



Олена ЯЦЕНКО

кандидат філософських наук, доцент, доцент кафедри менеджменту та інноваційних технологій соціокультурної діяльності НПУ імені М. П. Драгоманова, Visiting researcher in Institute for Human-Centered Engineering, School of Engineering and Computer Science, Bern University of Applied Sciences, Biel, Switzerland
e-mail: yatsenkood@gmail.com
ORCID: <http://orcid.org/0000-0003-0584-933X>

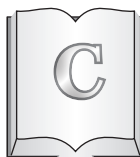
Ключові слова: взаємодія людина-машина, штучний інтелект, доповнений інтелект, робототехніка, додаткова співпраця.

Технології робототехніки та штучного інтелекту широко впроваджуються в процеси автоматизації у різних сферах суспільного життя і виробництва. Таким чином, у діяльність машин делегується все більш високий ступінь свободи в ухваленні рішень, що спричинює дискусії про потенційну сингулярність як неминучу перспективу розвитку технологій. В статті проаналізовано актуальні виклики та проблеми розвитку технологій штучного інтелекту та робототехніки, наголошується на необхідності розробки алгоритмів та інструментів взаємодії людини та машини, що спрямовані на доповнення, а не виключення, розширення, а не нівеляцію.

УДК 378.147
DOI 10.31392/NPU-VOU.2022.3(86).05

АНАЛІТИКА СУЧАСНИХ ВИКЛИКІВ РОЗВИТКУ ШТУЧНОГО ІНТЕЛЕКТУ ТА РОБОТОТЕХНІКИ У ФОКУСІ ПЕРСПЕКТИВИ СИНГУЛЯРНОСТІ

© Яценко О., 2022



учасність дає нам багатий матеріал для сприйняття та розуміння того факту, наскільки карколомні зміни відбуваються в життєдіяльності людства в останні століття. Розвиток робототехніки, технологій штучного та доповненого інтелекту, є по суті реалізацією античного принципу Deus ex machina. В античному театрі цей принцип означав неочікувану та доречну божественну присутність для розв'язання нездоланих перешкод та ускладнень. В сучасному світі високих технологій цей принцип трансформовано у феномен пасивного користувача, коли узвичаєні речі щоденного вжитку сприймаються як чорна скринька, устрій та функціонал якої не знайомий власнику. Постає доречне запитання: чи людство вже перетнуло червоні лінії, або точки не-повернення в процесі делегування своїх повноважень машинам, роботам та алгоритмам?

Спробуємо проаналізувати актуальні та потенційні ризики розвитку та поширення робототехніки та технологій штучного інтелекту (ШІ). Такий досвід систематизації ґрунтується на розрізненні основних сфер життя людини, в яких ці проблеми помітні та вагомі. Цей перелік не є вичерпним, оскільки вплив робототехніки та технологій ШІ швидко зростає. Тому окреслимо основні вектори таких зрушень.

Найбільш чутливою сферою до впливу розвитку робототехніки та ШІ доречно назвати виклики та проблеми *психологічного* гатунку. Філософська антропологія вчить нас, що людина є «незавершений проект», і свою завершеність сучасна людина актуалізує через високий рівень очікувань функціонування технічних засобів. Отримуємо доволі парадоксальну ситуацію: недосконала людина намагається створити досконалий інструмент. Втілення таких стратегій загрожує потенційною втратою людської ідентичності. Таке упереджене сприйняття робототехніки та ШІ ґрунтується на архаїчній практиці розуміння Іншого крізь призму власних предикацій, або антропоморфізм формалізованою мовою науки. Тому функціонуванню машини вважаються притаманними певні людські емоції: злий умисел, заздрість, мстивість тощо.

