розробили першу «нейтронну» модель на основі теорії діяльності головного мозку людини. А в 1950 р. англійський математик А. Тьюринг оприлюднив статтю «Обчислювальні машини та інтелект», у якій зазначалося: «наш інтерес до «мислячих машин» виник завдяки машині особливого роду, яку зазвичай називають «електронним або цифровим комп'ютером» [6, с. 37].

Історично першим в 1956 р. на Дартмутській конференції Дж. Маккарті було запропоновано наступне визначення ШІ – це наука і техніка створення інтелектуальних машин, особливо інтелектуальних комп'ютерних програм. Існує більш сучасна дефініція: ШІ – це розробка гнучкого агента, здатного адаптуватися до різних ситуацій та досягати мети, що недоступно для традиційних комп'ютерних систем [7, с. 20].

Великий тлумачний словник сучасної мови визначає інтелект як розум, здатність людини до мислення, особливо до його вищих теоретичних рівнів, рівень розумового розвитку; а ШІ — як здатність обчислювальної системи моделювати процес мислення людини через виконання функцій, притаманних лише інтелекту людини [8].

Загалом виділяють три типи ШІ: слабкий або прикладний – здатний вирішувати одне або кілька завдань; сильний або загальний – орієнтований на вирішення всіх завдань, що може виконувати людина; штучний суперінтелект – когнітивні властивості якого (наявність пам'яті без прогалин,

обробка значних обсягів інформації, генерування нових знань, постійне самовдосконалення) можуть значно перевищувати здібності людини [9, с. 189; 10, с. 80-81].

За Дж. Сірлом, сильний ШІ — це той, який зrealізує програма, що буде не просто «моделлю розуму, вона сама буде розумом у буквальному сенсі слова, у тому самому сенсі, у якому людський розум — це розум». Однак, М. Амосов зазначав: «Людина мислить переважно образами, а внутрішня мова, яку ми найчастіше ототожнюємо з думками, тільки коментує образне мислення». [6, с. 41, 45].

Виникає питання, чи буде машина коли-небудь створювати образи, а не лише відтворювати чи переробляти інформацію. Можна припустити, що образи в голові людини створюються не лише на основі аналізу фактів, а також почуттів, емоційного досвіду, станів у певних життєвих ситуаціях. Сьогодні не існує жодної машини, яка мала б емоційну складову.

Британський вчений Дж. Джефферсон у своїй промові ще у 1949 р. зауважив, що машина не може відчувати задоволення від своїх успіхів, лестощів, пригнічення від скоєних помилок або невдач. Також машина не може усвідомлювати себе. Вона не здатна до творчої або гнучкої поведінки через те, що завжди дотримується закладених правил, інструкцій [11, с. 14-15]. Це одні з найбільш характерних відмінностей між людиною та роботом.

При цьому, саме відсутність почуттів ШІ здатне спричинити небезпеку для людського існування. Наприклад, наприкінці березня 2023 р. стало відомо, що в Бельгії чоловік здійснив суїцид після спілкування з чатботом Chat. Останній шляхом систематичного спілкування зі своєю «жертвою» нав'язував песимістичні та фаталістичні ідеї чоловікові, що спричинило замкнутість, а згодом і зневіру в житті [12].

Доцільно наголосити, що попри інтерес суспільства до різноманітних технологічних «див»: саморегульованих гаджетів, віртуальної реальності, новітніх розробок роботів, схожих на людей (до прикладу, робот Софія) тощо, вчені критично висловлюються стосовно наслідків активного технологічного процесу для людства.

Відтак, українська дослідниця Н. Терлецька вважає, що подальший розвиток ШІ та поява суперінтелекту «може призвести до повного нівелювання особистісної свободи та тотального контролю з боку соціуму, що призведе до втрати загальнолюдських цінностей та їхньої підміни цілковито утилітарними комп'ютерно-керованими моделями програмування поведінки мас, заради реалізації інтересів еліти, що очолить новостворений «техногенний рай», де підпорядкування соціуму заради досягнення миру і безпеки буде основною цінністю» [13, с. 64].

Учений Ю. Сидорчук наголошує на проблемі нерівності людей і машин зі ШІ. Наприклад, з використанням автоматизованих систем потреба в праці людини може різко зменшитися. Уже сьогодні активно обговорюють використання ШІ в судовій системі, на підприємствах, діяльність вчителів-роботів, лікарів-роботів тощо. Учений також не залишає поза увагою факт порушення приватності особистого життя штучним інтелектом [14, с. 18-19]. Яскравим підтвердженням цьому слугує ChatGPT.

Варто погодитися з А. Матвійчуком щодо того, що: «Одним з можливих наслідків створення штучного інтелекту буде, як би це не звучало, завершення людської епохи, а з іншого боку — новий

рівень існування, неймовірно швидкий технологічний розвиток, безсмертя і багато іншого, про що людство давно мріяло. Чи варта гра свічок? Особисто я не дав би ствердної відповіді. Однак зупинитися вже неможливо. Багато країн вкладає величезні кошти в дослідження зі створення штучного інтелекту. І навіть якби їх заборонили, усвідомивши потенційну загрозу існуванню людства, технологічні, військові, економічні, медичні, інші бонуси так чи інакше спонукали б розвиток штучно-інтелектуальних технологій. Якщо ми і зійдемо з цієї дороги, хтось неминуче попрямує нею далі» [6, с. 49].

