

Отже, філософія технологій грає важливу роль у розумінні глибинних аспектів людського існування в умовах технологічної революції, викликаючи нас до високоморального та етичного вдумливого підходу до створення та використання нових технологій.

Слід підкреслити, що сучасні технології стали невіддільною частиною нашого існування, впливаючи на його глибинні аспекти. Вивчення філософії технологій приводить до розуміння широкого спектру викликів і можливостей, які вони приносять у наше повсякденне життя та ментальність.

Технології впливають на наше сприйняття себе та інших, розширюючи межі ідентичності через віртуальні платформи та інтерактивні інтерфейси. Вони викликають глибокі філософські питання щодо свободи, прав та етики, розвиваючи нові норми та цінності в цифровому віці.

Приватність та безпека даних стають критичними елементами обговорення, визначаючи межі між технологічними можливостями та захистом особистого простору. Це вимагає ретельного балансу між інноваціями та збереженням основних прав людини.

З іншого боку, технології вносять нові можливості у сферу освіти, сприяючи самовдосконаленню та здобуттю нових знань. Вони переписують правила економіки та робочого процесу, викликаючи роздуми про майбутнє праці та соціальних відносин.

Враховуючи вищенаведене, поява технологій в житті людини додає нових філософських аспектів, пов'язаних із сучасними технологіями, та підкреслює необхідність глибокого розуміння цих аспектів для того, щоб забезпечити етичний та справедливий розвиток технологій в інтересах всього суспільства.

#### **Література:**

1. Штучний інтелект : вчора, сьогодні, завтра : лекція. – Режим доступу : [//www.victoria.lviv.ua/html/ai/lecture/00.doc](http://www.victoria.lviv.ua/html/ai/lecture/00.doc).
2. Бусол О. Ю. Потенційна небезпека штучного інтелекту / О. Ю. Бусол // Інформація і право. - 2015. - № 2. - С. 121-128. - Режим доступу: [http://nbuv.gov.ua/UJRN/Infpr\\_2015\\_2\\_21](http://nbuv.gov.ua/UJRN/Infpr_2015_2_21)
3. Зуев В. М. Поняття технології в сучасній філософії / В. М. Зуев // Вісник НТУУ «КПІ». Філософія. Психологія. Педагогіка : збірник наукових праць. – 2010. – № 3(30). – С. 23–26. – Бібліогр.: 14 назв.

#### **Проблема штучного інтелекту в експериментах: «гра в імітацію» Тюрінга та «китайська кімната» Серля**

*Ілля КОСТРОВ, здобувач III освітньо-наукового ступеня (PhD)*

*Національний аерокосмічний університет ім. М.Є. Жуковського*

*«Харківський авіаційний інститут»*

На сьогодні навряд чи знайдеться людина, яка б чула про комп'ютери, проте, не чула про Штучний Інтелект. Улітку 1956 року в Університеті Дартмута у США пройшла перша робоча конференція науковців з проблематики штучного інтелекту. Організаторами семінару були Джон Маккарті, Марвін Мінські, Клод Шеннон і Натаніель Рочестер. Вони запросили всіх відомих американських дослідників, так чи інакше пов'язаних з питаннями теорії управління, теорії

автоматів, нейронних мереж, теорії ігор і дослідженням інтелекту. Саме тоді і з'явився сам термін «Artificial Intelligence» - «Штучний Інтелект». З того часу не вщухають дискусії, як у науковій спільноті, так і у науково-фантастичній, художній літературі та кінематографі, які порушують усе більше нових питань: від чи можна створити машину, яка імітує/відтворює людську діяльність, чим відрізняється людина від створеної машини до етичних питань правомірності та обмеженості таких пошуків. Спроби вирішити ці та багато інших питань, ми бачимо у фільмах, читаємо у книгах, при тому, кожна з запропонованих варіацій зводиться до спроби віднайти відповідь на споконвічне філософське питання: Що таке людина?

