

УДК 159.9

DOI <https://doi.org/10.32782/2707-0409.2023.2.2>

Кириченко В. В.

доктор психологічних наук, доцент, професор кафедри соціальної та практичної психології

Житомирський державний університет імені Івана Франка

e-mail: viklorenc@gmail.com

ORCID ID: 0000-0002-5210-6304

СОЦІАЛЬНО-ПСИХОЛОГІЧНА ПАРАДИГМА РОЗУМІННЯ ЕВОЛЮЦІЇ ШТУЧНОГО ІНТЕЛЕКТУ

У статті здійснюється історико-методологічний екскурс до проблеми розвитку технологій штучного інтелекту, його впливу на соціокультурну діяльність людини. Описано основні етапи створення ШІ, передумови його виникнення, описані вихідні положення розробників та принципові моделі взаємодії людини з розумними інформаційними системами. У межах соціально-психологічної парадигми еволюції технологій штучного інтелекту ми виокремили три етапи розвитку імітації психічної діяльності людини, такі як: імітація інтелектуальних процесів, імітація емоцій та імітація творчої діяльності. Соціально-психологічна парадигма еволюції штучного інтелекту базується на тому, що розвиток інтелектуальних систем має соціально-психологічні наслідки, які є результатом впливу ШІ на суспільство та взаємодію між людьми. Ця парадигма розглядає не лише технічні аспекти розвитку ШІ, але й вплив цих технологій на міжособистісні відносини, етику, культуру та інші соціальні явища. Вона враховує психологічні аспекти сприйняття інтелектуальних систем людьми та взаємодію з ними в різних сферах життя. Соціально-психологічна парадигма еволюції штучного інтелекту визнає, що розвиток ШІ повинен бути врегульований не лише з технічної точки зору, а й з урахуванням його соціокультурних та психологічних наслідків впливу на людську культуру. Одним із соціально-психологічних наслідків взаємодії людини із системами штучного інтелекту є формування специфічних страхів стосовно машин, атрибуція їм необґрунтованих властивостей та суто людських рис. Але такі тенденції взаємодії з розумними інформаційними системами не змінили загального культурного тренду на інтеграцію штучного інтелекту у широке коло суспільних процесів. Ознаками продуктивної взаємодії з розумними інформаційними системами є розуміння їх користі у тісній взаємодії з людиною як інструмент полегшення рутинної праці.

Ключові слова: штучний інтелект, мислення, творчість, нейронні мережі, інформаційне суспільство.

Постановка проблеми. Останніми роками (2020–2023) інформаційно-комунікаційні технології, які пов'язані з виконанням аналітико-синтетичних операцій, що по своїй суті стоїть в основі імітації людського мислення, переживають чергове піднесення розвитку. Цього разу цікавість громадськості до ШІ (штучного інтелекту) пов'язана зі значним удосконаленням вихідних алгоритмів обробки інформації та завершенням машинного навчання на відкритих базах даних, що дає можливість керувати велетенськими масивами даних, піддаючи їх не меншій за кількістю обробці спеціальними алгоритмами. Штучний інтелект став помічником не лише студентів, але й буквально кожної людини, пропонуючи «розумні»

технології у всіх сферах, де присутня інформаційна діяльність. Проблема взаємодії людини з цифровими технологіями, як і у попередні роки, полягає у тому, що, власне, людство створює «досконалу машину» чи «імітацію людини», яка має бути максимально схожою на неї візуально, що інтенсивно зараз розробляється, і психологічно. У своєму розвитку штучний інтелект пройшов непросту дорогу від локального вирішення інтелектуальних завдань до суцільної інтеграції у суспільні процеси, що стало сильним викликом для людства у плані розуміння того, що (чи хто) проєктувався під багатозначною назвою «штучний інтелект».

