УДК 159.9:005.32:004.8

Галина СВІДЕНСЬКА
кандидат психологічних наук, доцент,
ДВНЗ «Донбаський державний педагогічний університет»,
https://orcid.org/0000-0003-1580-8881
e-mail: svidenskag@gmail.com
Наталія КАСЬЯНОВА
доктор економічних наук, професор,
Національний університет «Києво-Могилянська академія»
https://orcid.org/0000-0001-7729-2011
e-mail: nat_kas@ukr.net
Сергій КОВЕРГА
доктор економічних наук, професор,
ДВНЗ «Донбаський державний педагогічний університет»
https://orcid.org/0000-0003-4094-8165
e-mail: kovergaserg1970@gmail.com

ПСИХОЛОГІЧНА БЕЗПЕКА ЯК МЕХАНІЗМ ДОВІРИ ДО ШТУЧНОГО ІНТЕЛЕКТУ

Статтю присвячено обґрунтуванню психологічної безпеки взаємодії людини з ШІ як механізму формування довіри до ШІ-рішень. Показано, що психологічна прийнятність алгоритмічної рекомендації визначається поєднанням технічної точності, функціональної корисності й пояснюваності рішення з переживанням контролю, автономії судження, процедурної справедливості, можливості перевірки та визначеності відповідальності за наслідки управлінського вибору. Сформовано концептуальну структурну модель, у якій психологічна безпека виконує медіаційну функцію між оцінкою характеристик ШІ-системи та довірою до ШІ-рішень. Здійснено апробацію окремих змінних моделі на матеріалі взаємодії здобувачів освіти з генеративними ШІ-системами. У межах апробації розмежовано сприйняття компетентності ШІ, ШІ-тривожність, довіру до ШІ-відповідей, критичний контроль і попередні ознаки психологічної безпеки.

Ключові слова: психологічна безпека; довіра до ШІ; алгоритмічне лідерство; критичний контроль; автономія судження; управлінська відповідальність.

Halyna SVIDENSKA, Serhii KOVERHA
Donbas State Pedagogical University
Nataliia KASIANOVA
National University of Kyiv-Mohyla Academy

PSYCHOLOGICAL SAFETY AS A MECHANISM OF TRUST IN ARTIFICIAL INTELLIGENCE

The article substantiates psychological safety in human–AI interaction as a mediating mechanism in the formation of trust in AI decisions within the system of algorithmic leadership. It is shown that the psychological acceptability of an algorithmic recommendation is determined by the combination of technical accuracy, functional usefulness, and explainability of the decision with the experience of control, autonomy of judgment, procedural fairness, verifiability,

and clearly defined responsibility for the consequences of managerial choice. Trust in AI decisions is substantiated as a regulated psychological state in which reliance on an algorithmic conclusion is combined with the preservation of critical control.

Algorithmic leadership is revealed as a form of organizational influence in which a managerial decision is partially mediated by automated systems of data analysis, monitoring, evaluation, and recommendation generation. Within the structure of trust in AI decisions, cognitive, emotional, socio-psychological, and ethical prerequisites are differentiated, through which the characteristics of an algorithmic system acquire psychological significance for the subject of interaction. The mediating function of psychological safety in the transition from the evaluation of an AI system to the readiness to rely on its recommendation is substantiated.

A conceptual structural model for subsequent PLS-SEM testing is developed. In the model, explainability, transparency, perceived algorithmic fairness, and perceived AI competence are defined as positively directed predictors of PSAFE; AI anxiety is defined as a negative predictor; psychological safety is defined as a mediating variable; and leader emotional intelligence is defined as a moderator of the influence of predictors on psychological safety. A pilot descriptive approbation of selected model variables is carried out using data on students' interaction with generative AI systems. Within the approbation, perceived AI competence, AI anxiety, trust in AI responses, critical control, and preliminary indicators of psychological safety are differentiated. It is clarified that, at the pilot stage, psychological safety is interpreted through autonomy of judgment and control over AI influence, without equating these indicators with a fully developed latent scale.

