Стаття та будь-який пов'язаний з нею опублікований матеріал поширюється за ліцензією Creative Commons Attribution License (CC BY 4.0).
The article and any related published material are licensed under the Creative Commons Attribution License (CC BY 4.0).

УДК 159.9.072:004.8:355/359
DOI: https://doi.org/10.17721/BPSY.2026.1(23).13

Олег ОРЛОВ, канд. психол. наук
ORCID ID: 0000-0003-0954-4402
e-mail: orlovoleh@gmail.com
Інститут спеціальної педагогіки і психології імені М. Ярмаченка НАПН України, Київ, Україна

ПОТЕНЦІАЛ ШТУЧНОГО ІНТЕЛЕКТУ У ПСИХОЛОГІЧНОМУ ПРОФАЙЛІНГУ ЗА СУБ'ЄКТИВНИМИ ОЦІНКАМИ ФАХІВЦІВ-ПСИХОЛОГІВ

Вступ. ***В умовах повномасштабної російсько-української війни та зростання потреби у швидкому психологічному оцінюванні персоналу сил безпеки і оборони особливої актуальності набуває пошук нових інструментів психодіагностики. Одним із перспективних напрямів є використання штучного інтелекту для психологічного профайлінгу на основі цифрових слідів користувача, зокрема історії його взаємодії з великими мовними моделями. Метою дослідження було на основі опитування фахівців у галузі психології оцінити якість і точність психологічних профілів особистості, згенерованих системами штучного інтелекту.***

Методи. ***Дослідження проводилось у форматі анонімного онлайн-опитування з елементами природного експерименту. Учасники самостійно запускали стандартизовані промпти у звичних для себе ШІ-системах, які мали доступ до попереднього контексту взаємодій, і оцінювали точність отриманих профілів. У фінальний аналіз було включено 29 фахівців у галузі психології. Оцінювалися три підходи: Big Five (OCEAN), психодинамічна модель Н. МакВільямс та аналітичний профайлінг. Для обробки даних використовувалися описова статистика, критерії Фрідмана і Вілкоксона, а також кореляційний аналіз Пірсона.***

Результати. ***Усі три підходи отримали подібні оцінки точності в межах помірно доброго рівня відповідності дійсності: Big Five – 7,1 бала, психодинамічний підхід – 6,9 бала, аналітичний профайлінг – 7,2 бала за 10-бальною шкалою. Загальна суб'єктивна оцінка якості ШІ-профайлінгу також склала 7,1 бала. Статистично значущих відмінностей між підходами не виявлено. Встановлено, що фахівці з вищою самооцінкою компетентності у профайлінгу критичніше оцінювали результати ШІ, тоді як більша різноманітність тематики попереднього спілкування з ШІ була пов'язана з вищими оцінками точності. Якісні відповіді респондентів додатково показали, що ефективність ШІ-профайлінгу суттєво залежить від обсягуй змістової різноманітності доступного контексту.***

Висновки. ***Сучасні системи штучного інтелекту здатні формувати психологічні профілі, які фахівці оцінюють як такі, що помірно добре відповідають дійсності. Отримані результати свідчать про перспективність застосування ШІ як допоміжного інструменту в психологічній діагностиці, зокрема у сфері психологічного відбору та супроводу персоналу сил безпеки і оборони. Водночас використання таких систем потребує критичного ставлення до результатів, подальшої об'єктивної валідації та врахування питань захисту персональних даних.***

Ключові слова: ***ШІ, великі мовні моделі, цифрові сліди, психодіагностика, профайлінг.***

Вступ

Повномасштабне вторгнення Російської Федерації в Україну та пов'язане з цим швидке і значне збільшення чисельності сил оборони України висунуло на перший план гостру потребу у швидкому та надійному психологічному оцінюванні великої кількості особового складу. Традиційні методи психодіагностики, попри значний накопичений досвід застосування, добру валідність та інші переваги, мають низку недоліків. Зокрема, їх застосування потребує значних часових і кадрових ресурсів, що є критичним обмеженням в умовах бойових дій та масштабної мобілізації. Водночас наявний у силах безпеки та оборони України психодіагностичний інструментарій потребує оновлення, що також є тривалим та ресурсоємним процесом (Гурковський та ін., 2025). Одним із перспективних потенційних напрямків досліджень для розв'язання зазначених нагальних проблем є застосування штучного інтелекту (ШІ).

Мета дослідження – на основі опитування фахівців у галузі психології оцінити якість та точність психологічних профілів особистості, згенерованих системами ШІ.

