

УДК 004.8:001.8

DOI <https://doi.org/10.30970/PPS.2026.64.11>

ШТУЧНИЙ ІНТЕЛЕКТ ЯК СУБ'ЄКТ НАУКОВОГО ПІЗНАННЯ

Антон Іскрик

*Український державний університет імені Михайла Драгоманова,
кафедра соціальної філософії, філософії освіти та освітньої політики
вул. Пирогова, 9, 01601, м. Київ, Україна
<https://orcid.org/0009-0000-2995-8376>*

Актуальність дослідження зумовлено стрімким поширенням ШІ у науково-дослідній діяльності та зростанням його впливу на процеси продукування, аналізу й інтерпретації наукового знання, що актуалізує потребу філософсько-епістемологічного осмислення статусу ШІ у структурі наукового пізнання. За відсутності усталених теоретичних підходів до визначення меж автономності, відповідальності та інтерпретованості результатів, отриманих з використанням ШІ, виникає ризик методологічної невизначеності та зниження наукової обґрунтованості досліджень.

Метою статті є теоретичне обґрунтування наукової доцільності трактування ШІ як специфічного елемента суб'єктної структури наукового пізнання та визначення епістемологічних і прикладних наслідків такого підходу для сучасної науково-дослідної практики.

Методи дослідження ґрунтуються на філософсько-науковому аналізі, системному та функціонально-епістемологічному підходах, узагальненні сучасних міждисциплінарних досліджень у галузі філософії науки та аналізі практик використання ШІ в різних наукових напрямках.

Результати дослідження свідчать про те, що ШІ виконує власне пізнавальні функції, здатні впливати на формування наукових результатів, зокрема через автономний аналіз даних, побудову моделей і генерацію гіпотез. Встановлено, що відсутність у ШІ інтенціональності, рефлексії та епістемологічної відповідальності обмежує можливість його трактування як повноцінного суб'єкта наукового пізнання. Доведено доцільність використання поняття квазісуб'єкта для фіксації проміжного статусу ШІ в науці. Виявлено ключові науково-практичні проблеми застосування ШІ, пов'язані з інтерпретованістю результатів, розмитістю відповідальності та алгоритмічними викривленнями.

Висновки полягають у тому, що ефективне й науково коректне використання ШІ можливе лише за умов чіткого методологічного контролю, збереження людської епістемологічної відповідальності та прозорості алгоритмічних процедур.

Перспективи подальших досліджень пов'язані з розробленням формалізованих моделей людино-алгоритмічного пізнання, удосконаленням інтерпретованих методів ШІ та поглибленим аналізуванням трансформації наукової раціональності в умовах цифровізації.

Ключові слова: філософія науки, епістемологія, наукова раціональність, цифровізація науки, квазісуб'єктність, інтерпретованість знань, алгоритмічні викривлення, наукова відповідальність, людино-алгоритмічна взаємодія.

Постановка проблеми. Стрімкий розвиток систем штучного інтелекту (ШІ) зумовлює необхідність переосмислення їхньої ролі у сучасному науковому пізнанні не лише як інструментів оброблення даних, але й як активних учасників пізнавального процесу. У межах класичної епістемології суб'єкт наукового пізнання традиційно ототожнювався з людиною-дослідником, наділеною свідомістю, інтенціональністю та здатністю до рефлексії, тоді як технічні засоби розглядалися як допоміжні. Однак поширення алгоритмів машинного навчання, здатних автономно виявляти закономірності, формувати гіпотези, оптимізувати моделі та генерувати науково значущі результати, актуалізує проблему



© Іскрик А., 2026

Стаття поширюється на умовах ліцензії відкритого доступу (CC BY 4.0)

розширення або перегляду цього уявлення. Виникає науково значуща суперечність між зростаючою пізнавальною спроможністю ШІ та відсутністю усталених теоретичних підходів до його трактування як можливого суб'єкта або квазісуб'єкта наукового пізнання.

Зазначена проблема має безпосередній зв'язок з важливими науковими та практичними завданнями сучасної науки, зокрема з розробленням нових епістемологічних моделей пізнання в умовах цифровізації, уточненням меж відповідальності та авторства наукових результатів, отриманих і використаних ШІ, а також із забезпеченням надійності, відтворюваності та інтерпретованості знань, згенерованих алгоритмічними методами. У практичному вимірі проблема статусу ШІ в науковому пізнанні пов'язана з потребами підвищення ефективності досліджень у галузях з великими масивами даних, формуванням стандартів використання ШІ в науковій діяльності, а також з мінімізацією ризиків алгоритмічних викривлень і некритичного перенесення машинно згенерованих висновків у наукову та прикладну практику. Таким чином, осмислення ШІ як потенційного суб'єкта наукового пізнання є комплексною міждисциплінарною проблемою, що перебуває на перетині філософії науки, інформатики та прикладних дослідницьких практик і має істотне значення для подальшого розвитку наукового знання.