Аналіз останніх досліджень і публікацій. Історія розумних машин починається з 1921 року, коли письменник-фантаст Карл Чапек публікує свою комедійну п'єсу «Rossumovi univerzální roboti», яка і дала назву терміну «робот», що так інтенсивно вживається зараз у повсякденні. За комедійним футуристичним змістом стояла глибока соціальна проблема того, що робот (гуманоїд) може подолати людські недоліки, фактично започаткувати новий виток еволюції людини. Робот у своїх витках проєктувався не просто «розумним», він мав бути максимально «людиноподібним», адже тільки людина може бути наділена інтелектуальними рисами та здатною до цілеспрямованої діяльності. У роботах Роберта Сапольського прослідковується думка про те, що еволюція homo sapiens у майбутньому вийде за межі біологічного пристосування, а тому пришвидшиться за рахунок розумного симбіозу з машинами [0]. У коридорах прогресу людство намагається «обдурити Бога» в розумінні постійного прагнення до вдосконалення за рахунок пристосування рукотворного світу, а інформаційний світ є найбільшим втіленням людського генію на сьогодні. Перша «холодна зима» штучного інтелекту, яка супроводжувалася певним послабленням інтересу до розумних технологій, відбулася в 1950 році, після проведення експерименту Алана Тюрінга, відомого всім як «гра в імітацію» [5]. Багато дослідників вважають, що саме цей експеримент є відправною точкою появи штучного інтелекту як предмета наукового дослідження. Його автор намагався знайти відповідь на питання: «Чи може машина мислити як людина? І які межі існують у машини?». Теоретично відповідь на це питання стояла у назві експерименту, адже імітація не має на меті створення «точного» відповідника. «Тест Тюрінга» описаний у статті «Computing Machinery and Intelligence» («Обчислювальні машини та інтелект»). Його мета полягала в оцінці здатності машини виконувати інтелектуальні завдання та імітувати людський розум. Важливим аспектом тесту є те, що обмеження не ставиться на конкретні види інтелектуальної діяльності. Тест не вимагав вирішення конкретних завдань, але завдання полягає у вмінні вести розмову, висловлювати думки та реагувати на запитання так, щоб інша людина не змогла визначити, хто її співрозмовник – людина чи машина. Важливо відзначити, що тест Тюрінга не залишився без своїх критиків. Часто висувається аргумент, що розв'язання тесту не обов'язково свідчить про інтелект або свідомість машини в науково-психологічному розумінні цих термінів. Однак цей тест залишається важливим концептуальним внеском

Тюрінга у сферу розуміння відмінностей між машинним та людським інтелектом. **Найважливішим внеском експерименту Алана Тюрінга в розуміння проблеми взаємодії зі штучним інтелектом є вивчення здатності диференціювати інтелектуальні дії людини від їх імітації.** І поки людина справляється з цим завданням добре.

Мета статті. Ми виокремлюємо три основні періоди розвитку штучного інтелекту, спираючись на вихідні положення соціально-психологічного дискурсу розуміння цієї проблеми, кожен з яких позначає окремий період їх розвитку у напрямі імітації психічної діяльності людини. Перехід між ними був детермінований зміною концептуальних поглядів суспільства на роль та значення штучного інтелекту в еволюції соціокультурних відносин (рис. 1). У статті аналізують психологічні передумови виникнення кожного з етапів розвитку штучного інтелекту та соціально-психологічні особливості взаємодії з ним людини.

Результати дослідження. Мислення людини та імітація людського мислення машини відрізняється рамками логіки, а найголовніше – здатністю людей мислити, не маючи конкретних логічних рамок. Імітація мислення людини штучним інтелектом відбувається через відтворення програмою ключових аспектів когнітивних функцій людського мислення. Це може включати у себе розуміння мови, прийняття рішень, вирішення проблем, самонавчання та інші аспекти інтелекту. На сьогодні ШІ може використовувати різні підходи, такі як машинне навчання, нейронні мережі та обробка природної мови, для моделювання імітації мислення.

Найголовнішим, на нашу думку, атрибутом людського мислення є здатність до усвідомлення «абсурду» як результату існування парадоксів логіки. Будь-який результат для машини є результатом дії логічно зумовленого алгоритму, тому якщо алгоритм відбувся (операцію виконано), то цей результат не може містити логічної суперечності. Основою людського мислення є пошук суперечностей та природне прагнення до генерації абсурду, який на всіх етапах суспільної еволюції відігравав вирішальне значення науково-технічного, культурного та суспільного прогресу. У цьому плані здатність розуміти гумор у будь-якій його формі є швидким, але дуже дієвим способом ідентифікації людини в інформаційному просторі. Деякі прийоми на основі розуміння «абсурдів» уже використовуються



Рис. 1. Хронологія еволюції технологій штучного інтелекту у контексті соціально-психологічної парадигми

на інформаційному фронті гібридного протистояння України проти росії з метою виявлення «ботів» (штучно створених профілів користувачів глобальних інформаційних систем).