Вважається, що засновником наукової думки про штучний інтелект в області комп'ютерних наук, є Алан Тюрінг, людина, яка приймала участь в створенні перших цифрових програмованих обчислювальних машин – *машини Тюрінга*. У своїй статті “Computing Machinery and Intelligence” (1950) він поставив за мету розглянути питання, чи є комп'ютер розумним у людському сенсі слова та запропонував тест задля визначення здатності машини імітувати інтелектуально обумовлену поведінку, що тотожна до поведінки людини і яку неможливо відрізнити від поведінки людини. Спираючись на статтю пропонує розглянути його погляди на проблематику питання та звернутися до критики його поглядів, яка представлена Джоном Роджерсом Серлем, автором концепції уявного експерименту під назвою «китайська кімната».

У своїй роботі Тюрінг пише: «Пропоную розглянути питання "Чи можуть машини мислити?" Це слід почати з визначення значення термінів «машина» і «мислити». Визначення можуть бути сформульовані таким чином, що б, наскільки це можливо, відображати звичайне вживання слів, але таке ставлення є небезпечним (не однозначним), якщо значення слів «машина» та «мислити» будуть визначатись, у результаті дослідження того, як вони зазвичай використовуються...» [1].

Таким чином, А. Тюрінг вказує на одну з суттєвих проблем філософії, яку певним чином намагалися вирішити Рене Декарт у роботі «Роздуми про метод» (1637), Дені Дідро у роботі «Філософські думки» (1746), Альфред Айєр у книзі «Мова, істина та логіка» (1936) та інші.

Новизна запропонованого Тюрінгом вирішення питання полягає у експериментальному тесті - «*гра у імітацію*». «У неї грають троє людей: чоловік (А), жінка (В) і слідчий (С)... Мета гри для допитувача полягає в тому, щоб визначити, хто з двох інших гравців є чоловік, а хто жінка. Він знає інших гравців під псевдонімами Х і Y, а наприкінці гри він має сказати або «Х – це А, а Y – це В», або «Х – це В, а Y – це А». Допитувачу дозволяється ставити запитання (А) та (Б)... Мета гравця (А) змусити гравця (С) зробити неправильну ідентифікацію... Мета гри для третього гравця (В) – допомогти слідчому... Тепер ми ставимо запитання: Що станеться, коли машина візьме участь у цій грі у ролі гравця (А)? Чи буде допитувач приймати неправильні рішення так само часто, як

коли така гра ведеться між чоловіком і жінкою? Ці запитання замінюють наше початкове: Чи можуть машини думати?» [1].

Постановою такого мисленевого, на той час, експерименту Тюрінг акцентує увагу на питанні спроможності цифрової системи до імітації інтелектуальної діяльності. На сьогодні цей експеримент проводиться практично: якщо оцінювач не може достовірно відрізнити машину від людини, можна сказати, що машина пройшла тест. При цьому, результати тесту залежатимуть не від здатності машини давати правильні відповіді на запитання, а лише від того, наскільки її відповіді схожі на відповіді людини.

Тест Тюрінга став важливим кроком у розвитку концепції Штучного Інтелекту. Пізніше цей експеримент став об'єктом критики з боку різних дослідників, проте залишається впливовим і зараз. Критика тесту Тюрінга дозволила висвітлити нові напрямки розвитку Штучного Інтелекту та визначати обмеженості його використання.

Одним з перших, хто звернувся до критики тесту Тюрінга, став Джон Роджерс Серль. У своїй роботі “Is the brain’s mind a computer program?” він змінює акценти у постанові питання та зазначає наступне: «Чи може машина мислити? Чи може машина мати свідомі думки в тому ж сенсі, що й ми з вами?» Отже, питання про імітацію мислення замінюється на питання про імітацію свідомості, як то ставив у своїй роботі Д. Дідро, який писав, що якщо знайдеться папуга, яка буде давати відповіді на запитання без вагань, то він назве її розумною. Джон Серль ставить питання, чи можна давати відповіді на запитання, які інші будуть сприймати як усвідомленні, розумні, навіть не розуміючи самого питання?