Аналіз останніх досліджень і публікацій. Огляд сучасних досліджень засвідчує системну трансформацію класичного уявлення про суб'єктність у науці та поступове формування розширених моделей пізнання, у межах яких інтелектуальні функції частково делегуються технічним системам. У цьому контексті О. Варипаєв обґрунтовує деконструкцію класичного суб'єкта науки та формування нової онтології пізнання, в якій штучний інтелект виступає не лише інструментом, але й структурним компонентом когнітивного процесу [1]. Подібну логіку розвивають О. Петрунько та О. Плющ, які розглядають взаємодію людини і штучного інтелекту як співучасть у продукуванні знання, наголошуючи на обмеженості традиційних пояснювальних моделей і необхідності нової епістемологічної парадигми [2]. Н. Козаченко, звертаючись до епістемології чеснот, підкреслює, що використання штучного інтелекту змінює характер пізнавальної відповідальності, ставлячи під сумнів сталі уявлення про авторство та академічну добросовісність у науковому пізнанні [3].

Цивілізаційний і антропологічний вимір проблеми суб'єктності штучного інтелекту розкривається у роботах, де він осмислюється як чинник глибоких змін у структурі наукової раціональності. О. Петрунько та О. Плющ розглядають добу штучного інтелекту як етап радикального перегляду меж людської автономії та ролі людини в процесах наукового пізнання [4]. В. Нікітенко та співавтори акцентують увагу на філософській рефлексії впливу штучного інтелекту на людину, суспільство й освіту, наголошуючи на ризиках втрати гуманістичних засад науки в умовах алгоритмізації пізнавальних процесів [5]. О. Іванова аналізує етичні виклики штучного інтелекту з антропологічної перспективи, підкреслюючи небезпеку редукції людської суб'єктності та відповідальності в умовах зростання автономності алгоритмічних систем [6].

Важливий пласт досліджень зосереджується на когнітивному аналізі можливостей і меж штучного інтелекту як квазісуб'єкта пізнання. О. Варипаєв зі співавторами, аналізуючи співвідношення емоційного та штучного інтелекту, демонструє принципові обмеження алгоритмічних систем у сфері смислотворення та рефлексії, що є ключовими характеристиками суб'єктного пізнання [7]. Г. Ксіє (H. Xie) розглядає перспективи зближення когнітивної науки та штучного інтелекту, водночас наголошуючи на збереженні фундаментальної відмінності між людською та штучною когніцією [8]. Дж. Сіменс (G. Siemens) та співавтори аналізують людську і штучну когніцію як взаємопов'язані, але асиметричні системи, що формують спільний пізнавальний простір сучасної науки [9]. А. Банді (A. Bundy) та

співавтори вводять поняття когнітивного штучного інтелекту, підкреслюючи його здатність до навчання й узагальнення без набуття повноцінної епістемологічної автономії [10].

Прикладний вимір проблеми штучного інтелекту як суб'єкта пізнання представлений дослідженнями, що аналізують реальні науково-дослідні практики взаємодії людини і алгоритмічних систем. В. Федорець (V. Fedorets) та співавтори досліджують когнітивні аспекти функціонування системи «людина – штучний інтелект», у якій ШІ виступає активним учасником процесів прийняття наукових рішень, не набуваючи, однак, статусу автономного суб'єкта [11]. Й. Донг (Y. Dong) та співавтори аналізують вплив людського інтелекту, свідомості та когнітивних моделей на розвиток штучного інтелекту, підкреслюючи його похідний характер щодо людського пізнання [12]. Й. Маруяма (Y. Maruyama), досліджуючи раціональність і когнітивні викривлення, виявляє структурні обмеження штучного інтелекту, що унеможлиблюють його повноцінне ототожнення із суб'єктом наукового пізнання [13].

Попри активне дослідження ролі ШІ в сучасній науці, залишаються не вирішеними питання цілісного епістемологічного осмислення трансформації суб'єкта наукового пізнання, чіткого розмежування пізнавальних функцій ШІ за рівнями автономності та визначення теоретично обґрунтованих меж трактування ШІ як суб'єкта або квазісуб'єкта. Недостатньо розробленими є методологічні підходи до інтерпретації результатів, отриманих за участю ШІ, а також проблеми відповідальності та алгоритмічних викривлень, що обмежує наукову валідність і практичну застосовність таких результатів.

Запропоноване дослідження спрямовано на подолання цих прогалин шляхом системного аналізу еволюції уявлень про суб'єкта наукового пізнання, концептуалізації пізнавальних функцій ШІ та обґрунтування його статусу як квазісуб'єкта. Це дає змогу зменшити методологічну невизначеність і сформувати науково обґрунтовані підходи до відповідального використання ШІ в науково-дослідній діяльності.