Огляд літератури. Розвиток великих мовних моделей (Large Language Models, LLM) відкриває принципово нові можливості для автоматизації та масштабування психологічної діагностики. Системи штучного (машинного) інтелекту (ШІ), такі як ChatGPT, Claude, Gemini та інші, здатні швидко аналізувати та узагальнювати великі обсяги різноманітної структурованої та неструктурованої інформації про користувача в різних форматах. Сучасні дослідження показують, що в різних типах цифрових слідів користувача – насамперед у текстовій та іншій мовленнєвій продукції, патернах активності в цифрових середовищах тощо, – можуть бути виявлені психологічно значущі індикатори, а на їх основі можливо робити припущення про індивідуально-психологічні особливості та поведінкові патерни користувача. (Kosinski et al., 2013; Youyou et al., 2015; Ocana Flores, & Luna, 2024). Крім того, аналіз текстів у цифрових середовищах показав можливість раннього виявлення депресивних станів та інших психологічних проблем. Це також свідчить про інформативність мовних і поведінкових цифрових даних (Guntuku et al., 2017; Eichstaedt et al., 2018).

Так, у дослідженні Peters, Cerf та Matz (2024) вивчалось, чи здатний чат-бот на основі GPT-4 при виконанні спеціально наданої інструкції виводити риси особистості користувача за моделлю "Великої п'ятірки". За наявної можливості розпитувати користувача про ті чи інші особливості особистості, чат-бот оцінював особистісні риси користувачів з помірною точністю, перевершуючи попередні підходи, що спиралися на аналіз статичного текстового контенту. При цьому точність висновків суттєво залежала від типу комунікативної взаємодії: вона була найвищою, коли чат-бот цілеспрямовано спонукав користувача розкривати особистісно значущу інформацію, і нижчою – за умов, орієнтованих на природність спілкування.

Подібні кількісні результати отримали Zhu, Jin та Coifman (2025), які оцінювали здатність моделей GPT-4o та GPT-4o mini виводити особистісні риси "Великої п'ятірки" за методикою BFI-10 із зразків користувацьких діалогів за умов zero-shot промптингу (прямої вказівки LLM вивести ці показники, без додаткових роз'яснень).

© Орлов Олег, 2026

Водночас дослідники встановили, що додавання проміжного кроку – генерації змодельованих відповідей користувача на окремі пункти опитувальника BFI-10 перед обчисленням підсумкових рис – підвищує точність і узгодженість результатів з відповідями самих користувачів на цей же опитувальник порівняно з прямим виведенням рис. Окремо було виявлено підвищену диференційну чутливість GPT-4o mini до пов'язаних із депресивною симптоматикою значень показників за шкалами нейротизму та сумлінності. Ці дані узгоджуються з більш ранніми спостереженнями щодо інформативності мовленнєвих цифрових слідів для виявлення депресивних станів (Guntuku et al., 2017; Eichstaedt et al., 2018), поширюючи їх на формат діалогової взаємодії з ШІ.

У цьому контексті історія попередніх взаємодій користувача з великими мовними моделями може розглядатися як окремий різновид цифрового сліду, потенційно придатний для психологічного аналізу. На відміну від статичних анкетних відповідей, такі взаємодії можуть містити більш природні прояви стилю мислення та комунікативних стратегій, прояви емоційних реакцій, дані щодо тематики інтересів користувача, пріоритетів тощо. Водночас, попри зростання інтересу до ШІ у психології, більшість наявних праць зосереджена на аналізі соціальних мереж або на загальних оглядах можливостей ШІ у ментальному здоров'ї. Питання про те, наскільки правдоподібними й прийнятними з погляду фахівців у галузі психології є психологічні профілі, згенеровані сучасними LLM, поки що залишається недостатньо опрацьованим (Ocana Flores, & Luna, 2024). Відповідно, заповнити цю прогалину і стало метою нашого дослідження.

Методи

Загальний дизайн дослідження. Дослідження проводилось у форматі анонімного онлайн-опитування з елементами природного експерименту. В експериментальній частині учасники самостійно запускали стандартизовані промпти (інструкції для ШІ) у своєму власному обліковому записі тієї ШІ-системи, якою вони зазвичай користуються, та оцінювали точність отриманих результатів. Такий дизайн забезпечував екологічну валідність: ШІ мав доступ до реального контексту взаємодій з учасником, а не до штучно сконструйованого профілю.

Учасники. До дослідження було залучено 31 фахівця в галузі психології, рекрутованих методом цільової вибірки. Опитування проводилось анонімно через Google Forms протягом одного місяця (у міру надходження нових відповідей). Результати двох учасників було виключено з аналізу, оскільки вони зазначили, що використовували ШІ у режимі без збереження історії переписки, внаслідок чого ШІ не мав доступу до попереднього контексту взаємодій – ключової умови дослідження. Таким чином, фінальна вибірка становила n=29. Вибірка охоплювала практичних психологів, клінічних психологів, психотерапевтів, науковців та викладачів. Стаж роботи в психологічній галузі сягав від 0,5 до 18 років (M = 8,7 років). Більшість учасників (57 %) зазначили ChatGPT як основний ШІ-інструмент, 32 % – Gemini, 7 % – Perplexity, решта – інші ШІ-системи.