Для вирішення багатьох проблем взаємодії із сучасними технічними системами, особливо після появи специфічних «технофобій», варто розуміти, що жодна машина не може мислити як людина та не може діяти як людина (з позицій відсутності мотивації та цілепокладання як такого). Спроба створити розумну, а саме головне людиноподібну істоту призвела до того, що в окремої частини людського суспільства почали формуватися численні антропоморфізми щодо різноманітних технічних систем. Люди почали інтерпретувати діяльність машини з точки зору власної діяльності та тих ціннісно-мотиваційних детермінант, які стоять в її основі. Це дає відповідь на питання виникнення штучного інтелекту саме як соціально-психологічного феномену. У філософсько-антропологічних пошуках трансгуманістів відкривається відповідь на ще одне питання, яку кінцеву мету переслідує «імітація» людського мислення [4]. Створення «людиноподібної машини» мало на меті відкрити людству технічний простір для еволюції, тому машина наділялася людськими рисами задля подолання психологічного бар'єру подальшого симбіозу. У далекому майбутньому машини матимуть властивості людини: мотивацію, цінності, потреби саме тому, що цей симбіоз уже має відбутися, але поки ці філософсько-антропологічні пошуки є предметом футуристичних передбачень далекого майбутнього. Але люди уже почали сприймати розумні технології з позицій того, що вони все ж можуть «помститися людству», «захопити владу» саме через надмірну екстраполяцію власних рис особистості на об'єкти, які цих рис не можуть мати в жодному разі. І чим більше розумні машини стають схожими на своїх творців, тим більше проявляється антропоморфізм у світі інформаційних технологій. Технофобії як необґрунтовані страхи або відчуття ворожості до технологій та автоматизації мають різні причини виникнення, найбільш поширені пов'язані з об'єктивним розумінням втрати роботи на фоні стрімкого заміщення праці людини на автоматизовану працю машини. З цього приводу виникає багато етичних проблем, які пов'язані з межами використання ШІ, їх впливом на приватне життя людини. У США та країнах Європейського Союзу уже відбуваються антиглобалістичні рухи, які мають на меті захист прав пересічних громадян на працю та приватне життя [8]. Значною проблемою у формуванні об'єктивного розуміння користі штучного інтелекту для покращення життя людей є технологічний розрив між окремими країнами та державними конгломератами. Не всі країни та їх економіки перейшли у рамки інформаційного суспільства, це питання не лише можливості, але й внутрішнього бажання це робити. Тому не всі люди, які проживають на планеті Земля, мають рівномірний доступ до інформації та ІКТ, які цей доступ забезпечують (наприклад, є країни, де доступ до Internet є конституційним правом громадянина). У межах антропологічної парадигми основних підходів до розробки штучного інтелекту зверта-

ється увага на принцип обчислюваності об'єкта аналізу та недоліки створення ШІ шляхом копіювання людських патернів. Вона пропонує посередницький підхід, розроблений на основі культурно-історичної психології, який розглядає процеси мислення як об'єктивні дії, які не завжди описуються у метричній системі знаків мовної системи. Цей підхід веде до нової стратегії розвитку штучного інтелекту як розумної допоміжної системи для людей, яка є джерелом інтелектуальних алгоритмів та їх виконавцем, але залишає людині право ініціації цього процесу [3].