Мета і завдання статті. Метою статті є обґрунтування наукової доцільності розгляду ШІ як специфічного суб'єкта наукового пізнання та визначення епістемологічних і практичних наслідків такого підходу для сучасної науково-дослідної діяльності.

Для досягнення поставленої мети у статті передбачено вирішення таких завдань:

- проаналізувати еволюцію уявлень про суб'єкта наукового пізнання та пізнавальні функції ШІ в сучасних наукових дослідженнях;
- визначити теоретичні підстави, концептуальні межі й ключові науково-практичні проблеми трактування ШІ як суб'єкта або квазісуб'єкта наукового пізнання;
- обґрунтувати рекомендації щодо відповідального застосування ШІ в науково-дослідній діяльності.

Виклад основного матеріалу. У філософії науки уявлення про суб'єкта наукового пізнання зазнали суттєвої трансформації, що відображає зміну характеру наукової раціональності та засобів пізнавальної діяльності. У класичних концепціях суб'єкт пізнання ототожнювався з індивідуальним дослідником, здатним до раціонального мислення, спостереження та логічного узагальнення, а процес пізнання розглядався як відносно лінійний і детермінований. Некласична філософія науки змістила акцент на обмеженість суб'єкта, залежність результатів пізнання від методів, теоретичних рамок і контексту дослідження, визнаючи роль інтерсуб'єктивності та наукових спільнот. У постнекласичних підходах суб'єкт наукового пізнання дедалі частіше осмислюється як розширена система, що включає не лише людину, але й складні технічні, інформаційні та алгоритмічні компоненти, серед яких ШІ відіграє дедалі вагомішу роль. Така еволюція відображає перехід від антропоцентричної моделі пізнання до системно-мережевої, у якій когнітивні функції розподіляються між людськими та штучними агентами (табл. 1).

Таблиця 1

**Еволюція уявлень про суб'єкта наукового пізнання
у філософсько-наукових концепціях**

Етап розвитку науки	Домінуюче уявлення про суб'єкта пізнання	Характерні риси пізнавальної діяльності
Класичний	Індивідуальний дослідник	Раціональність, об'єктивність, детермінізм, інструментальний характер техніки
Некласичний	Обмежений суб'єкт у науковій спільноті	Залежність знання від методів і теорій, інтерсуб'єктивність, контекстуальність
Постнекласичний	Розширений суб'єкт (людина + технічні системи)	Системність, нелінійність, інтеграція цифрових і алгоритмічних засобів, участь ШІ

Джерело: сформовано автором на основі [2; 4; 5;8; 11]

У сучасних умовах постнекласичне розуміння суб'єкта наукового пізнання набуває практичного втілення через системну інтеграцію ШІ в процеси продукування, аналізу та інтерпретації наукового знання, що трансформує традиційну модель індивідуальної дослідницької діяльності у розподілену когнітивну систему. Показовим у цьому контексті є досвід застосування ШІ в структурній біології, де алгоритмічна система AlphaFold забезпечила принципово новий рівень точності прогнозування просторової структури білків. Функціонально ШІ в цьому разі реалізує складні пізнавальні операції моделювання та узагальнення, які раніше базувалися переважно на емпіричних методах і експертній інтуїції, тоді як науковець виконує роль концептуального координатора, інтерпретатора та верифікатора результатів, що свідчить про зміну статусу суб'єкта пізнання з індивідуального на системно-розширений [14]. Аналогічні тенденції спостерігаються у фізиці високих енергій, зокрема в дослідженнях, пов'язаних з аналізом даних Великого адронного колайдера. Застосування методів машинного навчання для оброблення надвеликих масивів експериментальних даних дає змогу ідентифікувати рідкісні події та статистично значущі закономірності, які не можуть бути виявлені традиційними аналітичними підходами. У цьому разі ШІ виступає як елемент колективного суб'єкта наукового пізнання, інтегрованого в інфраструктуру наукової спільноти, де когнітивні функції розподіляються між дослідниками, експериментальними установками та алгоритмічними системами, забезпечуючи нову якість наукового знання [15]. Ще одним прикладом практичної реалізації постнекласичної моделі є використання ШІ в матеріалознавстві для прискорення відкриття та проєктування нових матеріалів. Алгоритмічні системи аналізують великі масиви експериментальних і симуляційних даних, прогнозуючи фізико-хімічні властивості матеріалів і пропонуючи перспективні напрями досліджень, які не завжди є очевидними для дослідника. У такому підході ШІ фактично бере участь у формуванні наукових гіпотез, що зумовлює перехід від суто інструментального використання цифрових технологій до їх включення в пізнавальну структуру науки. Практичний ефект цього полягає у скороченні циклу «гіпотеза – експеримент – інтерпретація», водночас актуалізуючи проблеми інтерпретованості моделей, контролю наукової обґрунтованості та відповідальності за результати, отримані за участю ШІ [16]. Наведені приклади в сукупності ілюструють не окремі випадки застосування цифрових інструментів, а системну трансформацію суб'єкта наукового пізнання, у межах якої ШІ інтегрується в когнітивну архітектуру сучасної науки, змінюючи співвідношення між людською раціональністю та алгоритмічними формами пізнання.