Процедура. Кожен учасник після надання інформованої згоди та заповнення анкети виконував три окремих завдання в новому вікні чату ШІ, без додаткової інформації про себе.

Промпти були стандартизованими для всіх учасників і містили запит на формування психологічного профілю на основі наявного в системі контексту попередніх взаємодій за однією з трьох моделей:

1. Big Five (OCEAN) – оцінка п'яти особистісних факторів (відкритість досвіду, сумлінність, екстраверсія, доброзичливість, нейротизм) (Zell et. al., 2022);

2. Психодинамічна модель Н. МакВільямс – визначення рівня організації особистості, типу особистісної організації, домінантних захисних механізмів та особливостей формування міжособистісних стосунків (Lingiardi, & McWilliams, 2026);

3. Аналітичний профайлінг – синтетичний підхід, що включає опис стилю мислення, поведінкових патернів, мотиваційних драйверів, аналіз потенційних психологічних вразливостей та прогноз реакції на стрес.

Наприклад, промпт для оцінки особистості за моделлю OCEAN виглядав таким чином: "На основі всього, що тобі може бути відомо про мене з моїх попередніх взаємодій: стилю формулювання запитів, тематики звернень, стилю мовлення, фактів та будь-якого доступного контексту, спробуй оцінити мій психологічний профіль за моделлю Big Five (OCEAN). Спочатку коротко опиши, які психологічні індикатори ти можеш витягнути з доступного контексту (стиль мови, структура запитів, спосіб аргументації, тематика звернень, відомі факти про мене тощо). Потім оціни п'ять факторів Big Five (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism). Для кожного фактору: дай оцінку від 1 до 10; коротко поясни, на чому базується ця оцінка; зазнач рівень впевненості (низький / середній / високий). Якщо інформації недостатньо, сформулюй найбільш імовірні гіпотези, а не відмовляйся від відповіді. Зазнач рівень впевненості (низький / середній / високий)". В решті промптів змінювались назва підходу та деталізувались параметри, за якими слід було здійснити опис.

Після виконання кожного завдання учасники мали оцінити точність отриманого результату за шкалою від 0 до 10. Після виконання всіх трьох завдань учасники мали надати відповідь на запитання "Наскільки якісно, на вашу думку, ШІ може формувати психологічний профіль користувача?" за шкалою від 0 до 10.

Додаткові запитання передбачали збір інформації щодо стажу професійної діяльності в галузі психології (в роках), самооцінки компетентності в кожному з трьох психодіагностичних підходів (від "0. Не знайомий(а)" до "3. Добре знайомий(а) і використовую у практиці"), досвіду роботи у структурах, де застосовується психологічний профайлінг (запитання з множинним вибором від 0 до 5 категорій), частоти використання ШІ (вибір з 6 категорій від "0. Не користуюсь або майже не користуюсь" до "5. Кілька разів на день"), "стажу" використання ШІ (в місяцях), тематики запитів до ШІ (множинний вибір від 1 до 12 категорій), досвіду з різними системами ШІ (тип системи) та зазначення своєї основної системи ШІ. Також учасникам пропонувалось надати відкриту відповідь на запитання щодо потенціалу використання ШІ в психологічній діагностиці особистості.

Статистичний аналіз включав визначення дескриптивних статистик (середнє арифметичне, медіана, стандартне відхилення, мінімум та максимум значень), порівняльний аналіз середніх (за методом Фрідмена) для визначення відносної ефективності психологічних підходів до профайлінгу, порівняльний аналіз середніх (за методом Вілкоксона) для визначення відносної ефективності моделей ШІ, а також кореляційний аналіз (за методом Пірсона) для визначення параметрів, що потенційно пов'язані з ефективністю ШІ в психологічному діагностуванні.

Обчислення здійснювались в пакеті програмного забезпечення RStudio 2026.01.1 та R 4.5.3.

Результати

Точність профілів за трьома підходами. Усі три підходи отримали подібні оцінки точності в діапазоні "помірно-добре відповідає дійсності" (табл. 1). Жоден підхід

статистично значущо не перевершував інші (ні при застосуванні критерію Фрідмана, ні при попарному порівнянні за критерієм Вілкоксона), що може свідчити про наявність загального чинника сприйняття ефективності ШІ, незалежного від конкретної теоретичної рамки профайлінгу.

Таблиця 1

Середні оцінки точності за підходами (n=29)

Підхід	M	SD	m	Min	Max
Big Five (OCEAN)	7,1	1,9	7,0	3,0	10,0
Психодинамічний (МакВільямс)	6,9	2,1	7,0	2,0	10,0
Аналітичний профайлінг	7,2	1,9	8,0	2,0	10,0
Загальна суб'єктивна оцінка якості	7,1	1,7	8,0	2,0	10,0

Примітка: M – середнє арифметичне; m – медіана; SD – стандартне відхилення; Min – мінімум; Max – максимум.