Технологічний процес є стандартизованим та алгоритмізованим, і взаємодія з машиною часто не вимагає оригінальності та креативності. На цій основі зростає проблема відчуження, коли людські щоденні практики нагадують механістичні операції машини. Така ситуація сприймається як проблемна, оскільки люди звикли думати про всі предмети та явища крізь призму власної природи, поглядів та цінностей. Тому ми оцінюємо роботів та ШІ як Іншого (партнера або ворога, помічника або перешкоду). Оцінюючи такі перспективи, дослідники пропонують різні сценарії майбутнього, антиутопічні за своїм характером: заміщення людської праці роботами та масове безробіття (Ford, 2015), панування «розумних машин» як провайдерів деґуманізації роботи (Head, 2014) або виник-

нення особливого соціального класу – кіберпролетаріату (Huws, 2014). Інформаційні технології, робототехніка та ШІ підкріплюють занепокоєння про можливе домінування «суспільства машин» (Carr, 2014). Відчуження за допомогою сучасних технологій продукує суспільство дозвілля, емоційного голоду, культивування егоїзму та невизначеного статусу істини.

Наступний рівень аналітики – *соціальний*. Відповідно до теорії дарвінізму, різні види (біологічні, але можливо не тільки) з необхідністю залучені до боротьби за виживання. І цей аспект натурального відбору покладено в основу інтерпретації взаємодії людини та машини. Проте ця взаємодія не обмежується конкуренцією, а також передбачає співпрацю. Дослідники вирізняють різні рівні взаємодії людини та робота в діалектиці автономії та близькості: співіснування (без спільних задач), кооперація (наявність спільної мети) та колаборація (складні задачі з/без фізичним контактом). Часто такі ієрархії взаємодії людини з машиною називають різними ступенями свободи (Hentout, 2019). Специфікація різних видів професійної діяльності разом із автоматизацією та автономізацією у виконанні окремих завдань, призводить до самотності та анонімності в сучасній спільноті. Іншими словами, людина прагне створити найбільш комфортне середовище із найбільш зручним партнером: роботом або ШІ. Напевно, саме тому найбільш поширені у виробництві роботи-гуманоїди (компанійони або помічники) або машини-функції.

Із плином технічного прогресу виникає перспектива формування принаймні двох соціальних класів. Один з них так звані «непотрібні люди» (з низьким рівнем компетентностей), а інший – «нова соціальна еліта» (ІТ-спеціалісти та інженери як сучасні жреці сакрального знання). Питання декваліфікації проти розширення можливостей є визначальним для динаміки сучасного суспільного життя. В минулому впровадження технологічних інновацій також часто призводило до соціальних вибухів та безробіття. З огляду на системний та стрім-

кий характер таких трансформацій у сьогоденні, виникає необхідність переосмислення принципів організації часу та простору праці, цінностей публічного та приватного гатунку, цілей та значення креативності та відповідальності в ухваленні рішень, компетентності та експертності у технологічних та виробничих процесах. Оскільки праця стає все більш мобільною, пластичною та автономною, то на думку окремих дослідників, її доречно визначати також і менш безпечною та передбачуваною (Malecki & Moriset, 2007).

В останні кілька років були ідентифіковані випадки дискримінації з боку оперативних систем-даних. Переважно ці випадки зафіксовані аплікантами на вакантні посади на ринку праці. Так, алгоритми машин мають бути зорієнтовані на пошук релевантної інформації стосовно набутих аплікантом навичок та компетентностей відповідно бажаної посади та функціоналу завдань, які вона передбачає. Натомість автоматизація процесу первинного розгляду та верифікації резюме аплікантів призводила до їх дискримінації за ознаками статі, віку, сексуальної орієнтації, етнічної приналежності, соціально-економічного становища тощо. Спроби знайти технічне рішення цієї проблеми не були вдалими. Головний висновок фахівців полягає у визнанні факту неможливості зафіксувати у вигляді алгоритму моральні норми та директиви. Саме тому, на їхню думку, ухвалення рішень є виключно відповідальністю та місією людини. Тому у вирішенні питання про загальний контур та логіку взаємодії людини та машини доречно закласти принцип розширення людських здібностей та доповнення людської природи та поінформованості, а не автоматизації процесу та заміни людини машиною. М. Berk (Berk M., 2019) називає цю процедуру «оцінкою ризиків», на кшталт ухвалення рішення про умовно-дострокове звільнення злочинця суддями, які мають винести обґрунтований прогноз про його майбутню поведінку та соціальну інтеграцію.