У сучасній науці ШІ дедалі частіше виконує не лише допоміжні, але й власне пізнавальні функції, що зумовлює зміну його статусу в структурі наукового дослідження. Якщо на ранніх етапах цифровізації ШІ розглядався переважно як інструмент автоматизації обчислень і оброблення даних, то сьогодні його застосування характеризується зростаючим рівнем автономності та активною участю у формуванні наукових результатів [7]. Пізнавальна діяльність ШІ реалізується через здатність самостійно виявляти закономірності у складних багатовимірних даних, будувати моделі об'єктів і процесів, оптимізувати параметри досліджень і генерувати проміжні або кінцеві результати, що мають наукову новизну. За таких умов наукове пізнання набуває характеру спільної діяльності людини й алгоритмічних систем, де функції аналізу, узагальнення та прогнозування частково делегуються ШІ, а роль дослідника зосереджується на постановці проблем, методологічному контролі та інтерпретації результатів (табл. 2).

Таблиця 2

Пізнавальні функції ШІ в наукових дослідженнях за рівнем автономності та участі у формуванні результатів

Пізнавальна функція ШІ	Характер автономності	Участь у формуванні наукових результатів
Оброблення та структурування даних	Обмежена, за заданими алгоритмами	Забезпечує підготовку даних для подальшого аналізу
Виявлення закономірностей і кореляцій	Часткова автономність	Формує нові аналітичні узагальнення
Побудова моделей і прогнозування	Висока автономність у межах заданої задачі	Генерує науково значущі моделі та прогнози
Генерація гіпотез і дослідницьких напрямів	Обмежено висока, з опорою на дані	Сприяє формуванню нових наукових результатів
Оптимізація експериментів і рішень	Висока автономність	Підвищує ефективність і точність досліджень

Джерело: сформовано автором на основі [1; 5; 8; 9; 12]

Пізнавальні функції ШІ в сучасній науковій практиці реалізуються не декларативно, а через конкретні дослідницькі механізми, у межах яких автономність алгоритмів безпосередньо впливає на формування наукових результатів і зміну ролі людського дослідника. У галузі кліматичних досліджень ШІ застосовується для побудови складних прогностичних моделей, здатних відтворювати нелінійну динаміку кліматичної системи Землі на підставі великих масивів супутникових і реаналітичних даних. Зокрема, використання глибоких нейронних мереж для даунскейлінгу кліматичних моделей дає змогу автономно формувати детальні просторові прогнози, які безпосередньо використовуються у наукових публікаціях і прикладних сценаріях адаптації до кліматичних змін. У цьому разі ШІ виконує пізнавальну функцію моделювання та узагальнення, тоді як людина забезпечує концептуальну інтерпретацію результатів і їх включення в теоретичні кліматичні рамки [17]. У сучасній астрономії пізнавальні функції ШІ проявляються через автономний аналіз спостережних даних задля відкриття нових астрономічних об'єктів. Алгоритми машинного навчання застосовуються для оброблення світлових кривих, отриманих космічними телескопами, та виявлення сигналів, характерних для транзиту екзопланет. Такі системи здатні самостійно відбирати кандидати для подальшого підтвердження, що фактично означає участь ШІ у первинному етапі наукового відкриття. Науковий результат у цьому разі формується як наслідок людино-алгоритмічної взаємодії, де ШІ здійснює когнітивно насичений відбір

і класифікацію, а дослідник відповідає за верифікацію та теоретичне осмислення отриманих даних [18]. У фармацевтичних і біомедичних дослідженнях ШІ дедалі активніше залучається до процесів відкриття нових лікарських засобів, виконуючи функції прогнозування біологічної активності молекул і генерації перспективних хімічних структур. Алгоритмічні системи аналізують великі хімічні бібліотеки та біомедичні бази даних, автономно пропонуючи кандидатів, які раніше не розглядалися в межах класичних підходів. У такій практиці ШІ фактично бере участь у формуванні наукових гіпотез, що зумовлює зміщення меж між інструментальним використанням технологій і їх включенням до когнітивної структури дослідження. Людський дослідник при цьому виконує функцію епістемологічного контролю, забезпечуючи експериментальну перевірку та наукову легітимацію результатів, згенерованих за участю ШІ [19]. Наведені приклади свідчать про те, що автономність ШІ в сучасній науці є не абсолютною, проте достатньою для виконання власне пізнавальних функцій, які впливають на формування наукових результатів. У філософсько-науковому вимірі це означає перехід до моделі розширеного суб'єкта пізнання, у межах якої ШІ інтегрується в наукову раціональність як активний когнітивний компонент, зберігаючи водночас потребу в людській інтерпретації та відповідальності.