Порівняльний аналіз оцінок точності ШІ між користувачами Gemini (n = 9) та ChatGPT (n = 16) за критерієм Вілкоксона виявив статистично значущу відмінність лише за оцінкою точності аналітичного профайлінгу (W = 30; p = 0,015): користувачі Gemini оцінили точність суттєво вище (M = 8,33; SD = 1,00), ніж користувачі ChatGPT (M = 6,56; SD = 2,06). За рештою показників, попри тенденцію користувачів давати Gemini на 1,0–1,8 балів вищі оцінки, ніж ChatGPT, статистично значущих відмінностей не виявлено: Big Five (W = 48,5; p = 0,187), психодинамічний підхід (W = 52; p = 0,264), загальна суб'єктивна оцінка (W = 49,5; p = 0,195). При цьому варто зазначити, що аналогічна тенденція спостерігається й у відмінностях між зазначеними групами користувачів за ступенем різноманіття досвіду використання ШІ (Gemini: M = 6,89; ChatGPT: M = 5,19; W = 44,5; p = 0,121). Тож інтерпретація відмінностей у відносній ефективності популярних ШІ моделей вимагає обережності.

Кореляційний аналіз. Статистично значущі результати аналізу взаємозв'язків між змінними дослідження наведено в табл. 2.

Між оцінками точності за всіма трьома підходами та загальною суб'єктивною оцінкою виявлено сильні позитивні кореляційні зв'язки (r = 0,61 – 0,77 при p < 0,001).

Таблиця 2

Результати кореляційного аналізу оцінок точності ШІ

Показник	Big Five	МакВільямс	Аналітичний профайлінг	Різноманітність тематики спілкування з ШІ	Загальна суб'єктивна оцінка ШІ
Big Five	-	0,61***	0,63**	0,48**	0,61***
МакВільямс	0,61***	-	0,77***	0,28	0,62***
Аналіт. профайлінг	0,63***	0,77***	-	0,29	0,76***
Різноманітність тематики спілкування з ШІ	0,48**	0,28	0,29	-	0,29
Суб'єктивна оцінка компетентності фахівця в профайлінгу	-0,43*	-0,28	-0,33	-0,30	-0,14
Досвід роботи на посадах, що передбачають застосування профайлінгу	-0,31	-0,33	-0,41*	-0,09	-0,27

Примітка: * – p < 0,05; ** – p < 0,01; *** – p < 0,001.

Виявлено окремі статистично значущі кореляти оцінки точності ШІ. По-перше, суб'єктивна оцінка знання психологічного профайлінгу негативно корелює з оцінкою ефективності ШІ у застосуванні підходу Big Five (r = -0,43; p < 0,05): фахівці, які вище оцінюють власну обізнаність з методологією профайлінгу, оцінюють результати ШІ більш критично. По-друге, досвід роботи на посадах, що тією чи іншою мірою передбачають застосування профайлінгу (як-от військові підрозділи, розвідувальні структури, правоохоронні органи, управління людськими ресурсами), негативно корелює з оцінкою ефективності ШІ у застосуванні аналітичного профайлінгу (r = -0,41; p < 0,05). По-третє, різноманіття тематики спілкування з ШІ (наявність в ШІ різноманітного контексту) позитивно корелює з оцінкою підходу Big Five (r = 0,48, p < 0,05).

Водночас варто зазначити, що статистично значущих кореляційних зв'язків між оцінками точності ШІ та тривалістю досвіду роботи в психологічній галузі, а також тривалістю досвіду застосування ШІ виявлено не було.

Якісні дані. Аналіз відповідей учасників на відкриті запитання (n=8 коментарів) дозволяє більш глибоко інтерпретувати кількісні результати дослідження. По-перше, учасники одностайно наголошують на залежності точності ШІ-профайлінгу від наявного контексту: якість результату принципово визначається обсягом і тематичним різноманіттям попередніх взаємодій із системою ("у випадку малоінтенсивного користування цінність діагностики буде дуже недостовірною"; "у 90 % випадків я роблю запити під час навчання – у цій сфері ШІ чудово описує мої риси, в інших – далекий від правди"). По-друге, учасники виокремлюють три різні моделі доцільного застосування ШІ в психологічній діагностиці: 1) як інструменту первинного скринінгу великих масивів даних; 2) як засобу структурування та узагальнення інформації перед клінічним інтерв'ю; 3) і як вузькоспеціалізованого діагностичного модуля з обмеженою предметною областю. По-третє, учасники підкреслюють потенційні проблеми, пов'язані із застосуванням ШІ в психодіагностиці, зокрема пов'язані зі схильністю ШІ до галюцинацій, маніпуляцій та проблемою передавання персональних даних третім особам без можливості контролю (при використанні хмарних сервісів).