Keywords: psychological safety, trust in AI, algorithmic leadership, critical control, autonomy of judgment, managerial responsibility.

https://doi.org/10.31891/PT-2026-4-17
Стаття надійшла до редакції / Received 24.09.2026
Прийнята до друку / Accepted 06.11.2026
Опубліковано / Published 26.11.2026

This is an Open Access article distributed under the terms of the Creative Commons CC-BY 4.0

© Галина СВІДЕНСЬКА, Наталія КАСЬЯНОВА, Сергій КОВЕРГА

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

Поширення систем штучного інтелекту змінює не лише способи роботи з інформацією, а й психологічні умови прийняття рішень. Алгоритмічні системи можуть використовуватися для аналізу даних, прогнозування, оцінювання результатів, розподілу завдань, контролю виконання та підготовки рекомендацій. У таких умовах людина взаємодіє вже не тільки з керівником або колегами, а й із системою, яка бере участь у формуванні управлінського висновку. Форму організаційного впливу, коли частина управлінського рішення формується або підтримується автоматизованою системою, розглядають як алгоритмічне лідерство. Це призводить до того, що працівник має оцінити не лише результат рішення, а й з'ясувати наскільки зрозумілим є спосіб його отримання, чи можна перевірити

рекомендацію, чи зберігається можливість власного судження та хто відповідає за наслідки прийнятого рішення.

Технічна точність комп'ютерних систем не гарантує їх психологічного прийняття. Людина може визнавати систему компетентною, але водночас не бути готовою покладатися на її рекомендації, якщо не розуміє підстав рішення або не відчуває можливості його перевірити. І навпаки, готовність довіряти може зростати тоді, коли взаємодія із системою сприймається як контрольована, передбачувана та така, що не позбавляє людину автономії. У зв'язку з цим виникає питання про психологічний механізм переходу від оцінки системи до довіри до її рішенням. Пропонується таким механізмом вважати психологічну безпеку взаємодії людини з алгоритмічною системою, а саме зі штучним інтелектом (ШІ). Психологічна безпека в цьому контексті означає не відсутність будь-яких сумнівів щодо технології, а можливість взаємодіяти з нею без втрати відчуття контролю, автономії та відповідальності за власне рішення. Саме тому довіра до рішення ШІ доцільно розглядати не як безумовне прийняття автоматизованого висновку, а як готовність покладатися на нього за збереження критичного контролю.

Аналіз останніх досліджень і публікацій

Проблема довіри до штучного інтелекту формується на перетині кількох напрямів психологічних і міждисциплінарних досліджень. У моделях прийняття технологій Ф. Девіса та В. Венкатеша і співавторів основну увагу приділено корисності технології, простоті її використання, соціальному впливу та умовам застосування [1; 2]. Ці підходи пояснюють, чому людина починає використовувати певну технологію, однак недостатньо описують психологічні умови довіри до конкретного алгоритмічного рішення.

Дослідження алгоритмічного управління зміщують увагу до проблеми контролю та організаційної влади. М. Лі, А. Розенблат, Л. Келлогг та інші автори показують, що алгоритмічні системи можуть змінювати досвід працівника через непрозорість критеріїв оцінювання, інформаційну асиметрію та обмеження суб'єктивного контролю [3-5]. У такій ситуації довіра до алгоритму пов'язана не тільки з його точністю, а й зі сприйняттям справедливості та можливістю контролювати наслідки його використання.

Важливим напрямом є дослідження довіри до автоматизації. Р. Лі та К. Сі розглядають довіру як основу відповідного рівня покладання на автоматизовану систему [6]. Водночас надмірна довіра також може бути проблемою, оскільки автоматизований висновок може

прийматися без належної перевірки. Дослідження людського використання автоматизації демонструють, що можливі як недостатнє використання системи, так і надмірне покладання на неї [7].

Окреме значення має пояснюваність алгоритмічних рішень. Дослідження Т. Міллера, А. Ададі та М. Берради, Д. Ганнінга та інших авторів показують, що зрозуміле пояснення може зменшувати невизначеність щодо того, як система формує рекомендацію [8-10]. Для психології важливим є не лише технічне існування пояснення, а й те, чи допомагає воно людині зрозуміти рішення та співвіднести його із власним судженням. Соціально-етичний аспект довіри пов'язаний зі справедливістю алгоритмічних процедур, використанням даних і розподілом влади. К. Кроуфорд і Ш. Зубофф звертають увагу на ризики дискримінації, непрозорого контролю та нерівномірного розподілу впливу в цифровому середовищі [11; 12]. У таких умовах довіра залежить від того, чи сприймається процедура як справедлива та чи має людина можливість реагувати на рішення системи.