Теоретичні підстави трактування ШІ як суб'єкта або квазісуб'єкта наукового пізнання формуються в межах функціонально-епістемологічного підходу, у якому суб'єкт визначається не через наявність свідомості, а через здатність виконувати пізнавальні операції, що призводять до отримання нового знання. У цьому сенсі ШІ може бути включений до пізнавальної структури науки як агент, що автономно реалізує аналіз, моделювання та узагальнення в межах заданих методологічних рамок [3]. Водночас концептуальні межі такого включення зумовлені відсутністю в ШІ інтенціональності, рефлексії та епістемологічної відповідальності, що унеможлиблює його повноцінне ототожнення з людським суб'єктом і обґрунтовує доцільність використання поняття квазісуб'єкта наукового пізнання (табл. 3).

Таблиця 3

Теоретичні підстави та концептуальні межі трактування ШІ як суб'єкта наукового пізнання

Критерій	Суб'єкт наукового пізнання	Квазісуб'єкт (ШІ)
Джерело пізнавальної активності	Внутрішня мотивація та інтенціональність	Зовнішньо задані цілі та алгоритми
Здатність до рефлексії	Притаманна	Відсутня
Участь у формуванні знання	Осмислення, інтерпретація, оцінка істинності	Генерація моделей і результатів у межах заданих рамок
Епістемологічна відповідальність	Персоналізована	Делегована людині
Статус у науковому пізнанні	Повноцінний суб'єкт	Функціональний квазісуб'єкт

Джерело: сформовано автором на основі [3; 4; 9; 11; 12]

Теоретичні підстави та концептуальні межі трактування ШІ як суб'єкта або квазісуб'єкта наукового пізнання отримують чітке практичне підтвердження в сучасних формах організації досліджень, де автономні алгоритмічні системи інтегруються в пізнавальний процес за умови збереження людської епістемологічної відповідальності. Характерним прикладом є автоматизована наукова система Adam, здатна самостійно здійснювати повний цикл пізнавальних операцій у межах функціональної геноміки, включно з формуванням гіпотез, плануванням експериментів і аналізом результатів. Незважаючи на високий

рівень автономності, концептуальні межі її «суб'єктності» визначаються зовнішньо заданими дослідницькими цілями, методологічними обмеженнями та критеріями наукової значущості, що обґрунтовує трактування ІІІ як квазісуб'єкта, позбавленого власної інтенційності та відповідальності за істинність отриманого знання [20].

Аналогічна логіка простежується у використанні фізично-інформованих нейронних мереж для розв'язання задач математичного моделювання фізичних процесів. Такі системи здатні автономно будувати узгоджені з фундаментальними законами моделі на основі даних і формальних обмежень, що надає їм ознак пізнавальної активності. Водночас вони не здійснюють теоретичної рефлексії щодо статусу отриманих моделей у структурі наукового знання та не визначають їхньої пояснювальної цінності, що зберігає вирішальну роль людини у визнанні результатів як науково значущих. Це підтверджує функціональний, а не онтологічний характер «суб'єктності» ІІІ та концептуально фіксує його статус як квазісуб'єкта наукового пізнання [21].

Використання ІІІ в науковому пізнанні супроводжується комплексом науково-практичних проблем, що безпосередньо впливають на якість, надійність і легітимність отриманих результатів. Однією з базових проблем є обмежена інтерпретованість алгоритмічних моделей, особливо у разі глибоких нейронних мереж, результати яких часто мають характер статистично коректних, але концептуально непрозорих залежностей. За таких умов ускладнюється встановлення причинно-наслідкових зв'язків, що є критично важливим для теоретичного пояснення явищ і перевірки відповідності результатів наявним науковим теоріям. Непрозорість моделей знижує відтворюваність досліджень і ускладнює наукову верифікацію, перетворюючи окремі результати на «чорні скриньки» без належного епістемологічного обґрунтування. Тісно пов'язаною з інтерпретованістю є проблема наукової відповідальності за результати, отримані із застосуванням ІІІ. У ситуації розподіленої когнітивної діяльності між дослідником і алгоритмічною системою розмивається суб'єкт відповідальності за помилки, хибні узагальнення або некоректні висновки [4]. Це особливо критично в міждисциплінарних і прикладних дослідженнях, де наукові результати можуть мати соціальні, технологічні або регуляторні наслідки. Відсутність чітких методологічних критеріїв розмежування ролей людини і ІІІ у формуванні знання ускладнює як етичну, так і наукову оцінку результатів. Суттєвою науково-практичною проблемою є також наявність алгоритмічних викривлень, зумовлених якістю навчальних даних, вибором архітектури моделей і параметрів оптимізації. Такі викривлення можуть відтворювати або підсилювати приховані статистичні, методологічні чи соціальні упередження, що призводить до систематичних помилок у наукових висновках [13]. У науковому пізнанні це створює ризик хибної універсалізації результатів і зниження їхньої теоретичної валідності, особливо за умов обмежених або нерепрезентативних даних. Додатковою проблемою є залежність наукових результатів від програмних реалізацій і обчислювальної інфраструктури ІІІ, що ускладнює відтворюваність досліджень і перевірку отриманих висновків незалежними науковими групами. Нерідко алгоритмічні моделі є складними для повторної реалізації через закритість коду, відсутність повної документації або залежність від специфічних обчислювальних ресурсів, що суперечить базовим принципам наукової відкритості [11].