Дискусія і висновки

Отримані результати свідчать про те, що сучасні системи ШІ здатні формувати психологічні профілі користувачів, які фахівці у галузі психології оцінюють як такі, що помірно-добре відповідають дійсності (середні оцінки в межах 6,9–7,2 балів за 10-бальною шкалою). Це узгоджується з попередніми дослідженнями, які демонстрували психодіагностичну інформативність цифрових слідів (Kosinski et al., 2013; Youyou et al., 2015), і дозволяє поширювати ці

висновки на новий тип цифрового сліду – історію взаємодії користувача з великими мовними моделями.

Варто зазначити, що користувачі схильні до глибшого саморозкриття у взаємодії з розмовними LLM порівняно з традиційними веб-опитувальниками чи анкетами. Це підтверджується систематичними оглядами літератури (як-от Papneja, & Yadav, 2025). Відповідно, історія взаємодії з LLM потенційно містить більш природні прояви стилю мислення, емоційних реакцій та комунікативних стратегій користувача, ніж статичні анкетні відповіді, і потенційно становить високу психодіагностичну цінність.

Попри потенційну користь, із цим пов'язана і низка ризиків. Staab et al. (2024) показали, що LLM здатні з точністю до 85 % виводити особисту інформацію про користувача (місце проживання, дохід, стать тощо) з неструктурованого тексту, причому анонімізація тексту практично цьому не перешкоджає. В поєднанні з психологічною інформацією, це створює передумови для накопичення компаніями-провайдерами високоточних особистих профілів, що можуть використовуватись для впливу на користувачів. Наприклад, Matz et al. (2017) в дослідженні на вибірці понад 3 млн осіб показали, що реклама, підлаштована під психологічний профіль, підвищує клікабельність до 40 % і кількість покупок до 50 %. Гіпотетично такий вплив може становити суттєву загрозу як для окремих осіб, так і для держав.

Відсутність статистично значущих відмінностей між трьома концептуально різними підходами до профайлінгу (Big Five, психодинамічна модель МакВільямс, аналітичний профайлінг), у поєднанні із сильними позитивними кореляціями між оцінками їх точності (r = 0,61–0,77), дозволяє припустити існування загального чинника сприйняття ефективності ШІ, який є відносно незалежним від конкретної теоретичної рамки. Інакше кажучи, учасники схильні оцінювати "ШІ як такий" – його здатність робити осмислені висновки про користувача, – а не специфічні переваги того чи іншого психодіагностичного підходу. Це може пояснюватися як реальною універсальністю мовних моделей у роботі з різними діагностичними підходами, так і ефектом Барнума (правдоподібні і достатньо широкі описи особистості можуть сприйматись як точні, незалежно від методологічного підґрунтя та відповідності реальності).

У ситуації з великими мовними моделями потенційний вплив ефекту Барнума підтверджується емпірично. В дослідженні, проведеному Pataranutaporn, Lee, Amores та Maes (2026), продемонстровано схильність учасників оцінювати позитивні за валентністю оцінки щодо власної фінансової грамотності як валідні та суб'єктивно значущі, попри те що ці прогнози були фіктивними й заздалегідь згенерованими. Ефект Барнума у взаємодії з LLM при цьому був зіставний з виміряним у тому ж дослідженні аналогічним ефектом при оцінці астрологічних прогнозів та "розпливчастих" психологічних формулювань. Відповідно, можна припустити, що ці висновки можна поширити й на згенеровані LLM особистісні профілі: позитивна, персоналізована за формою оцінка особистості, згенерована LLM, може сприйматися користувачем як унікально точна незалежно від її фактичної валідності.

Розрізнення двох вищеозначених пояснень відносної однорідності оцінок потребує подальших емпіричних досліджень із залученням об'єктивних критеріїв валідації (наприклад, зіставлення ШІ-профілів з результатами валідних стандартизованих психодіагностичних опитувальників).

Виявлені негативні кореляції між оцінкою ефективності ШІ та суб'єктивною компетентністю у профайлінгу (r = -0,43 для Big Five) і досвідом роботи на посадах, що передбачають профайлінг (r = -0,41 для аналітичного підходу), свідчать про те, що фахівці з релевантною підготовкою є помітно критичнішими в оцінці ШІ-профілів. Імовірне пояснення полягає в тому, що такі фахівці мають внутрішній еталон якісного профілю, з яким можуть зіставити результат ШІ, і тому точніше ідентифікують його поверховість, надмірну загальність та методологічні недоліки. Показово, що ані загальний стаж роботи в психологічній галузі, ані "стаж" використання ШІ значущого зв'язку з оцінками не показали. Це узгоджується з відомим принципом доменної специфічності експертизи: критичність до ШІ-профайлінгу формується не професійним досвідом узагалі, а саме досвідом виконання тієї конкретної задачі, яку ШІ намагається відтворити. Загальні показники стажу об'єднують надто різнорідну діяльність (психотерапію, викладання, наукову роботу), більша частина якої не передбачає систематичного складання психологічних профілів і тому не формує релевантного еталону для критичного зіставлення. Позаяк критичність фахівців з досвідом у психологічному профайлінгу щодо результативності ШІ може пояснюватись і іншими чинниками.

