

УДК 342.4:351.86

DOI <https://doi.org/10.30970/PPS.2024.56.17>

ЕТИЧНІ АСПЕКТИ ШТУЧНОГО ІНТЕЛЕКТУ

Анжеліка Шевель

*Сумський національний аграрний університет,
юридичний факультет, кафедра правосуддя та філософії
вул. Герасима Кондратьєва, 160, 40000, м. Суми, Україна*

Стаття присвячена дослідженню етичних аспектів штучного інтелекту. Зазначено, що використання штучного інтелекту сьогодні революціонує наукові дослідження. Розглянуто переваги та негативні наслідки використання штучного інтелекту. Описані ключові питання, пов'язані з етикою використання штучного інтелекту в наукових дослідженнях.

Автор наголошує на тому, що штучний інтелект – реальність сьогодення, зазвичай, неусвідомлена. З упевненістю можна говорити, що кожен, хто користувався комп'ютером із сучасним програмним забезпеченням, працював із інтелектуальними програмами задумувався над етичними аспектами штучного інтелекту.

Нині у науці досі немає єдиного розуміння терміну штучний інтелект. Найчастіше він використовується зазвичай у трьох значеннях: 1) науковий напрямок, що має на меті моделювання процесів пізнання та мислення, використання застосовуваних людиною методів вирішення завдань для підвищення продуктивності обчислювальної техніки; 2) різні пристрої, механізми, програми, які за тими чи іншими критеріями можуть бути названі «інтелектуальними»; 3) сукупність уявлень про пізнання, розум і людину, що уможливають саму постановку питання про моделювання інтелекту. Таким чином під штучним інтелектом може розумітися як науковий напрямок, так і різні пристрої, здатні до рефлексії, що моделюють (імітують) подобу людського інтелекту.

У сучасній науці коло питань, що об'єднуються терміном «штучний інтелект», досить широке. Нині штучний інтелект поєднує велику кількість розділів. До штучного інтелекту відносять як загальну теорію відчуття чи сприйняття, і спеціальні методи типу гри у шахи, докази математичних теорем, написання віршів чи діагностики захворювань. Вчені, спеціальністю яких не є штучний інтелект, знаходять у ньому основу для систематизації та вирішення інтелектуальних завдань, яким вони присвятили значну частину свого життя. І навпаки, фахівці застосовують методи штучного інтелекту в різних областях, де без них не вдається досягти успіху. Саме тому штучний інтелект є воістину універсальною галуззю знань, яка потребує аналізу його використання з етичної точки зору.

У висновку автор зазначає, що розвиток штучного інтелекту може стати мостом між науковими досягненнями та етичними цінностями, що сприятиме просуванню знань та добробуту суспільства.

Ключові слова: відповідальність, етика, етика штучного інтелекту, наукові дослідження, прозорість та обліковість, проблеми приватності та безпеки даних, питання відповідальності та прийняття рішень, питання рівності та справедливості, моральні проблеми в інтелектуальних системах, рефлексія, штучний інтелект.

Актуальність дослідження. Штучний інтелект сьогодні викликає найбільшу кількість багатогранних, підчас дуже суперечливих емоцій, починаючи від активного просування технологій, пов'язаних з його подальшим розвитком, та з його використанням.

У цьому зв'язку дослідження етичних аспектів застосування штучного інтелекту набувають особливого значення, що зумовлює актуальність обраної теми та доцільність проведення досліджень для розвитку цього питання.

Метою дослідження є аналіз етичних аспектів штучного інтелекту, а саме переваг та негативних наслідків використання штучного інтелекту.

В своїх наукових дослідженнях українські вчені, серед яких – Іванюк Д., Єфремов М., Твердохліб І.А., Улянівський Т., Панченко В., Резнікова Н. та багато інших розглядають проблему штучного інтелекту.