Відповідальне застосування ІІІ в науково-дослідній діяльності має ґрунтуватися на принципі методологічної підконтрольності алгоритмічних результатів людині-досліднику та на пріоритеті наукової відтворюваності над продуктивністю обчислень. Доцільно фіксувати роль ІІІ в дизайні дослідження як інструмента або квазісуб'єкта з чітким описом меж автономності, умов застосування, параметрів моделей і точок, у яких рішення переходять

до експертної інтерпретації. Необхідним є запровадження процедур інтерпретаційної валідації, коли кожний суттєвий висновок, сформульований за участю ШІ, супроводжується пояснювальними аргументами, перевірками чутливості, аналізом альтернативних моделей і контролем щодо причинно-наслідкової узгодженості з теоретичними положеннями предметної галузі.

Для мінімізації алгоритмічних викривлень слід забезпечувати доказову якість даних шляхом документування їх походження, критеріїв включення/виключення, репрезентативності та потенційних упереджень, а також застосовувати стандартизовані процедури тестування моделей на зсуви вибірок, нестійкість до змін середовища й приховані конфаундери. Важливою є практика незалежної перевірки результатів, яка передбачає відокремлення етапів навчання та оцінювання, використання зовнішніх наборів даних, реплікацію ключових експериментів іншими дослідницькими групами та обов'язкове розкриття метрик невизначеності, меж застосовності і сценаріїв, у яких модель може давати хибні висновки. Для забезпечення відтворюваності досліджень доцільно впроваджувати політики повного протоколювання обчислювальних експериментів, що включає версії даних, коду, середовища виконання, гіперпараметрів, випадкових зерен, апаратних обмежень і критеріїв зупинки навчання. Рекомендованою є публікація репозитаріїв або, за наявності правових обмежень, надання контрольованого доступу до коду й синтетичних або деперсоналізованих даних, достатніх для незалежної верифікації результатів. У високоризикових галузях слід застосовувати принцип «людина в контурі» як обов'язкову умову прийняття кінцевих наукових інтерпретацій і прикладних рішень, а також використовувати моделі з вищою інтерпретованістю, якщо це не погіршує валідність результатів. Задля відповідальності та наукової добросовісності необхідно інституціоналізувати прозоре зазначення участі ШІ у підготовці даних, аналізі, генерації тексту або візуалізації, з розмежуванням авторства, інструментальної підтримки та автоматизованого внеску. Доцільно закріплювати внутрішні протоколи рецензування ШІ-результатів, де окремо перевіряються відтворюваність, інтерпретованість, стійкість до зміни даних і коректність висновків. Етична складова частина відповідального застосування ШІ має включати оцінку потенційних наслідків помилок і викривлень, механізми їх виявлення та виправлення, а також процедури керування ризиками в публікації та трансфері результатів у практику. У сукупності ці рекомендації формують рамку відповідального використання ШІ, за якої алгоритмічна ефективність підпорядковується вимогам наукової обґрунтованості, прозорості та контрольованості знання.