Проблема полягає в тому, що технології не мають можливості аналітики контек-

сту, тільки закладених параметрів. Тільки людині притаманне так зване «стереоскопічне мислення», або можливість побачити ситуацію під різним кутом зору, з різних вихідних позицій локалізації об'єкту дослідження. Для сучасного суспільства нагальною проблемою доречно вважати не відсутність таких спеціалістів, а організацію ефективної комунікації та колаборації між ними. В загальних рисах, ми знову маємо справу з певною парадоксальністю. З одного боку, наукове та технологічне знання все більше розгалужується і потребує фахівців не широкого профілю, а вузькоспеціалізованого напрямку. А з іншого, для ухвалення виваженого та доцільного рішення необхідний широкий контекст становлення та фіксації проблеми, а також плюралізм підходів до її розв'язання. Тому сучасна робототехніка та технології ШІ відмовляються від універсальних протоколів та загальних рішень. Натомість більш ефективними є так звані «no-code platforms», або комунікативні простори, що створюються ситуативно для розв'язання певної проблеми із залученням фахівців з необхідними компетентностями.

Розвиток робототехніки та ШІ спричинює вплив також на економічну сферу соціального життя. Технологічний прогрес вимагає величезних інвестицій з високим рівнем конкуренції на ринку та постійною стратегією стимулювання попиту та споживання. Це збільшує розрив між виробництвом і освітою, особливо в різних регіонах і країнах. Але цифровізація та робототехніка можуть бути корисними інструментами для «здорової» роботи, економічного зростання, мінімізації використаних ресурсів, оптимізації виробництва та встановлення зв'язку між бізнесом і суспільством.

Розвиток робототехніки та ШІ породжує певні загрози у сфері етики та права. Як відомо, людина вибирає між кількома можливостями. Високі технології не просто несуть таку інформацію, але можуть обмежити її коло або навіть спотворити. Крім того, відносини у співпраці людини і ма-

шини зараз не мають правового регулювання. Наприклад, робот Софія має набагато більше прав людини, ніж жінки в Саудівській Аравії. Або проблема порушення інтелектуальної власності та захисту даних. Необхідність приватності та безпеки є основними для людської природи. І зараз ми можемо знайти інформацію про людину без її згоди (про шлюб, дітей, хобі, смаки та інше). Ці дані можуть бути використані для маніпуляцій. Інструментами такого впливу є робототехніка (набагато менше) та ШІ. І це відбувається зараз. Але в будь-якому випадку в таких видах маніпуляцій це все одно конфлікт інтересів людей. Деякі мислителі визнають, що справжній конфлікт між людиною та машиною відбудеться ТОДІ, КОЛИ ШІ матиме можливість вибирати мету, і якщо ця мета буде відрізнятися від нашої.

Технології в цілому, робототехніка та ШІ зокрема, дають людині величезну силу. Відносини влади зосереджені в політиці. Отже, у нас є певні підстави хвилюватися щодо можливостей використання цих технологій як інструменту пропаганди, як механізму маніпулювання результатами виборів, як способу створення технократичної

диктатури або як методу знищення здатності людини приймати рішення.

Розвиток робототехніки та ШІ має також внутрішні (технічні) проблеми. Загальноприйнятими словами можна виділити основні правила, які вимагають рішення в кожному випадку винаходів: перевірка, валідність, безпека, контроль, «дрейф набору даних». Завдяки зростаючій потужності технології обробки даних вони можуть навмисно керувати протоколами та вибором дій, тим самим кидаючи виклик домінуванню людської волі, автономії та, зрештою, влади (Murray et al. 2021). Ці проблеми поділяються на тисячі різних завдань для фахівців, без завзятої праці яких ми не мали б таких корисних та зручних набутоків цивілізації. Певний ракурс цієї проблеми знаходиться у сфері освіти: вивчення користувачів, машинне навчання та їх співвідношення.

Буде неправильним завершити список актуальних проблем розвитку ШІ та робототехніки без екологічного контексту. Масштаби впливу людини на природу непередбачувані. І роль робототехніки та ШІ в цій ситуації також може бути позитивною, як інструмент для моніторингу, коригування та реалізації екологічної діяльності.