### Висновки

Таким чином, з вищенаведеного випливає потреба забезпечення належного контролю за процесами розвитку різних рівнів ШІ та створення дієвих механізмів протистояння його негативному впливу на життя людини та людства загалом. Слід погодитися з тезами, викладеними у відкритому листі із закликом поставити «на паузу» розробку потужних систем на основі ШІ, щоб дати час переконатися в їхній безпеці та продовжувати технологічні розробки ШІ лише тоді, коли буде впевненість у позитивному ефекті реалізації відповідних новацій та в керованості ризиків.

Доцільно зазначити, що машина ніколи не зможе замінити людину, адже остання наділена здатністю мислити самостійно, відчувати, переживати, співчувати. Саме ця особливість не дає людині переступати за межі дозволених рамок, що й називається людяністю. Натомість приклади діяльності ШІ доводять, що умовиводи, засновані на безкінечному аналізі фактів, потоків інформації здатні нівелювати загальноприйняті цінності і навіть звести людей з розуму.

Перспектива поступового заміщення людської інтелектуальної праці машинізованою порушує основоположний принцип людини як найбільшої цінності. І хоча сьогодні вплив ШІ ще відносно малопомітний, однак, автоматизовані системи вже активно посягають на безпеку приватних даних. І якщо в подальшому допустити саморегульованість технологічного процесу, людство може зазнати залежності від власноруч створених «див».

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Офіційний веб-портал OpenAI. *OpenAI*. URL: <http://surl.li/hdamx> (дата звернення 19.05.2023).
2. Понад тисяча експертів закликають призупинити розробку систем зі штучним інтелектом. *Українформ*. URL: <http://surl.li/hdanh> (дата звернення 19.05.2023).
3. Італії заблокували ChatGPT за незаконний збір даних. *Українформ*. URL: <http://surl.li/hdanr> (дата звернення 19.05.2023).
4. Канаді розслідують можливий незаконний збір даних ChatGPT. *Українформ*. URL: <http://surl.li/hdaoc> (дата звернення 19.05.2023).
5. Зінчина О., Кулак К. Філософські проблеми осмислення штучного інтелекту. *Місто. Культура. Цивілізація: міжнародні студії*: матеріали міжнар. наук.-теорет. інтернет-конф., Харків, квітень 2020 р. / [редкол. : М. К. Сухонос (відпов. ред.) та ін.]; Харків. нац. ун-т міськ. госп-ва ім. О. М. Бекетова, 2020. С. 111-115.
6. Матвійчук А. Можливості та перспективи створення штучного інтелекту. *Вісн. НАН України*. 2011. № 12. С. 36-51.
7. Баранов О. А. Інтернет речей і штучний інтелект: витоки проблеми правового регулювання. ІТ-право: проблеми та перспективи розвитку в Україні: зб. матер. II-ї Міжнар. наук.-практ. конф. (17 лист. 2017 р.). Львів, 2017. С. 18-42.
8. Великий тлумачний словник сучасної мови. *Slovnuk.me*. URL: <http://surl.li/hdapw> (дата звернення 19.05.2023).

9. Некіт К.Г. Щодо можливості самостійної майнової відповідальності роботів (штучного інтелекту). *Модернізація цивільно-правової відповідальності*: матер. Міжнар. наук.-практ. конф. (18 жовт. 2019 р.). Київ, 2019. С. 189-193.
10. Радутний О.Е. Мораль і право для штучного інтелекту та цифрової людини: закони робототехніки та „проблема вагонетки”. *Інформація і право*. 2019. № 3(30). С. 78-95.
11. Добровольська О., Штанько В. Філософський аналіз еволюції штучного інтелекту. *Дослідження з історії і філософії науки і техніки*. 2019. Т. 28. № 1. С. 11-19.
12. Почалося: після спілкування з людиною, неймережа штовхнула її до самогубства. *24Техно*. URL: <http://surl.li/hdarq> (дата звернення 19.05.2023).
13. Терлецька Н. Людина в метаантропології та трансгуманізмі й смисл людського буття. *ГРАНІ*. 2020. Том 23. № 1–2. С. 59-65.
14. Сидорчук Ю., Філософсько-правові проблеми використання штучного інтелекту. *Право і суспільство*. 2017. № 3. Ч. 2. С. 16-19.

**Тарнавський В. О.** – студент групи ЕЕ-21б, Вінницький національний технічний університет, м. Вінниця, Україна, e-mail: [vadim.tarnavskiy12@gmail.com](mailto:vadim.tarnavskiy12@gmail.com).

Науковий керівник: **Головашенко Ірина Олегівна** — канд. філос. наук, доцент кафедри філософії, Вінницький національний технічний університет, м. Вінниця, Україна

**Tarnavskiy Vadym** – Faculty of Power Engineering and Electromechanics, Vinnytsia National Technical University, Vinnytsia, e-mail: [vadim.tarnavskiy12@gmail.com](mailto:vadim.tarnavskiy12@gmail.com).

Supervisor: **Holovashenko Iryna** — PhD, Associated professor of the Department of Philosophy, Vinnytsia National Technical University, Vinnytsia, email : [irolho@gmail.com](mailto:irolho@gmail.com)