Українські дослідження також розглядають психологічні та соціальні аспекти взаємодії людини зі штучним інтелектом. М. Яворський аналізує соціальну перцепцію штучного інтелекту, зокрема роль антропоморфізації та довіри до цифрових агентів [13]. Г. Андрощук досліджує рівень суспільної довіри до штучного інтелекту та його сприйняття в Україні [14]. Л. Головкова та К. Ханнуф пов'язують емоційний інтелект лідера із взаємодією людини та генеративного ШІ [15]. Дослідження цифрового та алгоритмічного лідерства також звертають увагу на зміну управлінського впливу під дією автоматизованих систем [16].

Отже, попередні дослідження дають змогу визначити окремі чинники довіри до ШІ, але потребують подальшого психологічного об'єднання. На нашу думку, таким об'єднувальним механізмом може бути психологічна безпека. Вона дає змогу пояснити, чому одна й та сама характеристика системи може мати різне значення для різних людей: важливим є не лише те, якою є система, а й те, як людина почувається під час взаємодії з нею.

Формулювання цілей статті

Мета статті – обґрунтувати психологічну безпеку як проміжний механізм формування довіри до рішень штучного інтелекту в умовах алгоритмічного лідерства, визначити основні психологічні чинники цього процесу, запропонувати структурну модель його подальшої емпіричної перевірки.

Виклад основного матеріалу

Алгоритмічне лідерство розглядається як форма організаційного впливу, за якої частина управлінського рішення підтримується автоматизованими системами аналізу, оцінювання, прогнозування та формування рекомендацій. На відміну від традиційної управлінської взаємодії, де рішення переважно пов'язується з особистою позицією керівника, тут до процесу включається система, правила роботи якої можуть бути не повністю зрозумілими працівникові [3]. Це змінює психологічні умови довіри. Людина оцінює не тільки те, чи здається рекомендація правильною, а й те, наскільки вона зрозуміла та передбачувана. Важливим стає питання: чи можу я перевірити це рішення, чи можу з ним не погодитися і чи зберігаю відповідальність за власний вибір? Тому довіру до рішення штучного інтелекту доцільно розуміти як готовність покладатися на його рекомендацію за умови збереження можливості критичної перевірки. Така довіра відрізняється як від повної недовіри, так і від некритичного прийняття автоматизованого висновку.

У психологічному плані можна виділити кілька пов'язаних складових такого ставлення:

– людина оцінює компетентність системи – чи здатна вона надати корисну, точну та послідовну рекомендацію;

– має значення зрозумілість та прозорість рішення, чому система надала саме таку рекомендацію;

– справедливість, бо якщо алгоритмічне рішення сприймається як упереджене або дискримінаційне, готовність довіряти йому зменшується;

– вплив емоційної реакції, тривожність щодо можливих помилок, втрати автономії або негативного впливу технології на власне мислення;

– збереження власного контролю щодо можливості використання рекомендацій системи, право перевірити її, уточнити або не прийняти.

Таким чином, психологічно обґрунтована довіра до ШІ не передбачає відмови від власного судження. Навпаки, її важливою умовою є можливість поєднати користування рекомендацією з критичним контролем. Якщо людина вважає систему компетентною, але не розуміє її рекомендацій і не може перевірити результат, позитивна оцінка компетентності може не перетворитися на довіру. Натомість зрозуміла процедура, можливість уточнення та збереження автономії можуть зробити взаємодію психологічно прийнятною. На наш погляд, між оцінкою характеристик системи ШІ та довірою її рішенням

існує певний психологічний процес, пов'язаний із переживанням безпечності взаємодії.

Психологічна безпека в запропонованій моделі охоплює три основні компоненти: відчуття контролю, автономію судження та можливість перевірки результату. До них додається відчуття передбачуваності взаємодії та захищеності від небажаного впливу алгоритмічної системи. З цієї точки зору, пояснюваність і прозорість мають значення не самі по собі. Справедливість важлива тому, що знижує очікування упередженого ставлення. Сприйняття компетентності підтримує впевненість у корисності рекомендації. Натомість тривожність може послаблювати відчуття безпеки.

Отже, основну логіку моделі можна представити у вигляді ланцюгу:

оцінка характеристик ШІ → психологічна безпека → довіра до рішення.

Результатом довіри може стати готовність приймати алгоритмічне управління як допустимий спосіб підтримки управлінського вибору.