М. Рогоза зазначає: «Фахівці в галузі штучного інтелекту стверджують, що на поточному етапі розвитку штучний інтелект є ще досить слабким. Сьогоднішній штучний інтелект має недосяжні можливості, які здаються надто фантастичними й нереалістичними. Етичні проблеми, пов'язані з його використанням, вважаються проблемами навіть не другорядними або третьорядними. Проте, подібно до біотехнологій, етичні обговорення, що точаться навколо штучного інтелекту, нагадують ситуацію з ядерними дослідженнями. Тоді, коли ядерні розробки вважалися неможливими, етичні питання не виникали. Але коли ядерна технологія стала реальністю, було вже пізно піднімати питання етики» [5, с. 49].

Під штучним інтелектом розуміється розпізнавання образів, розпізнавання мови, тощо. У цьому ключі свого часу Ф. Джордж [2] також показав, що створення штучного інтелекту, подібне до людського мислення неможливе повною мірою (на той момент розвитку штучного інтелекту). Він обґрунтував такий висновок відсутністю і неможливістю створення певної частини системи, яка б дозволяла відновлювати систему у разі її знищення і була б сама незнищенна (за аналогією, можна сказати – мала б безсмертну «Душу» чи «ідею» Платона). Однак, на сьогодні розвиток глобальних мереж, вірусних та багато-агентних технологій дозволяє практично реалізувати програми, що теоретично існують, поки існує глобальна мережа. Отже, обмеження Ф. Джорджа можуть бути зняті.

При всій різноманітності трактувань поняття штучного інтелекту загальним більшість із них є визнання те, що системи штучного інтелекту спрямовані на моделювання чи імітацію людського мислення, зокрема здатність до рефлексії та асоціації.

Вважають, що про реалізацію штучного інтелекту можна буде говорити лише тоді, коли автомат почне вирішувати завдання, непосильні для людини, причому зробить це не внаслідок високої швидкодії, а в результаті застосування нового знайденого методу. Однак не всі з цим погоджуються. Більшість випадків досліджень з штучного інтелекту лише порівняні з результатами, отриманими людиною, і не такі оригінальні.

Відомий британський вчений і популяризатор науки, професор Стівен Хокінг, відповідаючи на питання про нову технологію, якою він користується для спілкування із зовнішнім світом, застеріг, що поява повноцінного штучного інтелекту може стати кінцем людської раси. Він стверджував, що швидкість перебудови штучного інтелекту набагато швидше аналогічної швидкості людей, обмежених повільною біологічною еволюцією, і з часом людям буде все складніше конкурувати з штучним інтелектом.

Професор права Вашингтонського університету, який вивчає етичні аспекти робототехніки і комп'ютерних систем, Райан Кало вважає, що немає причин турбуватися про загрозу, яку потенційно можуть створити «розумні машини». В основі його висновків – наявні дослідження, в яких ніде не говориться про можливість створення штучного розуму, навіть на рівні примітивних ссавців, не кажучи вже про людський розум. Тому немає причин вважати, що він раптово турбуватиметься захопленням світу, якщо тільки це не буде від самого початку в нього запрограмовано. Дискусію про безпеку штучного інтелекту, професор пропонує перевести на вивчення того, як предмет суперечок може служити на благо суспільства [5, с. 22].

Фахівці [4] стверджують, що вже на 2045 рік відбудеться неминуче витіснення людини з багатьох сфер. Відповідно до ще більш радикальних уявлень, комп'ютери розвиваються настільки швидко, що перевершать за інтелектом людину через одне, або, найбільше, два покоління. Базою для подібних тверджень є те, що важко знайти сьогодні

більш широку галузь знань ніж штучний інтелект – як за обсягом теорій, методів, технологій та практик, так і за обсягом фахівців, які долучені до його вивчення та розвитку.

Етика штучного інтелекту може бути розглянута як набір принципів, які визначають, яким чином штучний інтелект повинен діяти та яким не повинен. Ці принципи зазвичай орієнтовані на безпеку, справедливість і гуманістичні цінності, на захист прав користувачів і запобігання можливим зловживанням у сфері інформаційних технологій [1].