У 60–70 х. роках ХХ ст. починає розвиватися напрям використання ШІ для вирішення локальних, простих завдань. Експертні системи до цього часу використовуються як комплексне вирішення локальних завдань бізнесу, науки, політики, рідше – культури. У них реалізовано досить потужний механізм розв'язання завдань на основі використання математичної логіки різного ступеня складності [2]. Як нами зазначалося раніше, в основі інтелектуального алгоритму штучного інтелекту є опрацювання інформації за допомогою максимальної кількості математично вимірюваних показників. Якщо показник не може бути математично представленим, система намагається його примусово обрахувати, що і призвело до появи парадоксів мислення в обробці інформації ШІ. Друга «холодна зима ШІ» (1960–1980) була детермінована надмірними цілепокладаннями користувачів експертних систем щодо точності та достовірності отриманих даних. Якщо інтелектуальна система позбавлена вад і обмежень людського мислення, їй можна всіляко довіряти, а це, як виявилось, не можна робити. В психології праці давно в науковому вжитку поняття «людський фактор виробництва», проте також варто брати до уваги «технічний фактор», врахування якого на технічно складних та відповідальних процесах унеможлиблює роботи технічних систем в автономному режимі. Статистична ймовірність інтелектуальної помилки штучного інтелекту в рази менша за ймовірність помилки людини, але, як виявилось, люди готові до ймовірності власних помилок та не готові миритися з помилками штучних інтелектуальних систем. Тому друга «холодна зима» ШІ була результатом технологічного обмеження у обробці даних машинами і необхідністю постійної верифікації отриманих результатів з боку людини.

На початку 2000-х відбулося чергове потужне поживлення у застосуванні штучного інтелекту у широкому спектрі суспільної практики. Це сталося через розширення інструментарію (насамперед технічного) «машинного навчання». В наданому документі спостерігаються принципи моделювання штучного інтелекту, пов'язані з лінгвістичним аналізом та обробкою мови. Були розроблені правила для обробки мови, такі як визначення форм слів, синтаксичних структур речень, а також використання логічних умов для обробки мовних конструкцій. Вони були використані для створення систем штучного інтелекту, які здатні аналізувати та розуміти мову, виконувати синтаксичний та семантичний аналіз текстів, а також взаємодіяти з мовними конструкціями. Поява ШІ, що мав

можливість розуміти людську мову, призвела до вибуху технологій машинного навчання і однозначно наблизила його до пересічного користувача. У цей період штучний інтелект виходить на новий рівень – **імітації емоцій людини**. Розробники зробили роботу над помилками базового тесту Алана Тюрінга та визначили, що для повноцінної імітації інтелектуальних процесів потрібен емоційний відгук між діями реципієнта та комунікатора (людини та ШІ). Розробники зараз займаються машинним навчанням систем розпізнавання емоцій людини, що уможливить здатність імітувати та виражати емоції аналогічно тому, як це роблять люди [6]. Це може включати у себе розпізнавання та виведення емоцій у текстовій або голосовій комунікації, а також взаємодію на більш емоційному рівні. Одним з напрямів емоційної імітації є розробка систем, які здатні розпізнавати не лише слова, але й тон, інтонацію та вираз обличчя, щоб визначити емоційний стан особи. Це може бути важливим у розробці віртуальних асистентів та інших інтерфейсів, де важлива емоційна взаємодія, наприклад у роботі психолога [7].

Проблемою взаємодії людини з технічними системами, які вона наділяє додатковими інтелектуальними функціями, є постійна необхідність верифікації даних через можливу помилку логіки алгоритму, який обробляє конкретний запит. Природа помилки у людини та машини різна, тому не варто екстраполювати детермінанти спотворення змісту людини на інструменти генеративного інтелекту. Людина може помилятися через недостатню обізнаність, суб'єктивність, індивідуальні особливості мовної системи, упередженість та з корисливих намірів. Штучний інтелект не має жодного з цих атрибутів, проте він може вводити в оману через брак даних та їх нерепрезентативність. Найпоширенішою помилкою штучного інтелекту є помилки генерації, які виникають у результаті хибного запиту, або запиту, який не відповідає меті людини. Цікавою аналогією є те, що в школі Сократа за дві тисячі років до появи штучного інтелекту вчили правильно ставити запитання, адже завжди отримаєш не те, що хочеш, а те, що просиш. Однією з актуальних професій майбутнього буде оператор систем штучного інтелекту та машинного навчання: фахівці, які будуть тренувати ШІ та ставити релевантні запити відповідно до вихідних завдань.