На основі наведених положень запропоновано модель формування довіри до рішень штучного інтелекту (рис. 1), основна ідея якої полягає в тому, що довіра до рішень ШІ формується не безпосередньо під впливом технічних характеристик системи, а через психологічну безпеку взаємодії з нею.

Пояснюваність, прозорість, сприйняття справедливості та компетентності ШІ сприяють відчуттю контролю, передбачуваності й збереженню автономії користувача, тоді як тривожність щодо ШІ може послаблювати це відчуття. Сформована психологічна безпека, в свою чергою, підвищує готовність довіряти ШІ та приймати рекомендації як основу для професійного або управлінського рішення, не відмовляючись від критичного контролю.

Для подальшої емпіричної перевірки модель може бути представлена як система латентних психологічних змінних та перевірена методом PLS-SEM. На цій основі сформульовано структурні гіпотези концептуальної моделі (табл. 1).

Запропоновані структурні гіпотези конкретизують очікуваний перехід від оцінки характеристик ШІ-системи до психологічної безпеки, довіри до ШІ-рішень і організаційних наслідків алгоритмічного управління.

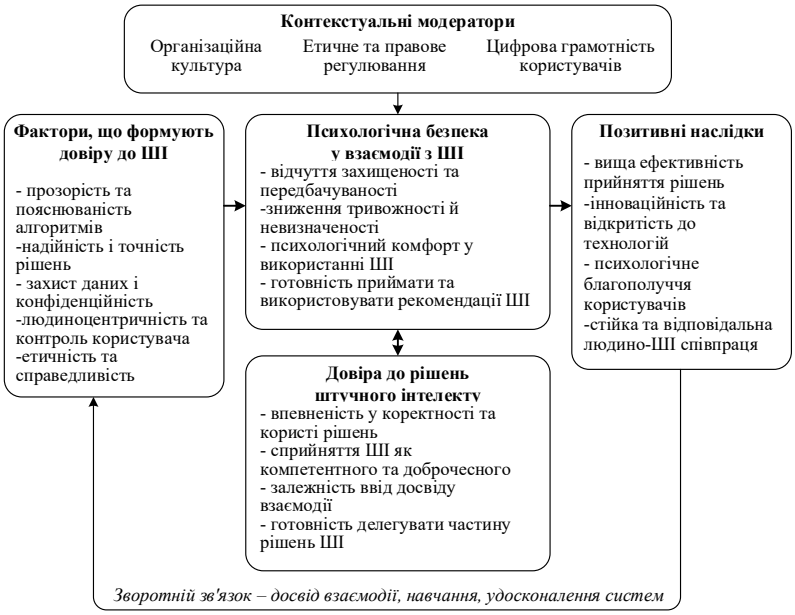


Рис. 1. Концептуальна модель психологічної безпеки та довіри до рішень штучного інтелекту

Джерело: розроблено авторами

На етапі емпіричної перевірки моделі було розглянуто психологічні передумови довіри до алгоритмічної рекомендації, які виникають до включення ШІ в організаційні процедури оцінювання, контролю та прийняття управлінських рішень. До складу цих передумов віднесено сприйняття компетентності ШІ, психологічну безпеку, ШІ-тривожність, готовність здобувачів освіти як майбутніх працівників покладатися на алгоритмічні рекомендації та критичний контроль автоматизованих висновків. Здобувачі освіти як майбутні працівники використовують генеративний ШІ під час пошуку, пояснення, редагування й оцінювання інформації, а також під час вибору способу виконання навчального завдання. У цьому досвіді поєднуються орієнтація на автоматизований висновок, перевірка його підстав, оцінка коректності відповіді, ризик перенесення відповідальності на систему та потреба зберегти автономне судження.

На цьому етапі особливу увагу приділено чотирьом групам показників: сприйняттю компетентності ШІ, тривожності щодо його використання, довірі до відповідей та ознакам психологічної безпеки.