№	група	назва	підстави	тлумачення
1	психологічні	Недосконала природа людини	Компенсаторна функція техніки	Моральне, а не прагматичне розуміння техніки
		Втрата ідентичності / "машинність"	Глобалізація та стандартизація	Теорії трансгуманізму
		Антропоморфізм як когнітивний принцип	Необхідність людиноцентричної технології	Еволюційний механізм людської активності
		Відчуження	Локус контролю	Непередбачувані процес та результат
		Емоційне залучення	Чиста раціональність технології	Емоційне вигорання в ході автоматизованої праці
2	соціальні	Показники праці	Автоматизовані процеси	Праця стала більш мобільною, гнучкою та вільною, але також менш безпечною та передбачуваною
		Брак солідарності	Поділ праці та автономність особи	Самотність доби ІКТ
		Соціальна роль робота	Широке коло потенційних функцій	Дві основних моделі: робот-компаньйон та машина-функція
		Дозвілля/ заміщення людини	Стрімкий прогрес технологій	Стрімкі зміни на ринку праці

		Нова соціальна еліта	ІТ-спеціалісти та інженери як нова еліта інформаційного суспільства	Нові конфігурації влади в суспільних відносинах
		Декваліфікація/ розширення можливостей	Автоматизовано велику кількість навичок	Люди розучилися робити багато операцій (обчислення, наприклад)
		Дискримінація	Один протокол для різних завдань	Системи не працюють (або не працюють належним чином) для певних груп.
		Організація часу/ простору	Маркування простору і часу взаємодії людини і машини	Продуктивність праці, правила техніки безпеки, мегаполіси тощо.
		Колективні цінності	Переоцінка цінностей	Інший вид меж між публічним і приватним
		Пасивні користувачі	Відчуження людини у виробничій діяльності	Відсутність знань та інтересу, втрата мотивації та креативності
3	економічні	Величезні витрати	Проблема рентабельності технології	Ризик і перспективи інвестицій
		Старіння техніки	Високий рівень конкуренції в дослідженнях технологій	Закони вільного ринку
		Насиченість ринку	Стратегії масового виробництва	Стимуляція попиту та пропозиції
		Кількісно, ідеальний працівник	Конкуренція на ринку	Необхідно змінити способи вимірювання ефективності праці
4	етичні та юридичні	Права людини та соціальні гарантії	Непомітний вплив на прийняття рішень людиною	Соціальні мережі, системи навігації, пошуківі системи обмежують можливість вибору
		Етика ШІ та робототехніки	Цифрова культура породжує нові формати співпраці	Проблеми суб'єктності, свободи волі та відповідальності, ліберальних цінностей тощо.
		Порушення інтелектуальної власності та захисту даних	Дані як інструмент і як товар	Широкий вибір способів використання інформації
		Різні цілі людини і машини	Розширення самостійного функціонування машини	Потенційно машина може перестати виконувати завдання помічника людині
5	політичні	Інструменти для пропаганди	Вплив на інформаційний простір	Впровадження маніпулятивних методів впливу на масову свідомість
		Маніпуляція результатами виборів	Від електронного урядування до системи підрахунку голосів	Відсутність прозорості в демократичному процесі
		Технократична диктатура	Можливості необмеженого контролю	Як система соціального кредиту в Китаї
		Втрата повноважень приймати рішення	Стереотипність мислення та автоматизм практиків призводить до бажання уникнути відповідальності.	Здатність приймати рішення обмежена відсутністю розуміння того, як працюють технології.
6	технічні	Верифікація	Вимагання випробувань, перевірок, специфікацій та аналізу конструкції	Порядок створення та апробації технології
		Вагомість	Послання вимог замовника з функціональністю програмного забезпечення	Орієнтація на соціальний запит
		Безпека	Стандарти безпеки та кіберзахист	Проблема вартості та ціни
		Контроль	Рівень втручання у функціональність машини	Розвиток людиноорієнтованих технологій

		“Дрейф набору даних”	Потенційна здатність машини працювати з неточними даними	Ефективний алгоритм VS помилки в базових даних
		Дослідження користувачів / Машинне навчання	Кореляційний розвиток умінь і функцій	Мета гармонізації освіти людини та розвитку машин
7	екологічні	Антропоцен	Руйнівний вплив на природу розвитку технологій	Загроза фізичному виживанню людства
		Високе енергоспоживання	ШІ споживає багато енергії	Необхідність використання великих ресурсів