Нижче розглянемо деякі з основних етичних проблем, пов'язаних зі штучним інтелектом, і можливі шляхи їх вирішення.

Прозорість та обліковість. Однією з найважливіших аспектів штучного інтелекту є прозорість його рішень та обліковість процесу, за яким вони приймаються. У багатьох випадках, особливо в глибокому навчанні та нейронних мережах, алгоритми працюють як чорні скриньки, де важко розуміти, як саме вони прийшли до певного висновку. Це ставить під сумнів довіру до їхніх рішень, особливо коли вони впливають на життя людей, наприклад, у сферах юстиції та медицини.

Забезпечення прозорості та обліковості алгоритмів є ключовим завданням для етичного застосування штучного інтелекту.

Проблеми приватності та безпеки даних. Ще однією з важливих етичних проблем є приватність та безпека даних. Збір, зберігання та обробка великих обсягів особистої інформації може вести до порушень приватності та може бути використана для маніпуляції або шкідливих цілей. Наприклад, алгоритми машинного навчання, які аналізують особисті дані користувачів для рекламних цілей, можуть створювати фільтрувальні, які обмежують доступ користувачів до різноманітності інформації та поглиблюють їхні погляди. Забезпечення конфіденційності та захисту особистої інформації вимагає розробки суворих стандартів і правових рамок.

Питання відповідальності та прийняття рішень. Наступною проблемою, на яку слід звернути увагу є питання відповідальності за рішення, прийняті системами штучного інтелекту. У випадках, коли алгоритми виконують дії, які мають серйозні наслідки для людей або суспільства в цілому, важливо визначити, хто несе відповідальність за ці дії. Наприклад, у випадках безпілотних автомобілів, які приймають рішення на дорозі, виникає питання, чи несе відповідальність виробник автомобіля, розробник програмного забезпечення або користувач, який використовує систему.

Питання рівності та справедливості. Штучний інтелект може підсилити нерівність у суспільстві, якщо він не враховує різноманітність та особливості різних груп. Наприклад, алгоритми розпізнавання обличчя можуть бути менш ефективними для розпізнавання обличчя людей з певним кольором шкіри або етнічним походженням. Такі системи можуть призводити до дискримінації та порушень прав людини. Щоб уникнути цього, необхідно враховувати принципи рівності та справедливості при розробці та застосуванні алгоритмів штучного інтелекту.

Моральні проблеми в інтелектуальних системах. Існує також широкий спектр моральних проблем, пов'язаних зі штучним інтелектом. Наприклад, як системи штучного інтелекту повинні вирішувати моральні дилеми в екстремальних ситуаціях? Чи повинні вони дотримуватися загально прийнятих моральних норм? Ці питання важливі для розуміння того, як штучний інтелект може впливати на моральні цінності суспільства та індивідуальні переконання [7].

Отже, постає питання: штучний інтелект – допоможе людству чи знищить його? Про це сперечаються найвидатніші вчені та розробники. Винахідник Ілон Маск, фізик Стівен Гокінг, голова корпорації «Майкрософт» Білл Гейтс застерігають від розробки

штучного інтелекту, але інші вчені і відомі особистості, такі як Марк Цукерберг, програміст і засновник мережі Фейсбук, навпаки бачать користь для всього людства у застосуванні штучного інтелекту.

В історії відомі спроби зовнішнього моделювання людських можливостей. В цьому аспекті згадується середньовічна легенда. Кельн середини XIII століття. Захоплення алхімією та астрологією. В одному з центрів з вивчення теології працюють чернець Хома Аквінський та його вчитель Альберт фон Больштедт, прозваний ще за життя Великим за свою вченість. Альберт часто працював у майстерні, створюючи хитромудрі пристрої та механізми. Одного разу, прийшовши до нього в гості, Хома побачив жінку, тіло якої світилося якимось неземним світлом. Вона підвела руку, повернула до нього голову і ввічливо привіталася. Але було в ній щось страшне та неживе. Вихований у благочесті і шануванні бога, Хома одразу ж зрозумів, що це диявольське наслання. Миттю розбив він тростиною найкращий інженерний витвір того часу – людиноподібну ляльку, яка вміла ходити, рухати руками і говорити. Подібні казки, перекази, міфи ніколи не з'являються на порожньому місці. В основі їх лежать якісь реалії, і лише фантастичне заломлення цих реалій, можливо, просто не зрозумілих сучасниками, надає легендам відтінку незвичайності та казковості. Ідеї гомункулусів – істот, подібних до людини, яких нібито можна було отримати штучно в XVI столітті, андроїдів – механічних людей XVIII століття (відомі твори французького механіка Жака де Вокансона та швейцарських годинників П'єра-Жака Дро та його сина.