Таблиця 1

Структурні гіпотези концептуальної моделі довіри до ШІ-рішень

Гіпотеза	Зв'язок	Напрям	Зміст
H1	Оцінка характеристик ШІ → психологічна безпека	+	Зрозумілість, прозорість, справедливість і сприйняття компетентності системи пов'язані з вищою психологічною безпекою взаємодії.
H2	Тривожність щодо ШІ → психологічна безпека	−	Тривожність щодо помилок, втрати автономії або негативного впливу системи послаблює психологічну безпеку.
H3	Психологічна безпека → довіра до ШІ-рішення	+	Відчуття контролю, автономії та можливості перевірки підтримує готовність покладатися на рекомендацію системи.
H4	Оцінка характеристик ШІ → психологічна безпека → довіра	опосеред-кований	Психологічна безпека опосередковує зв'язок між оцінкою системи та довірою до її рішення.
H5	Довіра до ШІ-рішення → прийняття алгоритмічного управління	+	Вища довіра пов'язана з більшою готовністю приймати алгоритмічну рекомендацію як допоміжний елемент управлінського рішення.
H6	Емоційний інтелект керівника × оцінка ШІ → психологічна безпека	модерація	Емоційний інтелект керівника може змінювати силу зв'язку між оцінкою системи та психологічною безпекою.

Джерело: складено авторами.

Емпіричну основу становили відповіді 262 здобувачів освіти різних курсів і спеціальностей одного закладу вищої освіти. Онлайн-анкета містила запитання про частоту та напрями використання генеративного ШІ, довіру до його відповідей, перевірку інформації та твердження за шкалою Лікерта від 1 до 5. На основі цього масиву відповідей здійснено описову апробацію сприйняття компетентності ШІ, ШІ-тривожності, психологічної безпеки, довіри та критичного контролю автоматизованих висновків. Частота використання генеративного ШІ характеризує практичну насиченість досвіду респондентів у роботі з алгоритмічно сформованими відповідями, тому її розподіл використано як вихідний описовий показник апробації (табл. 2).

Таблиця 2

Частота використання генеративного ШІ у навчальній діяльності

Частота використання	Кількість	%	Накопичений %
Декілька разів на день	75	28,6	28,6
Раз на день	33	12,6	41,2
Декілька разів на тиждень	122	46,6	87,8
Декілька разів на місяць	29	11,1	98,9
Майже ніколи	3	1,1	100,0
Разом	262	100,0	–

Примітка. Не рідше декількох разів на тиждень генеративний ШІ використовують 230 респондентів, або 87,8 % вибірки.

Джерело: складено авторами

Для переважної більшості учасників взаємодія з генеративним ШІ є регулярною. Це створює достатню основу для описового вивчення того, як користувачі оцінюють можливості системи та як вони поєднують довіру до неї з перевіркою отриманих відповідей. Характер цієї взаємодії уточнюється через завдання, для виконання яких респонденти звертаються до генеративного ШІ (табл. 3).

Таблиця 3

Напрями використання генеративного ШІ

Напрям використання ШІ	**Кількість**	**%**	**Психологічний зміст відповіді**
Пошук та узагальнення інформації	215	82,1	**звернення до ШІ як джерела інформаційної підтримки**
Пояснення складних тем	199	76,0	**використання ШІ для зниження когнітивної складності матеріалу**
Генерування ідей	159	60,7	**залучення ШІ до формування варіантів міркування зі збереженням власного висновку**
Написання / редагування текстів	130	49,6	**ризик заміщення самостійного формулювання автоматизованим текстом**
Переклад	90	34,4	**мовна підтримка навчального завдання**
Вирішення задач	81	30,9	**перевірка способу розв'язання або альтернативного ходу міркування**
Створення презентацій	73	27,9	**структурування навчального матеріалу**
Генерування коду	68	26,0	**спеціалізована процедурна підтримка**

Джерело: складено авторами

Найчастіше ШІ було використано для пошуку й узагальнення інформації (82,1%) та пояснення складних тем (76,0%). Це свідчить, що в навчальній діяльності ШІ виконує переважно функцію інформаційної та пояснювальної підтримки. Генерування ідей використовують 60,7% респондентів. У цьому випадку особливо важливим є питання збереження власного судження: система може запропонувати варіанти, але остаточний висновок залишається за людиною.

Для попередньої оцінки сприйняття компетентності було використано кілька показників (табл. 4). Середнє значення за цими показниками становить 3,90 (за шкалою Лікерта 1–5), що свідчить про достатньо позитивне сприйняття можливостей ШІ.