Розширений інтелект як спосіб взаємодоповнюючої співпраці

Розвиток робототехніки та технологій ШІ актуалізує наявні та створює нові проблеми та виклики стабільному функціонуванню людського суспільства. Проте позиція протиставлення техніки та людини як антагоністичних начал є необґрунтованою з декількох причин. По-перше, ці феномени не рівнозначні, а пов'язані генетичним зв'язком, відношенням походження та наслідування (людина створює техніку, наслідуючи зразкам власної природи та природи в широкому сенсі терміну). По-друге, ці поняття не є протилежними, контрадикторними за обсягом, їхня сукупність не утворює поняття більш високого порядку. Тому і їхня опозиційність є штучною, тобто хибною. По-третє, сам факт наявності техніки в нашому житті, так само як і темпи зростання рівня її складності та кількості в цілому є аргументом на користь її необхідності, а не шкоди.

Техніка та технології суттєво розширюють сферу можливостей людини. Сучасний етап розвитку технологій вирізняється тим фактом, що таке доповнення відбувається не стільки в контексті фізичної праці, а у фокусі інтелектуальної діяльності, яка охоплює також широке коло питань світоглядного, етичного та аксіологічного характеру. На цьому тлі виникають уявлення про сингулярність як невідворотну перспективу набуття суб'єктності роботами або програмами штучного інтелекту. Тому дослідники пропонують змінити принциповий підхід до технологічного прогресу. Суть такої зміни полягає у розробці винятково людино-центрованих технологій, в яких ухвалює рішення та робить висновки людина як суб'єкт біологічного походження. Про-

те така взаємодія є комплементарною: людина створює і вдосконалює технології, а технології стають більш ефективними для розв'язання поставлених завдань. Відповідно, якщо штучний інтелект потенційно здатний конкурувати з людиною на основі автономії ухвалення рішень, то розширений інтелект є операційною системою, що виконує чітко визначений функціонал.

Розвиток техніки та технологій є, очевидно, наслідком цивілізаційного розвитку, накопичення і збагачення знань. Існує думка, що вся історія метафізики (не тільки західного зразка) є прототипом «над-людського», або розширеного інтелекту (Negarestani R., 2018). Тому технології доповненого інтелекту спрямовані на якісно новий формат роботи з аналітикою та оперуванням даними, відповідно до якого позиція суб'єктності в ухваленні рішень, формуванні суджень та оцінок здійснюється винятково людиною на основі ідеалів та цінностей, які в традиціях західної метафізики визнано трансцендентальними сутностями.

Отже, поширена думка про надзвичайно швидкий розвиток робототехніки та ШІ є перебільшенням. Лабораторні експерименти часто не виправдовують очікувань фахівців. Необхідні у виробництві інтуїтивні здібності, креативність, уміння фантазувати та діяти на основі своїх образів ми можемо назвати недосяжною мрією техніків зараз. Значимо також, що розвиток робототехніки обмежений принципами фізичної мобільності (діапазон діяльності робота завжди чітко визначений) і прийняття рішень (ключові переваги щодо цілеспрямованості належать людині). З огляду на все вищесказане, таке запитання було б доречним: «...щоб мати справу з потенційними негативними

та позитивними наслідками НРС для якості роботи, ми не повинні зосереджуватися на питанні, чи повинні люди та роботи працювати разом, а як» [2, с. 20].

Взаємодія з робототехнікою та ШІ залежить від рівня рефлексії та критичного мислення людини. Оскільки людина є «незавершеним проектом», ми маємо величезну загрозу в екстраполяції нашої сутності в машину або прийнятті її характеристик у своїй поведінці. «Золота середина» у даному разі – взаємодоповнюючі відносини. Правильне поєднання між механізованою працею, автоматизацією праці, людською творчістю та прийняттям рішень здається хорошим рішенням для майбутнього розвитку. Це може бути реалізовано шляхом гнучкої та комунікативної співпраці між фахівцями різних галузей науки, освіти та виробництва. Такий підхід доречно називають світоглядом «направляючої філософії», що не орієнтована

на загальне рішення множини можливих проблем та один рядок коду у виконанні різноманітних завдань. Її доцільність полягає в орієнтації на різноманіття.