Експеримент Д. Серля отримав назву «китайська кімната». Він описує його так: «Для прикладу, візьміть мову, яку ви не розумієте. У моєму випадку я не розумію китайську... А тепер припустімо, що я опинився у кімнаті, де повні кошики китайських символів. Припустімо також, що мені надали інструкцію англійською (відомою мені) мовою для зіставлення китайських символів з іншими китайськими символами. Правила ідентифікують символи виключно за їх формою і не вимагають, щоб я розумів будь-який із них. Правила можуть містити такі речі як: «Візьміть знак N з кошика №1 та розташуйте його поруч із знаком M з кошика №2». Уявіть собі, що люди за межами кімнати, які розуміють китайську мову, передають у кімнату невеликі групи символів, а я у відповідь маніпулюю символами згідно з інструкцією та повертаю інші групи символів...» [2]. Д. Серль зазначає, що у такий спосіб можна отримати «відповіді» на запитання, що не відрізняються від відповідей носія китайської мови та створити уявлення про її знання, тобто, «пройти» тест Тюрінга. Він пише: «Не зважаючи на це, я все одно абсолютно не знаю китайської... Подібно до комп'ютера, я маніпулюю символами, але не надаю їм значення» [2]. Таким чином, Д. Серль звертає увагу на те, що тест Тюрінга «гра у імітацію» не є вичерпним засобом перевірки притаманності об'єкту властивостей

інтелектуальної активності. Тест може підтвердити тільки спроможність імітації людської поведінки, яку можна досягти як засобами з проявами інтелектуальної активності так і без.

Отже, Д. Серль, порушує питання про те, що розуміння об'єкту людиною не обмежується формальним описом. В той час як цифрові обчислювальні системи спроможні лише до формального опису без розуміння їх значення. Така позиція стане підґрунтям концепції «слабкого інтелекту». Тести А. Тюрінга та Д. Серля порушили цілу низку суттєвих питань щодо можливостей та обмеженості використання Штучного Інтелекту, які є актуальними темами дослідження у сучасних концепціях.

#### **Література:**

1. "COMPUTING MACHINERY AND INTELLIGENCE" by A. M. Turing <https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf>
2. "Is the brain's mind a computer program?" by John R. Searle [https://courses.edx.org/asset-v1:MITx+24.09x+3T2015+type@asset+block/1\\_searle\\_mind\\_s\\_brain\\_a\\_computer\\_program\\_.pdf](https://courses.edx.org/asset-v1:MITx+24.09x+3T2015+type@asset+block/1_searle_mind_s_brain_a_computer_program_.pdf)  
[https://www.cs.princeton.edu/courses/archive/spr06/cos116/Is\\_The\\_Brains\\_Mind\\_A\\_Computer\\_Program.pdf](https://www.cs.princeton.edu/courses/archive/spr06/cos116/Is_The_Brains_Mind_A_Computer_Program.pdf)

### **Штучний інтелект і свідомість**

*Артур ІВАНОВ, здобувач освіти*

*Науковий керівник – Наталя БІЛЬЧУК, канд. філос. наук, доцент*

*Національний аерокосмічний університет ім. М.С. Жуковського*

*"Харківський авіаційний інститут"*

Аристотель вважав, що розум - це властивість виключно людська. Так вважали до 20 століття. Це час, коли з'являються перші комп'ютери, які здатні обчислювати. До цього моменту подібне могла робити тільки людина. Пізніше, 1956 року, на Дартманському семінарі, вперше вживається термін «штучний інтелект». Одним із учених, Марвіном Мінським, було запропоновано визначити Штучний Інтелект як машину, здатну виконувати завдання, що вимагають інтелектуальних зусиль людини.

У 1980 році Джон Серл в одній зі своїх робіт вводить два терміни «сильний» і «слабкий» штучний інтелект. «Слабкий» - лише корисний інструмент людини, а «сильний» має не просто вирішувати інтелектуальні завдання, а й мати свідомість у людському розумінні. І ось з приводу свідомості філософія міркувала протягом усієї своєї історії. Платон під поняттям «душа» мав на увазі саме свідомість у її метафізичному сенсі. Середньовічні богослови стверджували, що душа це неповторне творіння бога.

До того ж 1950 року знаменитий учений Алан Тьюрінг публікує у філософському журналі статтю, в якій поставить питання, актуальне донині. Чи може машина здійснювати дії, які не відрізняються від усвідомлених? Саме той самий тест Тюрінга має відповісти на поставлене запитання. Під час тесту випробовуваний повинен вести анонімне листування з кількома