Таблиця 4

Описова статистика індикаторів сприйняття компетентності ШІ

Індикатор	Зміст індикатора	Середнє (M)	Стандартне відхилення (SD)	Психологічний зміст показника
IC1	ШІ допомагає знаходити різні точки зору	3,83	1,13	**оцінка здатності ШІ розширювати поле аналізу**
IC2	Став(-ла) краще формулювати запити завдяки ШІ	4,11	1,0	**оцінка керованості взаємодії через точніше формулювання запиту**
IC3	ШІ допомагає виявляти упередження у текстах	3,76	1,1	**оцінка аналітичної придатності ШІ**

Джерело: складено авторами

Найвище значення отримано за показником, пов’язаним із формулюванням запитів. Це важливо з психологічної точки зору – взаємодія із системою сприймається не як пасивне отримання готової відповіді, а як процес, у якому користувач може впливати на результат через спосіб постановки запиту. Отже, сприйняття компетентності ШІ поєднується із суб’єктивним відчуттям керованості взаємодії.

Отримані значення показника тривожності свідчать про помірну вираженість занепокоєння щодо використання ШІ. Найвищі показники стосуються можливого зменшення глибини самостійного аналізу та спрощення власної розумової роботи (табл. 5).

Таблиця 5

Описова статистика індикаторів ШІ-тривожності

Індикатор	Зміст твердження	Середнє (M)	Стандартне відхилення (SD)	Психологічний зміст показника
AIAN1	Менш глибоко аналізую тему самостійно	2,86	1,19	**зниження глибини самостійного аналізу під час використання ШІ**
AIAN2	Стаю «лінивішим» у мисленні	2,83	1,2	**суб'єктивне переживання когнітивного спрощення**
AIAN3	Складніше відрізняти факти від галюцинацій	2,50	1,1	**невпевненість у достовірності автоматизованого висновку**
AIAN4	Боюся погіршення критичного мислення	2,23	1,1	**тривожність щодо послаблення критичного мислення**
AIAN5	Іноді «сліпо» копіюю відповіді ШІ	2,57	1,2	**некритичне покладання на відповідь ШІ**

Джерело: складено авторами

Ці результати важливі для запропонованої моделі, оскільки показують подвійний характер взаємодії. ШІ може сприйматися як компетентний і корисний, але водночас викликати побоювання щодо збереження самостійності мислення. Саме ця суперечність робить психологічну безпеку важливою проміжною ланкою між оцінкою системи та довірою до неї.

Окремо було проаналізовано загальне ставлення до відповідей ШІ та частоту перевірки отриманої інформації (табл. 6). Найбільша частка респондентів (62,2%) обрала позицію часткової довіри, що поєднується з наміром перевіряти інформацію. Регулярну перевірку фактів зазначили 72,5% учасників. Цей результат є особливо важливим для запропонованої концепції довіри. Переважання часткової довіри не означає слабкого прийняття технології. Навпаки, воно може свідчити про форму обережного покладання, за якої людина використовує можливості ШІ, але не відмовляється від власної перевірки. Саме такий тип довіри відповідає розумінню психологічно безпечної взаємодії – покладання на рекомендацію без втрати критичного контролю.

Таблиця 6

Індикатори довіри до штучного інтелекту

Індикатор	Категорія відповіді	Кількість	%	Психологічний зміст показника
TRAI1	Повністю / здебільшого довіряють	72	27,5	**переважне покладання на відповідь ШІ**
TRAI2	Ставлення 50/50, намагаються перевіряти	163	62,2	**часткова довіра з перевіркою інформації**
TRAI3	Здебільшого / абсолютно не довіряють	27	10,3	**скептичне ставлення до відповідей ШІ**
Перевірка фактів	Завжди / часто перевіряють факти	190	72,5	**критичний контроль автоматизованого висновку**

Джерело: складено авторами

На наступному етапі для оцінки психологічної безпеки було використано два показники, які пов'язані з автономією судження та контролем над впливом ШІ (табл. 7). Обидва показники мають близькі середні значення. PS1 відображає збереження автономії судження, тоді як PSAFE2 стосується відчуття контролю над впливом ШІ на власне мислення.