Висновки. У ході дослідження встановлено, що розвиток ШІ зумовлює перехід від класичного антропоцентричного розуміння суб'єкта наукового пізнання до постнекласичної моделі розподіленої когнітивної діяльності, у межах якої алгоритмічні системи виконують власне пізнавальні функції та впливають на формування наукових результатів. Водночас доведено, що відсутність у ШІ інтенціональності, рефлексії та епістемологічної відповідальності обмежує можливість його трактування як повноцінного суб'єкта, що обґрунтовує доцільність розгляду ШІ як квазісуб'єкта наукового пізнання. Виявлено ключові проблеми використання ШІ в науці, зокрема обмежену інтерпретованість результатів, розмитість відповідальності за наукові висновки, алгоритмічні викривлення та ускладнення відтворюваності досліджень. Обґрунтовано, що подолання цих проблем потребує методологічного контролю автономності ШІ, обов'язкової людської інтерпретації та прозорості обчислювальних процедур. Сформульовано рекомендації щодо відповідального застосування ШІ, які передбачають закріплення людської епістемологічної відповідальності, інтерпретаційну та реплікаційну валідацію результатів і системний контроль

алгоритмічних викривлень. Перспективи подальших досліджень пов'язані з розвитком епістемологічних моделей людино-алгоритмічного пізнання, інтерпретованих методів ШІ та аналізом трансформації наукової раціональності в умовах цифровізації.

Список використаної літератури

1. Варипаєв О. Філософія науки та штучний інтелект: деконструкція суб'єкта і нова онтологія пізнання. *Вісник гуманітарних наук*. 2025. Вип. 7. DOI: <https://doi.org/10.5281/zenodo.15525177>.
2. Петрунько О., Плющ О. Взаємодія людини і штучного інтелекту: пошук пояснювальної парадигми. *Вчені записки Університету «КРОК»*. 2025. Вип. 78, № 2. С. 424–431. DOI: <https://doi.org/10.31732/2663-2209-2025-78-424-431>.
3. Козаченко Н. Штучний інтелект і академічна доброчесність в контексті епістемології чеснот. *Актуальні проблеми філософії та духовності*. 2024. Вип. 25, № 1. С. 315–342. DOI: <https://doi.org/10.55056/apm.7740>.
4. Петрунько О., Плющ О. Доба штучного інтелекту: камо грядеші?. *Вчені записки Університету «КРОК»*. 2025. Вип. 80, № 4. С. 363–370. DOI: <https://doi.org/10.31732/2663-2209-2025-80-363-370>.
5. Нікітенко В., Воронкова В., Кивлюк О., Олексенко Р., Сухенко В. Філософська рефлексія про штучний інтелект та його вплив на розвиток суспільства, людини та освіти. *Humanities Studies*. 2024. Вип. 96, № 19. С. 67–76. DOI: <https://doi.org/10.32782/hst-2024-19-96-07>.
6. Іванова О. Етичні виклики штучного інтелекту: антропологічний вимір людської суб'єктності. *Актуальні проблеми філософії та соціології*. 2024. Вип. 51. С. 72–76. DOI: <https://doi.org/10.32782/apfs.v051.2024.12>.
7. Варипаєв О., Байрамова О., Сільвестрова О. Емоційний інтелект та штучний інтелект: філософська рефлексія суб'єктності в процесах пізнання. *Філософія та управління*. 2025. Вип. 6, № 2. Article ID: epg0032. DOI: <https://doi.org/10.70651/3041-248X/2025.2.02>.
8. Xie H. The promising future of cognitive science and artificial intelligence. *Nature Reviews Psychology*. 2023. Vol. 2, № 4. P. 202–202. DOI: <https://doi.org/10.1038/s44159-023-00170-3>.
9. Fedorets V., Klochko O., Tverdokhlib I., Sharyhin O. Cognitive aspects of interaction in the “Human – Artificial Intelligence” system. *Journal of Physics: Conference Series*. 2024. Vol. 2871, № 1. Article ID: 012023. DOI: <https://doi.org/10.1088/1742-6596/2871/1/012023>.
10. Siemens G., Marmolejo-Ramos F., Gabriel F., Medeiros K., Marrone R., Joksimovic S., de Laat M. Human and artificial cognition. *Computers and Education: Artificial Intelligence*. 2022. Vol. 3. Article ID: 100107. DOI: <https://doi.org/10.1016/j.caeai.2022.100107>.
11. Bundy A., Chater N., Muggleton S. Introduction to “Cognitive artificial intelligence”. *Philosophical Transactions of the Royal Society A*. 2023. Vol. 381, № 2251. Article ID: 20220051. DOI: <https://doi.org/10.1098/rsta.2022.0051>.
12. Dong Y., Hou J., Zhang N., Zhang M. Research on how human intelligence, consciousness, and cognitive computing affect the development of artificial intelligence. *Complexity*. 2020. Vol. 2020, № 1. Article ID: 1680845. DOI: <https://doi.org/10.1155/2020/1680845>.
13. Maruyama Y. Rationality, cognitive bias, and artificial intelligence: a structural perspective on quantum cognitive science. *Lecture Notes in Computer Science*. 2020. P. 172–188. DOI: https://doi.org/10.1007/978-3-030-49183-3_14.
14. Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., Hassabis D. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021. Vol. 596, № 7873. P. 583–589. DOI: <https://doi.org/10.1038/s41586-021-03819-2>.