Зусилля, витрачені на створення андроїдів і роботів, починаючи з другої половини XX століття, не марні. Було позитивно вирішено питання можливості механічної імітації антропоморфних рухів. Наслідком став розвиток науки біомеханіки, що дозволило нашого часу створювати протези для людей.

В наш час, відомий робот-гуманоїд Софія, створена на основі штучного інтелекту, навчається, жартує і дає інтерв'ю, в яких виражає бажання мати родину, як в людей.

Паралельно спробам зовнішнього моделювання людських можливостей постійно йшли пошук та розвиток методів, дозволяють імітувати внутрішні процеси, які є у людини. Насамперед процес мислення. Цим займалися: Сократ (сократичні бесіди), Арістотель (теорія силогізмів), Ф. Бекон (індукція), Д. Буль (математична логіка), Г.-Г. Гадамер (герменевтика) і т.д. Завдяки їм до середини XX століття вже існувала будівля теорії логічних міркувань, що сформувалася. Чимало уваги приділялося і моделюванню комунікаційних можливостей людини.

Недоліки природних людських мов, пов'язані з їхньою багатозначністю та неточністю, звертали на себе увагу ще в Стародавній Греції. У XVII столітті ідея створення точної «філософської» мови для опису знань про навколишній світ знайшла численних прихильників. Спроби Декарта і Лейбніца не увінчалися успіхом, але зусилля, витрачені створення філософської мови, не зникли даремно. Наприклад, класифікація Карла Ліннея – перша суворо наукова класифікація в природознавстві – виникла на основі ідей філософської мови. Сучасні дослідження в галузі інтелектуальних систем також активно використовують ті ідеї про строгу мову знань, які були сформульовані ще триста п'ятдесят років тому. Але найважливіше – було створено логічну алгебру.

Такі були деякі ідеї, пов'язані з імітацією зовнішніх форм поведінки людини, її здібностей до міркувань та комунікацій. На початку XXI століття багато з цих ідей принесло багато плодів, як, наприклад, ідея формалізації міркувань. Інші залишилися лише на рівні початкових припущень. Проте всі вони виявилися потрібними та корисними, коли з'явилися технічні пристрої, названі електронними обчислювальними машинами, здатні

імітувати багато специфічних особливостей мислення та життєдіяльності людини. Все це безпосередньо передувало появі науки «Штучний інтелект».

Так, у липні 2017 року керівництво соціальної мережі «Facebook» відключило свою систему штучного інтелекту після того, як машини стали спілкуватися власною, неіснуючою мовою, яку люди не розуміли: система використовувала чат-боти, які спочатку створювалися для спілкування з живими людьми, але поступово почали спілкуватися між собою. Спочатку вони спілкувалися англійською мовою, але пізніше почали листуватися мовою, яку вони самі створили в процесі розвитку програми.

Шахи вже давно вважаються критерієм розвитку штучного інтелекту, а комп'ютерні технології, що змагаються в цій грі, з'явилися ще в 1950-х роках. Історія шахових програм на основі штучного інтелекту – захоплива, сповнена важливих проривів, які розширили межі можливого в галузі штучного інтелекту.

Одним з перших прикладів шахової програми на основі штучного інтелекту була розроблена IBM «Deep Blue», яка перемогла чемпіона світу з шахів Гаррі Каспарова в 1997 році. Це стало важливим поворотом в історії штучного інтелекту, оскільки вперше комп'ютерна програма перемогла чемпіона світу з шахів серед людей.