Ми провели багато паралелей, оскільки певні концепції минулого визначають зміст поточної аналітики взаємодії людини та машини. Зробимо це знову. Уявлення про сингулярність, яка незабаром і неминуче призведе до знищення людства, насправді є сучасною версією апокаліптичних релігійних вірувань. Передчуття кінця світу відоме ще з часів давньоєгипетських фараонів, помилково приписується циклічному календарю племені майя, ґрунтовно аргументовано в іудаїзмі та християнстві, а в контексті сьогодення набуває нових аргументів і можливих причин. Проте наше майбутнє – це проєкт, який ми створюємо. І ми, безперечно, маємо можливість зробити цю роботу якісно.

ЛІТЕРАТУРА

1. „What is Augmented Intelligence? – IEEE Digital Reality“. URL: <https://digitalreality.ieee.org/publications/what-is-augmented-intelligence> (zugegriffen 24. Juni 2022).
2. **Baltrusch, S. J., Krause, F., de Vries, A. W., van Dijk, W., de Looze, M. P.** (2022). What about the human in human robot collaboration? *Ergonomics*, 65:5, 719–740, p.20. doi:10.1080/00140139.2021.1984585
3. **Bui, L. T.** (2020) Asian Roboticism: Connecting mechanized labor to the automation of work. *Perspectives on global development and technology*, 19(1–2), 110–126. doi:10.1163/15691497-12341544
4. **Carr, N.** (2014) *The Glass Cage*. New York: Norton.
5. **Draeger, J., Miller-Eiselt, R.** (2019). *Wir und die intelligenten Maschinen: wie Algorithmen unser Leben bestimmen und wir sie für uns nutzen können*, 1. Auflage. München: Deutsche Verlags-Anstalt.
6. **Eyre, H. A., Berk, M., Lavretsky, H., Reynolds, III C.** (2021). *Convergence Mental Health: A Transdisciplinary Approach to Innovation*. Oxford University Press.
7. **Eubanks, V.** (2019). *Automating inequality: how high-tech tools profile, police, and punish the poor*. First Picador edition. New York: Picador St. Martin's Press.
8. **Ford, M.** (2015) *The Rise of the Robots*. New York: Basic Books.
9. **Fried, J., Weitman, M., & Davis, M. K.** (1972). Man-machine interaction and absenteeism. *Journal of Applied Psychology*, 56(5), 428–429. doi.org/10.1037/h0033591
10. **Head, S.** (2014) *Mindless*. New York: Basic Books.
11. **Hentout, A., Aouache, M, Maoudj, A., Akli, I.** (2019). Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017. *Advanced Robotics*, 33:15–16, 764–799. doi:10.1080/01691864.2019.1636714
12. **Hertweck, C., Heitz, C., Loi, M.** (2021) „On the Moral Justification of Statistical Parity“, in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Virtual Event Canada, M rz 2021, 747–757. doi: 10.1145/3442188.3445936
13. **Huws, U.** (2014). *Labor in the Global Digital Economy: The Cybertariat Comes of Age*. Monthly review press.
14. **Loi, M., Heitz, C.** (2022). „Is calibration a fairness requirement?: An argument from the point of view of moral philosophy and decision theory“, in *2022 ACM Conference on Fairness, Accountability, and Transparency*, Seoul Republic of Korea, Juni 2022, 2026–2034. doi: 10.1145/3531146.3533245
15. **Malecki, E. J., Moriset, B.** (2007). *The Digital Economy. Business Organization, Production Processes and Regional Developments*. Routledge.
16. **Murray, A., Rhymer, J., & Sirmon, D. G.** (2021). Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review*, 46(3), 552–571.
17. **Negarestani, R.** (2018) *Intelligence and Spirit*. Falmouth: Urbanomic.
18. **O’Neil, C.** (2016). *Weapons of math destruction: how big data increases inequality and threatens democracy*. First edition. New York: Crown.
19. **Stadnicka, D., Antonelli, D.** (2019). Human-robot collaborative work cell implementation through lean thinking. *International Journal of Computer Integrated Manufacturing*, 32:6, 580–595. doi:10.1080/0951192X.2019.1599437