Повертаючись до проблеми флуктаційного інтересу суспільства до інтелектуальних інформаційних систем, то він пояснюється нами через розвиток у суспільства завищених очікувань щодо його прикладної цінності та їх руйнування у процесі практичного застосування. Жодне технічне рішення ШІ не дало можливості людям відчутти полегшення від автоматизації життєвої рутини повною мірою. Жодна технічна система не може обійтися без людини як джерела запитів та джерела знань. Виявилось, що генеративний штучний інтелект виявився імітацією і у сфері створення контенту, він може комбінувати, робити підстановки та реконструювати, але не здатен до процесу створення як такого. Цього висновку дійшли фахівці Microsoft, Дельфтського технічного університету, Королівської галереї Mauritshuis і будинку-музею Рембрандта в

Амстердамі під час роботи над проєктом «The next Rembrandt» (<https://news.microsoft.com/ru-ru/sleduyushhij-rembrandt/>). У результаті машинного аналізу даних, отриманих у разі тривимірного сканування 346 картин Рембрандта, була створена картина в техніці живописця. Алгоритми ШІ дали можливість відтворити у деталях стиль, манеру живопису людини, яка жила 300 років тому, але не стала автором оригінальної ідеї. Штучний інтелект не зміг би намалювати картину у стилі Рембрандта, якби того не існувало і він не залишив репрезентативне джерело даних у вигляді предметів живопису. Тому *останньою імітацією штучного інтелекту є імітація творчості*. Останніми роками відбулося припорошення різноманітного контенту за рахунок його нескінченної генерації системами штучного інтелекту на відповідні запити користувачів. Amazon, Google уже потерпають від великої кількості інформаційної продукції авторства ШІ, це знецінює працю людини та демотивує до створення нового та оригінального продукту. У 2023 році німецький фотохудожник Борис Ельдагсен виграв у конкурсі Sony World Photography Award у номінації «Креатив» зі світлиною, створеною за допомогою штучного інтелекту. Таким чином він хотів переконатися, чи здатні люди відрізнити штучне мистецтво від оригінального. Виявилось, що не в змозі, що певним чином вплине на індустрію мистецтва та змусить митців шукати нові оригінальні форми відображення дійсності, які поки не опрацьовані системами машинного навчання ШІ.

Висновки та перспективи подальшого дослідження. Розуміння витоків технологій штучного інтелекту дає можливість зрозуміти їх першочергове призначення – позбавити людину від надмірної рутини. Але людське сприйняття цього технологічного процесу обросло зайвими світоглядними уявленнями, які перетворили математичні алгоритми на сутності, якими ті не були і не могли бути. Зараз відбувається певне стримування суспільних фантазій, які переросли в окремих випадках у фобії, щодо надмірного олюднення ШІ та екстраполяції йому людських рис. У зримому майбутньому ШІ не зможе перевершити людське мислення у частині здатності до генерації довільного оригінального продукту та створенні інтелектуальних умовисновків на основі нерепрезентативного та суперечливого набору даних, що з легкістю може робити людина. Відповідно, всі футуристичні передбачення про Світ, де панує штучний інтелект, є результатом інтелектуальної роботи людини, а не планом для реалізації наділених ознаками суб'єктності машин. Трендом 2023 року стало поступове зменшення інтересу бізнесу, науки, держави до технологій штучного інтелекту, проте це не свідчить про загальний тренд на їх постійному використанні у більшості технологічних процесів, які піддаються законам математичної логіки. Людство повинно бути готовим до зміни кон'юнктури ринку праці, зменшення частки робочої сили у секторах, які уже зараз проходять інтенсивне освоєння алгоритмами машинного навчання: сфера математичних розрахунків, лінгвістичний переклад, програмування. Це не означає, що людство стоїть на порозі масового безробіття, але це дасть можливість

застосовувати людську працю у тих випадках, де її доцільність буде зумовлена суто людською ексклюзивністю.

Список використаних джерел і літератури

1. Сапольські Р. Біологія поведінки. Причини доброго і поганого в нас. Київ : Наш Формат. 2021. 672 с.
2. Спірін О.М. Початки штучного інтелекту. Житомир : Вид-во ЖДУ ім. І. Франка, 2004. 172 с.
3. Aleksander I. From Turing to Conscious Machines. *Philosophies*. 2022, No. 7, 57 p.
4. Honavar S. Artificial intelligence in ophthalmology – Machines think! *Indian J Ophthalmol*. 2022 Apr; 70(4): 1075–1079.
5. Teuscher C. Alan Turing: Life and Legacy of a Great Thinker. Springer, First Edition. 2003. 570 p.
6. Ventura R., Custódio L., Pinto-Ferreira C. Artificial Emotions – Good Bye Mr. Spock! *Psychology*. 2005. URL: https://welcome.isr.tecnico.ulisboa.pt/wp-content/uploads/2015/05/672_ventura-spock.pdf.
7. Ventura R. Emotion-based Mechanisms for Decision Making in Autonomous Agents: Dissertação para obtenção do Grau de Doutor em Engenharia Electrotécnica e de Computadores, 2006. URL: https://welcome.isr.tecnico.ulisboa.pt/wp-content/uploads/2015/05/1803_FINAL.pdf.
8. Wall D. Babylon and Beyond: The Economics of Anti-Capitalist, Anti-Globalist and Radical Green Movements. Pluto Press. 2005. 220 p.