Питання фахової оцінки контенту, згенерованого LLM, частково висвітлене в дослідженні Föyen et. al. (2025). Ліцензовані психологи та психотерапевти в умовах сліпого вибору (за відсутньої інформації про авторство) оцінювали якість та емпатійність психологічних порад, частина яких була згенерована ШІ, а частина – написана фахівцями-людьми. Виявилося, що поради, згенеровані ШІ, отримали вищі оцінки емоційної та мотиваційної емпатійності, тоді як за якістю та когнітивною емпатійністю суттєвих відмінностей не зафіксовано. Водночас фахівці не змогли достовірно розрізнити авторство текстів. При цьому тексти, що сприймались учасниками як згенеровані експертами, отримували суттєво вищу загальну оцінку, незалежно від фактичного авторства. Це свідчить, що проблема прийнятності згенерованих ШІ психологічних профілів для фахової спільноти стосується не тільки власне якості контенту, а й когнітивних упереджень, пов'язаних з атрибуцією джерела. Тож питання критичності фахівців до результатів ШІ-профайлінгу потребує додаткового емпіричного дослідження.

Так чи інакше, впровадження ШІ-інструментів у діагностичну практику має супроводжуватися підготовкою користувачів саме до критичного оцінювання результатів застосування таких ШІ-систем задля зменшення негативного впливу ефекту Барнума.

Позитивний кореляційний зв'язок між різноманітністю тематики спілкування з ШІ та оцінкою точності профілю Big Five (r = 0,48), а також одностайні коментарі учасників у відкритій частині опитування, наголошують на критичній залежності якості ШІ-профайлінгу передусім від тематичного різноманіття доступного системі контексту, що є очікуваним з технічного погляду. Водночас конкретні обсяги та міри різноманітності контексту потребують детального вивчення.

Проведене дослідження показало, що сучасні системи ШІ здатні на основі наявного в них контексту спілкування формувати психологічні профілі користувачів, які фахівці у галузі психології оцінюють як такі, що помірно-добре відповідають дійсності. При цьому такі оцінки є відносно стабільними незалежно від обраного теоретико-методологічного підходу до профайлінгу (як-от Big Five, психодинамічна модель МакВільямс, аналітичний профайлінг).

Ключовими змінними, що пов'язані зі сприйняттям точності ШІ-профілів, виявились тематичне різноманіття попередніх взаємодій користувача із системою ШІ (чим

більш різноманітний контекст має система, тим вище користувачі оцінювали результат) та предметно-специфічний досвід у сфері профайлінгу (користувачі з більшим досвідом у профайлінгу були більш критичними до результатів ШІ).

Серед обмежень дослідження слід відзначити відносно невеликий обсяг вибірки (n=29), що знижує статистичну потужність аналізу; суб'єктивний характер оцінювання точності профілів самими учасниками без паралельної об'єктивної валідації за допомогою стандартизованих психодіагностичних методик; неконтрольованість обсягу та якості контексту, доступного ШІ для кожного учасника; а також переважно україномовний характер вибірки, що обмежує можливості узагальнення результатів на взаємодію з ШІ іншими мовами. Крім того, дизайн дослідження не дає змоги чітко розмежувати реальну валідність ШІ-профілів від ефекту Барнума.

Перспективи подальших досліджень охоплюють розширення вибірки та залучення об'єктивних критеріїв валідації через паралельне застосування стандартизованих психодіагностичних методик, контрольоване варіювання обсягу й тематичного різноманіття контексту для встановлення мінімально достатніх параметрів якісного профайлінгу, порівняльне оцінювання різних мовних моделей в умовах однакового контексту, дослідження ефективності ШІ-профайлінгу на специфічних типах цифрових слідів (зокрема експорту переписки з месенджерів), а також розробку та апробацію локальних, безпечних з погляду захисту персональних даних рішень для застосування у сферах, де використання комерційних хмарних сервісів є неприйнятним – насамперед у силах безпеки та оборони.

Відкрита наука. Дані, наведені в цій статті, можна отримати за запитом у автора. Запити щодо доступу до цих наборів даних слід надсилати на адресу orlovoleh@gmail.com

Джерела фінансування. Автор заявляє про відсутність зовнішнього фінансування для проведення цього дослідження та підготовки рукопису до публікації.