15. Guan W., Perdue G., Pesah A., Schuld M., Terashi K., Vallecorsa S., Vlimant J.R. Quantum machine learning in high energy physics. *Machine Learning: Science and Technology*. 2021. Vol. 2, № 1. Article ID: 011003. DOI: <https://doi.org/10.1088/2632-2153/abc17d>.
16. Pyzer-Knapp E.O., Pitera J.W., Staar P.W., Takeda S., Laino T., Sanders D.P., Curioni A. Accelerating materials discovery using artificial intelligence, high performance computing and robotics. *NPJ Computational Materials*. 2022. Vol. 8, № 1. Article ID: 84. DOI: <https://doi.org/10.1038/s41524-022-00765-z>.
17. Ravuri S., Lenc K., Willson M., Kangin D., Lam R., Mirowski P., Mohamed S. Skillful precipitation nowcasting using deep generative models of radar. *Nature*. 2021. Vol. 597, № 7878. P. 672–677. DOI: <https://doi.org/10.1038/s41586-021-03854-z>.
18. Yu L., Vanderburg A., Huang C., Shallue C.J., Crossfield I.J., Gaudi B.S., Quinn S.N. Identifying exoplanets with deep learning. III. Automated triage and vetting of TESS candidates. *The Astronomical Journal*. 2019. Vol. 158, № 1. Article ID: 25. DOI: <https://doi.org/10.3847/1538-3881/ab21d6>.
19. Tiwari P.C., Pal R., Chaudhary M.J., Nath R. Artificial intelligence revolutionizing drug development: Exploring opportunities and challenges. *Drug Development Research*. 2023. Vol. 84, № 8. P. 1652–1663. DOI: <https://doi.org/10.1002/ddr.22115>.
20. King R.D., Rowland J., Oliver S.G., Young M., Aubrey W., Byrne E., Clare A. The automation of science. *Science*. 2009. Vol. 324, № 5923. P. 85–89. DOI: <https://doi.org/10.1126/science.116562>.
21. Karniadakis G.E., Kevrekidis I.G., Lu L., Perdikaris P., Wang S., Yang L. Physics-informed machine learning. *Nature Reviews Physics*. 2021. Vol. 3, № 6. P. 422–440. DOI: <https://doi.org/10.1038/s42254-021-00314-5>.

ARTIFICIAL INTELLIGENCE AS A SUBJECT OF SCIENTIFIC COGNITION

Anton Iskryk

Dragomanov Ukrainian State University

Department of Social Philosophy, Philosophy of Education, and Educational Policy

Pyrohova str., 9, 01601, Kyiv, Ukraine

<https://orcid.org/0009-0000-2995-8376>

Relevance of the study is determined by the rapid integration of artificial intelligence (AI) into scientific research and its growing influence on the processes of knowledge production, analysis, and interpretation, which necessitates a philosophical and epistemological reconsideration of the status of AI within scientific cognition. In the absence of established theoretical frameworks for defining the limits of autonomy, responsibility, and interpretability of AI-generated results, risks of methodological uncertainty and reduced scientific validity emerge.

The purpose of the article is to provide a theoretical justification for the scientific expediency of considering AI as a specific element of the subject structure of scientific cognition and to determine the epistemological and practical implications of such an approach for contemporary research practice.

Research methods are based on philosophical and scientific analysis, systemic and functional-epistemological approaches, generalization of contemporary interdisciplinary studies in the philosophy of science, and analytical examination of current practices of AI application across different scientific domains.

Research results demonstrate that AI performs genuinely cognitive functions capable of influencing the formation of scientific results, particularly through autonomous data analysis, model construction, and hypothesis generation. It has been established that the absence of intentionality, reflexivity, and epistemological responsibility precludes the interpretation of AI as a full-fledged subject of scientific

cognition. The expediency of using the concept of a quasi-subject to denote the intermediate status of AI in science has been substantiated. Key scientific and practical problems related to the use of AI have been identified, including limited interpretability of results, diffusion of responsibility for scientific conclusions, and the presence of algorithmic biases.

Conclusions indicate that effective and scientifically valid application of AI is possible only under conditions of strict methodological control, preservation of human epistemological responsibility, and transparency of algorithmic procedures.

Prospects for further research are associated with the development of formalized models of human–algorithmic cognition, improvement of interpretable AI methods, and deeper analysis of transformations in scientific rationality under conditions of digitalization.

Key words: philosophy of science, epistemology, scientific rationality, digitalization of science, quasi-subjectivity, interpretability of knowledge, algorithmic bias, scientific responsibility, human-algorithm interaction.

Дата першого надходження статті до видання: 21.01.2026

Дата прийняття статті до друку після рецензування: 23.02.2026

Дата публікації (оприлюднення) статті: 19.05.2026