Таблиця 7

Описова статистика індикаторів психологічної безпеки у взаємодії з ШІ

Індикатор	Зміст твердження	Середнє (M)	Стандартне відхилення (SD)	Психологічний зміст показника
PS1	Використовую ШІ для ідей, висновки формулюю самостійно	3,79	1,26	**збереження автономії судження під час використання ШІ**
PS2	Не боюся, що ШІ погіршить моє мислення	3,78	1,1	**відчуття контролю над впливом ШІ на власне мислення**

Джерело: складено авторами

Ці результати не дають підстав говорити про повністю сформовану латентну конструкцію психологічної безпеки, однак

показують доцільність її подальшого вимірювання. У подальшому психологічну безпеку доцільно розглядати ширше: через автономію судження, контрольованість взаємодії, можливість перевірки та уточнення рекомендації, а також захищеність від непрозорого впливу системи.

Отримані показники мають різний емпіричний характер. Їх зіставлення формує первинний описовий профіль співвідношення між сприйняттям компетентності штучного інтелекту, ШІ-тривожністю, довірою та психологічною безпекою (табл. 8).

Таблиця 8

Зведена описова статистика показників моделі

Показник	Значення	Попередня інтерпретація
Сприйняття компетентності ШІ	3,90	достатньо позитивна оцінка можливостей системи
Тривожність щодо ШІ	2,60	помірна вираженість занепокоєння
Автономія судження	3,79	переважання самостійного формулювання висновків
Контроль над впливом ШІ	3,78	достатньо виражене відчуття контролю
Часткова довіра	62,2%	обережне покладання на відповіді
Регулярна перевірка фактів	72,5%	збереження критичного контролю

Джерело: складено авторами

Результати дослідження показують поєднання позитивного ставлення до можливостей ШІ з його обережним використанням. Респонденти досить високо оцінюють компетентність системи, але значна частина з них не переходить до безумовної довіри. Натомість переважає часткова довіра, яка супроводжується перевіркою інформації. Це співвідношення відповідає основній гіпотезі дослідження – людина може визнавати систему компетентною і водночас зберігати власне судження. Отже, психологічно прийнятною формою довіри може бути не повне покладання на автоматизований результат, а контрольоване використання рекомендації.

Проведений аналіз дає підстави розглядати довіру до рішень штучного інтелекту як психологічно складніше явище, ніж просте прийняття технології. Важливим є не тільки те, наскільки компетентною людина вважає систему, а й те, чи зберігає вона відчуття контролю над процесом взаємодії. З одного боку, респонденти досить позитивно оцінюють можливості ШІ. З іншого – більшість не демонструє безумовної довіри до його відповідей і регулярно перевіряє отриману інформацію.

Це дозволяє припустити, що критичний контроль не обов'язково є протилежністю довіри. Навпаки, у взаємодії з алгоритмічними системами він може бути її умовою. Людина може довіряти системі настільки, наскільки готова використовувати її рекомендацію, залишаючи за собою можливість перевірки та остаточного судження. У цьому контексті психологічна безпека виконує роль механізму, який поєднує технічні характеристики системи з психологічною готовністю покладатися на її рішення. Пояснюваність, прозорість, справедливість і компетентність набувають значення тоді, коли вони допомагають людині зберегти відчуття автономії та контролю.

Водночас тривожність щодо ШІ може послаблювати цей процес. Якщо людина очікує помилок системи або побоюється втрати власної здатності мислити самостійно, навіть позитивна оцінка її можливостей не обов'язково приведе до довіри. Для алгоритмічного лідерства це має особливе значення. Коли алгоритмічна рекомендація використовується у кадровому оцінюванні, контролі або розподілі завдань, психологічна безпека залежить не тільки від властивостей самої системи. Важливими стають правила її використання, можливість заперечити результат, пояснення з боку керівника та чітке визначення відповідальності.

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

Психологічну безпеку взаємодії людини з ШІ обґрунтовано як проміжну психологічну ланку між оцінкою характеристик алгоритмічної системи та довірою до ШІ-рішень у системі алгоритмічного лідерства. Пояснюваність, прозорість, сприйняття справедливості та компетентності можуть підвищувати психологічну безпеку, тоді як тривожність щодо використання ШІ може її послаблювати. Психологічна безпека пов'язується із відчуттям контролю, автономією судження, можливістю перевірки та передбачуваністю взаємодії.

У межах дослідження розмежовано сприйняття компетентності ШІ, ШІ-тривожність, обережне покладання на ШІ-відповіді та попередні ознаки психологічної безпеки. Регулярне використання генеративного ШІ характеризує релевантний досвід взаємодії з алгоритмічно сформованими відповідями. У цьому емпіричному профілі поєднуються помірно високе сприйняття компетентності ШІ, помірна вираженість ШІ-тривожності та переважання часткової довіри до ШІ-відповідей за умови перевірки фактів.