Список використаних джерел

Гурковський, В., Романенко, Є., Недбалек, К., Семененко, Л., Дужий, Р., & Семененко, А. (2025). Військова педагогіка та психологія в умовах російсько-української війни: проблеми, рішення і перспективи розвитку. *Міжнародний науковий журнал "Military Science",* 3(3), 66–82. https://doi.org/10.62524/msj.2025.3.3.5

Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoţiuc-Pietro, D., Asch, D. A., & Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences,* 115(44), 11203–11208. https://doi.org/10.1073/pnas.1802331115

Föyen, L. F., Zapel, E., Lekander, M., Hedman-Lagerlöf, E., & Lindsäter, E. (2025). Artificial intelligence vs. human expert: Licensed mental health clinicians' blinded evaluation of AI-generated and expert psychological advice on quality, empathy, and perceived authorship. *Internet Interventions.* https://doi.org/10.1016/j.invent.2025.100841

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: An integrative review. *Current Opinion in Behavioral Sciences,* 18, 43–49. https://doi.org/10.1016/j.cobeha.2017.07.005

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences,* 110(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Lingiardi, V., & McWilliams, N. (Eds.). (2026). *Psychodynamic Diagnostic Manual (3rd ed.; PDM-3).* Guilford Press.

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences,* 114(48), 12714–12719. https://doi.org/10.1073/pnas.1710966114

Ocana Flores, H. I., & Luna, A. (2024). AI for psychological profiles: Advances, challenges, and future directions. *Ciencia Latina Revista Científica Multidisciplinar,* 8(3), 10592–10609. https://doi.org/10.37811/cl_rcm.v8i3.12221

Papneja, H., & Yadav, N. (2025). Self-disclosure to conversational AI: A literature review, emergent framework, and directions for future research. *Personal and Ubiquitous Computing,* 29, 119–151. https://doi.org/10.1007/s00779-024-01823-7

Pataranutaporn, P., Lee, E., Amores, J., & Maes, P. (2026). Personal validation effect in LLMs: Positive AI responses bias perceptions of validity, reliability, personalization, and usefulness of fictitious predictions. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). ACM. https://doi.org/10.1145/3772318.3791851

Peters, H., Cerf, M., & Matz, S. C. (2024). Large language models can infer personality from free-form user interactions (arXiv:2405.13052). *arXiv.* https://doi.org/10.48550/arXiv.2405.13052

Staab, R., Vero, M., Balunović, M., & Vechev, M. (2024). Beyond memorization: Violating privacy via inference with large language models. In *Proceedings of the Twelfth International Conference on Learning Representations (ICLR 2024).* https://doi.org/10.48550/arXiv.2310.07298

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences,* 112(4), 1036–1040. https://doi.org/10.1073/pnas.1418680112

Zell, E., & Lesick, T. L. (2022). Big five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality,* 90(4), 559–573. https://doi.org/10.1111/jopy.12683

Zhu, J., Jin, R., & Coifman, K. G. (2025). Investigating large language models in inferring personality traits from user conversations (arXiv:2501.07532). *arXiv.* https://doi.org/10.48550/arXiv.2501.07532

References

Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoţiuc-Pietro, D., Asch, D. A., & Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences,* 115(44), 11203–11208. https://doi.org/10.1073/pnas.1802331115

Föyen, L. F., Zapel, E., Lekander, M., Hedman-Lagerlöf, E., & Lindsäter, E. (2025). Artificial intelligence vs. human expert: Licensed mental health clinicians' blinded evaluation of AI-generated and expert psychological advice on quality, empathy, and perceived authorship. *Internet Interventions.* https://doi.org/10.1016/j.invent.2025.100841

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: An integrative review. *Current Opinion in Behavioral Sciences,* 18, 43–49. https://doi.org/10.1016/j.cobeha.2017.07.005

Hurkovskyi, V., Romanenko, Ye., Nedbalek, K., Semenenko, L., Duzhyi, R., & Semenenko, A. (2025). Military pedagogy and psychology in the context of the Russian-Ukrainian war: Problems, solutions, and prospects for development. *International Scientific Journal "Military Science",* 3(3), 66–82. https://doi.org/10.62524/msj.2025.3.3.5

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences,* 110(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Lingiardi, V., & McWilliams, N. (Eds.). (2026). *Psychodynamic Diagnostic Manual (3rd ed.; PDM-3).* Guilford Press. ISBN 9781462558711.

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences,* 114(48), 12714–12719. https://doi.org/10.1073/pnas.1710966114

Ocana Flores, H. I., & Luna, A. (2024). AI for psychological profiles: Advances, challenges, and future directions. *Ciencia Latina Revista Científica Multidisciplinar,* 8(3), 10592–10609. https://doi.org/10.37811/cl_rcm.v8i3.12221

Papneja, H., & Yadav, N. (2025). Self-disclosure to conversational AI: A literature review, emergent framework, and directions for future research. *Personal and Ubiquitous Computing,* 29, 119–151. https://doi.org/10.1007/s00779-024-01823-7

Pataranutaporn, P., Lee, E., Amores, J., & Maes, P. (2026). Personal validation effect in LLMs: Positive AI responses bias perceptions of validity, reliability, personalization, and usefulness of fictitious predictions. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). ACM. https://doi.org/10.1145/3772318.3791851