References

1. Sapolska, R. (2021). *Bioložiia povedinky. Prychyny dobroho i pohanoho v nas [Biology of behaviour. The reasons for good and bad in us]*. Kyiv: Nash Format. 672 s. [in Ukrainian].
2. Spirin, O.M. (2004). *Pochatky shtuchnoho intelektu [The beginnings of artificial intelligence]*. Zhytomyr: Vyd-vo ZhDU im. I. Franka, 2004. 172 s. [in Ukrainian].
3. Aleksander, I. (2022). From Turing to Conscious Machines. *Philosophies*, 7, 57 p. [in English].
4. Honavar, S. (2022). Artificial intelligence in ophthalmology – Machines think! *Indian J Ophthalmol*. 70(4): 1075–1079 [in English].
5. Teuscher, C. (2003). Alan Turing: Life and Legacy of a Great Thinker. Springer, First Edition. 570 p. [in English].
6. Ventura, R., Custódio, L., & Pinto-Ferreira, C. (2005). Artificial Emotions – Good Bye Mr. Spock! *Psychology*. Retrieved from: https://welcome.isr.tecnico.ulisboa.pt/wp-content/uploads/2015/05/672_ventura-spock.pdf [in English].
7. Ventura, R. (2006). Emotion-based Mechanisms for Decision Making in Autonomous Agents: Dissertação para obtenção do Grau de Doutor em Engenharia Electrotécnica e de Computadores, Retrieved from: https://welcome.isr.tecnico.ulisboa.pt/wp-content/uploads/2015/05/1803_FINAL.pdf [in Spanish].
8. Wall, D. (2005). Babylon and Beyond: The Economics of Anti-Capitalist, Anti-Globalist and Radical Green Movements. Pluto Press. 220 p. [in English].

Kyrychenko V. V.

Doctor of Psychological Sciences, Associate Professor, Professor at the Department of Social and Practical Psychology

Zhytomyr Ivan Franko State University

THE SOCIO-PSYCHOLOGICAL PARADIGM OF UNDERSTANDING THE EVOLUTION OF ARTIFICIAL INTELLIGENCE

The article provides a historical and methodological overview of the development of artificial intelligence technologies and their impact on the socio-cultural activities of humans. It describes the key stages of AI development, the prerequisites for its emergence, outlines the initial positions of developers, and the fundamental models of human interaction with intelligent information systems. Within the socio-psychological paradigm of the evolution of artificial intelligence technologies, three stages of simulating human mental activity are identified: imitation of intellectual processes, imitation of emotions, and imitation of creative activities. The socio-psychological paradigm of the evolution of artificial intelligence is based on the idea that the development of intelligent systems has socio-psychological consequences resulting from the influence of AI on society and interactions among people. This paradigm considers not only the technical aspects of AI development but also the impact of these technologies on interpersonal relationships, ethics, culture, and other social phenomena. It considers the psychological aspects of how intelligent systems are perceived by humans and their interaction in various spheres of life. The socio-psychological paradigm of the evolution of artificial intelligence recognizes that the development of AI should be regulated not only from a technical standpoint but also considering its socio-cultural and psychological consequences for human culture. One of the socio-psychological consequences of human interaction with artificial intelligence systems is the formation of specific fears regarding machines, attributing unjustified properties to them, and purely human traits. However, such trends in interaction with intelligent information systems have not changed the overall cultural trend towards integrating artificial intelligence into a wide range of social processes. Signs of productive interaction with intelligent information systems include understanding their benefits in close collaboration with humans, serving as tools to facilitate routine work.

Key words: artificial intelligence, thinking, creativity, neural networks, information society.