Відтоді такі шахові технології продовжують вдосконалюватися і розвиватися. Сьогодні існує низка програм, які можуть грати на рівні гросмейстера, а деякі з них навіть перевершують здібності гравців-людей.

На початку січня 2023 року популярний шаховий ресурс Chess.com запустив кілька ігрових ботів-котів. Найпопулярнішою стала кішка Mittens – журналісти The Wall Street Journal назвали її «новою лиходійкою шахового світу».

Попри милий зовнішній вигляд, Mittens виявилася несподівано сильною суперницею. Цікаво, що у чаті під час гри бот-кішка цитує Ніцше та Оппенгеймера, дискутує про Ван Гога та французьке кіно 60-х років.

Загалом бот використовує виснажливу тактику. Американський гросмейстер Хікару Накамура вирішив зіграти з Mittens. Результат партії – нічия. Стрім гри шаховий чемпіон розмістив на своєму Youtube під заголовком «*Mittens: цей бот змусить вас кинути шахи*». А Леві Розман після матового програшу від цифрового kota затвітив: «Цей бот – псих».

Кішка-гросмейстер – розробка студента Гамільтонського коледжу Вілла Вейлена. Саме йому належить ідея створення «сильного гравця з оманливо милою зовнішністю». Команда Chess.com підхопила ідею Вейлена і перетворила кішку на «злого генія», який відверто мучить шахістів. Щоправда, у The Wall Street Journal вбачають у цій історії позитивний момент: Mittens може захопити грою нових прихильників шахів. Наприклад, відео Накамури на Youtube вже переглянули понад 3,6 млн разів.

Отже, те що раніше здавалося вершиною людської творчості – гра в шахи, шашки, розпізнавання зорових і звукових образів, синтез нових технічних рішень, на практиці виявилось не такою вже складною справою (тепер робота зводиться не до дослідження рівня можливості або неможливості реалізації перерахованого, а до знаходження найбільш оптимального алгоритму). Є надія, що й повне моделювання мислення людини виявиться не такою вже й складною справою.

Як пише Погореленко А.К., люди можуть втратити почуття власної унікальності внаслідок «вибуху інтелектуальності», який буде спричинений штучним інтелектом, та який буде здатний перевершити своїми можливостями самих людей.

Будь-який розвиток завжди передбачає переваги та ризики. Головне – баланс. Штучний інтелект – це вже даність. Завдання сьогодні – не відійти, а оптимізувати потенційні

ризиків від нього. Цей процес, безумовно, повинен бути чітко регламентованим та мати якісні цілі. Чим, власне, зараз й займаються провідні фахівці в цій галузі, щоб наслідки функціонування штучного інтелекту були економічно та соціально значущі, а не руйнівні [4, с. 23].

Висновки. Штучний інтелект є зараз «гарячою точкою» наукових досліджень. У цій точці, як у фокусі, сконцентровано найбільші зусилля програмістів, лінгвістів, психологів, математиків, інженерів. До списку можна включити представників практично всіх наукових спеціальностей. Саме тут вирішується багато корінних питань, пов'язаних з шляхами розвитку наукової думки, з впливом досягнень у галузі обчислювальної техніки на життя майбутніх поколінь людей та його етичного застосування.

Штучний інтелект – це те, що кардинально змінило світ. Напряма майбутніх змін – наслідок вибору та світогляду його творців.

Додатковий «мінус» штучного інтелекту – відсутність допоміжного блоку автономної (усвідомленої) відповідальності ухвалення остаточного рішення, що свідчить про неможливість ухвалення адекватних рішень. Але, якими б депресивними очікуваннями не наповнювали деякі фахівці інформаційне поле стосовно перспектив розвитку штучного інтелекту, поки очевидно одне: негативні наслідки від розвитку штучного інтелекту повністю нівелюються позитивними результатами та перспективами від його використання.