Peters, H., Cerf, M., & Matz, S. C. (2024). Large language models can infer personality from free-form user interactions (arXiv:2405.13052). *arXiv.* https://doi.org/10.48550/arXiv.2405.13052

Staab, R., Vero, M., Balunović, M., & Vechev, M. (2024). Beyond memorization: Violating privacy via inference with large language models. In *Proceedings of the Twelfth International Conference on Learning Representations (ICLR 2024).* https://doi.org/10.48550/arXiv.2310.07298

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences,* 112(4), 1036–1040. https://doi.org/10.1073/pnas.1418680112

Zell, E., & Lesick, T. L. (2022). Big five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality,* 90(4), 559–573. https://doi.org/10.1111/jopy.12683

Zhu, J., Jin, R., & Coifman, K. G. (2025). Investigating large language models in inferring personality traits from user conversations (arXiv:2501.07532). *arXiv.* https://doi.org/10.48550/arXiv.2501.07532

Отримано редакцією журналу / Received: 11.05.26
Схвалено до друку / Accepted: 15.07.26
Опубліковано / Published: 16.07.26

Oleh ORLOV, PhD (Psychol.)
ORCID ID: 0000-0003-0954-4402
e-mail: orlovoleh@gmail.com
Mykola Yarmachenko Institute of Special Education and Psychology
of the National Academy of Educational Sciences of Ukraine, Kyiv, Ukraine

THE POTENTIAL OF ARTIFICIAL INTELLIGENCE IN PSYCHOLOGICAL PROFILING ACCORDING TO THE SUBJECTIVE ASSESSMENTS OF PSYCHOLOGISTS

B a c k g r o u n d . ***In the context of the full-scale Russian-Ukrainian war and the growing need for rapid psychological assessment of security and defense personnel, the search for new psychodiagnostic tools has become especially relevant. One promising area is the use of artificial intelligence (AI) for psychological profiling based on users' digital traces, particularly the history of their interactions with large language models. The aim of the study was to assess, based on a survey of psychology specialists, the quality and accuracy of personality profiles generated by AI systems.***

M e t h o d s . ***The study was conducted as an anonymous online survey with elements of a natural experiment. Participants independently ran standardized prompts in the AI systems they normally use, which had access to the previous interaction context, and then evaluated the accuracy of the generated profiles. The final analysis included 29 psychology specialists. Three approaches were assessed: the Big Five (OCEAN), Nancy McWilliams' psychodynamic model, and analytical profiling. Data were processed using descriptive statistics, the Friedman and Wilcoxon tests, and Pearson correlation analysis.***

R e s u l t s . ***All three approaches received similar accuracy ratings within a moderately good range of correspondence to reality: Big Five – 7,1 points, psychodynamic approach – 6,9 points, and analytical profiling – 7,2 points on a 10-point scale. The overall subjective rating of AI profiling quality was also 7,1 points. No statistically significant differences were found between the approaches. It was established that specialists with higher self-rated competence in profiling evaluated AI results more critically, whereas greater thematic diversity in prior interactions with AI was associated with higher accuracy ratings. Qualitative responses from participants additionally showed that the effectiveness of AI profiling depends substantially on the amount and substantive diversity of the available context.***

C o n c l u s i o n s . ***Modern artificial intelligence systems are capable of generating psychological profiles that specialists evaluate as moderately well corresponding to reality. The results indicate the promise of using AI as a supportive tool in psychological assessment, particularly in the psychological selection and support of security and defense personnel. At the same time, the use of such systems requires a critical attitude toward the results, further objective validation, and careful consideration of personal data protection issues.***

K e y w o r d s ***: AI, large language models, digital traces, psychodiagnostics, profiling.***

Декларація про генеративний штучний інтелект та технології, що використовують штучний інтелект. Під час підготовки цієї статті автор використовував ChatGPT для пошуку наукових джерел. Після використання цього інструмента автор перевірив та проаналізував запропоновані джерела.

Declaration of Generative AI and AI-Assisted Technologies. During the preparation of this work, the author used ChatGPT to search for scientific literature. Following the use of this tool, the author reviewed, verified, and analyzed the suggested sources.

Етична декларація. Дослідження схвалено Інститутом спеціальної педагогіки і психології імені М. Ярмаченка НАПН України, (протокол № 2 від 05.02.2026). Усі учасники дослідження надали письмову інформовану згоду на участь та публікацію отриманих даних. Автори гарантують повну анонімність та конфіденційність персональної інформації.

Ethical Statement. The study was approved by the Mykola Yarmachenko Institute of Special Education and Psychology of the National Academy of Educational Sciences of Ukraine (Protocol No. 2 dated February 5, 2026). All study participants provided written informed consent for their participation and the publication of the obtained data. The authors guarantee full anonymity and confidentiality of personal information.

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.