Наступний етап дослідження передбачає розроблення повноцінної шкали психологічної безпеки взаємодії з ШІ та перевірку запропонованої моделі на організаційній вибірці. Особливий інтерес становить ситуація, у якій алгоритмічні рекомендації впливають на оцінювання результативності, розподіл завдань, контроль виконання або кадрові рішення. Перспективним напрямом є також перевірка того, чи є психологічна безпека саме механізмом переходу від позитивної оцінки системи до довіри, а не лише супутньою характеристикою ставлення до ШІ. Це дозволить точніше визначити психологічні умови прийняття алгоритмічних рішень у сучасному організаційному середовищі.

References

1. Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly.* 13 (3), 319-340. https://doi.org/10.2307/249008
2. Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*. 27 (3), 425-478. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3375136
3. Lee, M. K. (2018). Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management. *Big Data & Society*. 5 (1), 1-16. https://doi.org/10.1177/2053951718756684
4. Rosenblat, A. & Stark, L. (2016). Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers. *International Journal of Communication*. 10, 3758-3784. https://dx.doi.org/10.2139/ssrn.2686227
5. Kellogg, K. C., Valentine, M. A. & Christin, A. (2020). Algorithms at Work: The New Contested Terrain of Control. *Academy of Management Annals*. 2020. 14 (1), 366–410. https://doi.org/10.5465/annals.2018.0174
6. Lee, J. D. & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*. 46 (1), 50-80. https://doi.org/10.1518/hfes.46.1.50_30392
7. Parasuraman, R. & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*. 1997. 39 (2), 230-253. https://web.mit.edu/16.459/www/parasuraman.pdf
8. Miller, T. (2019). Explanation in Artificial Intelligence: Insights from the Social Sciences. *Artificial Intelligence*. 267, 1-38. https://doi.org/10.1016/j.artint.2018.07.007
9. Adadi, A. & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*. 6, 52138-52160. https://doi.org/10.1109/ACCESS.2018.2870052
10. Gunning, D. (2017). Explainable Artificial Intelligence (XAI). *Defense Advanced Research Projects Agency (DARPA),* 2017. https://doi.org/10.1145/3301275.3308446
11. Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.* New Haven : Yale University Press. https://www.essra.org.cn/upload/202105/The%20Atlas%20of%20AI%20Power,%20Politics,%20and%20the%20Planetary%20Costs%20of%20Artificial%20Intelligence.pdf
12. Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* New York : PublicAffairs.
13. Yavorskyi, M. V. (2025). Psykholohiia sotsialnoi pertseptsii shtuchnoho intelektu [Psychology of social perception of artificial intelligence]. *Psykholohichnyi zhurnal.* 15 (2), 112-124. https://doi.org/10.31499/2617-2100.15.2025.342104 [in Ukrainian]

14. Androshchuk, H. O. (2023). Riven doviry do shtuchnoho intelektu: analiz rezultativ hlobalnykh doslidzhen ta stan v Ukraini [Level of trust in artificial intelligence: analysis of global research results and the situation in Ukraine]. *Informatsiia i pravo*. 4 (47), 217-231. https://doi.org/10.37750/2616-6798.2023.4(47).291675 [in Ukrainian]

15. Holovkova, L. S. & Khannuf, K. Ye. (2025). Emotsiinyi intelekt lidera v umovakh heneratyvnoho ShI: "Human–AI Leadership" yak nova upravlinska kompetentnist [Emotional intelligence of a leader in the context of generative AI: "Human–AI Leadership" as a new management competency]. *Review of Transport Economics and Management*. 12 (28). https://doi.org/10.15802/rtem2025/351795 [in Ukrainian]

16. Vesolovska, M. K. (2025). Transformatsiia modelei liderstva pid vplyvom shtuchnoho intelektu: vid kharyzmatychnoho do alhorytmichnoho upravlinnia [Transformation of leadership models under the influence of artificial intelligence: from charismatic to algorithmic management]. *Visnyk Natsionalnoho universytetu «Lvivska politekhnika». Seriia «Problemy ekonomiky ta upravlinnia»*. 9 (2), 93-105. https://doi.org/10.23939/semi2025.02.093 [in Ukrainian]