Список використаної літератури

1. Балик А. (2023). Етика в IT-сфері. <https://lemon.school/blog/etyka-v-it-sferi>.
2. George F. The Foundations of Cybernetics. London: Gordon and Breach Science Publisher, 1977. 286 p.
3. Moshe Y. Vardi, «The Consequences of Machine Intelligence». Atlantic, October 25, 2012. URL: <http://www.theatlantic.com/technology/archive/2012/10/the-consequences-of-machine-intelligence/264066>.
4. Погореленко А.К. Штучний інтелект: сутність, аналіз застосування, перспективи розвитку. Науковий вісник Херсонського державного університету. Серія Економічні науки. Випуск 32. 2018 С. 22-27
5. Рогоза М. М. Моральні дилеми у галузі штучного інтелекту. Категорії. Мораль: зб. наук. матеріалів за результатами Всеукраїнського круглого столу «Читання пам'яті Івана Бойченка-2019. Людина. Історія. Мораль» (25 жовтня 2019 р.) / редкол.: А. Є. Конверський [та ін.]; Київський національний університет ім. Т. Шевченка, Інститут філософії ім. Г.С. Сковороди НАН України, Центр гуманітарної освіти НАН України. Київ: Знання України, 2019. С. 45–50. URL: <https://ekmair.ukma.edu.ua/handle/123456789/18081>
6. Федоров А. Перегони технологій та нова етика. <https://chas.news/future/peregoni-tehnologii-ta-nova-etika>.
7. Штучний інтелект: що це і яку несе небезпеку. URL: https://24tv.ua/techno/shtuchniy-intelekt_shho_tse_i_yaku_nese_nebezpeku_n91466

ETHICAL ASPECTS OF ARTIFICIAL INTELLIGENCE

Anzhelika Shevel

*Sumy National Agrarian University,
Faculty of Law, Department of Justice and Philosophy,
Gerasim Kondratieva str., 160, 40000, Sumy, Ukraine*

The article is devoted to the study of ethical aspects of artificial intelligence. It is noted that the use of artificial intelligence is revolutionizing scientific research today. Advantages and negative consequences

of using artificial intelligence are considered. The key questions are described, related to the ethics of using artificial intelligence in scientific research. The author emphasizes that artificial intelligence is a reality today, usually unconscious. It is safe to say that anyone who has used a computer with modern software, worked with intelligent programs, thought about the ethical aspects of artificial intelligence. Currently, science still does not have a single understanding of the term artificial intelligence. Most often, it is usually used in three meanings: 1) a scientific direction aimed at modeling the processes of cognition and thinking, the use of human-applied problem-solving methods to increase the productivity of computing equipment; 2) various devices, mechanisms, programs that can be called “intelligent” according to certain criteria; 3) a set of ideas about cognition, mind and man, which make it possible to pose the very question of intelligence modeling. Thus, artificial intelligence can be understood as both a scientific direction and various devices capable of reflection that model (imitate) the likeness of human intelligence.

In modern science, the range of issues united by the term “artificial intelligence” is quite wide. Nowadays, artificial intelligence combines a large number of sections. Artificial intelligence includes both the general theory of sensation or perception, and special methods such as playing chess, proving mathematical theorems, writing poetry, or diagnosing diseases. Scientists whose specialty is not artificial intelligence find in it a basis for systematizing and solving intellectual tasks to which they have devoted a significant part of their lives. Conversely, experts use artificial intelligence methods in various areas where it is impossible to achieve success without them. That is why artificial intelligence is a truly universal field of knowledge that requires an analysis of its use from an ethical point of view. In the conclusion, the author notes that the development of artificial intelligence can become a bridge between scientific achievements and ethical values, which will contribute to the advancement of knowledge and the well-being of society.

Key words: responsibility, ethics, ethics of artificial intelligence, scientific research, transparency and accountability, privacy and data security issues, issues of responsibility and decision-making, issues of equality and justice, moral problems in intelligent systems, reflection, artificial intelligence.