



Integration of AI-Assisted Mental Well-Being Metrics (MWI and PES) into Corporate Practice: Methodological Development and Results of a Pilot Study

Oleg Chaban

Educational and Research Institute of Mental Health,
Bogomolets National Medical University

Oleg Shevchenko

OSAI Workspace, SM Pulse Platform

Background. With growing demand for scalable mental wellbeing solutions in corporate settings, there is a need for objective metrics that monitor psychological state without privacy intrusion. Traditional questionnaires have limitations due to social desirability bias and low compliance.

Objective. To develop and validate a system of AI-assisted metrics for daily mental wellbeing monitoring (Mental Wellbeing Index, MWI) and platform engagement (Platform Engagement Score, PES) in corporate context, and to investigate the relationship between these indicators.

Methods. A pilot study was conducted over 6 weeks (November 10 - December 22, 2025) with 12 employees of a Ukrainian company. MWI was calculated daily based on AI analysis of dialogues across 5 dimensions (energy, clarity, balance, control, connection). PES was assessed weekly across 5 components (frequency, quality, completion, initiative, response speed). Pearson correlation analysis and quadrant matrix were used for user stratification.

Results. Mean MWI was 4.93 ± 0.52 (range 0-10), mean PES was 7.78 ± 0.89 (range 0-10). A paradoxical pattern emerged: 58% of users with lower MWI (< 5.5) showed high PES (> 7.0), indicating active help-seeking behavior. The strongest correlation was found between dialogue quality and MWI ($r=0.35$, $p<0.05$), 3 times stronger than frequency impact ($r=0.12$). Highly engaged users ($PES \geq 7$) had MWI 0.41 points higher (+9%) than low-engagement users.

Conclusions. The MWI/PES system shows potential for early burnout detection and wellbeing program effectiveness monitoring. Key factor is dialogue quality rather than mere quantity. The "help-seeking" pattern (low MWI + high PES) suggests the platform effectively engages users who need support most. Further longitudinal studies with larger samples are needed to validate long-term effectiveness.

Keywords: digital mental health, artificial intelligence, workplace wellbeing, cognitive-behavioral therapy, micro-interventions, burnout prevention, conversational AI

Вступ

Проблематика корпоративного ментального здоров'я

Професійне вигорання визнано Всесвітньою організацією охорони здоров'я як "синдром, що виникає внаслідок хронічного стресу на робочому місці" [1]. За даними Gallup (2022), 44%

працівників відчувають значний стрес щодня, що призводить до економічних втрат понад \$322 млрд щорічно через абсентеїзм та зниження продуктивності [2]. В Україні ситуація ускладнюється наслідками пандемії COVID-19 та триваючої війни, що створює додаткове навантаження на психічне здоров'я населення [3,4].

Традиційні підходи до корпоративної підтримки ментального здоров'я включають програми Employee Assistance Program (EAP), консультації психологів та періодичні опитування благополуччя. Однак вони мають суттєві обмеження (рис.1).

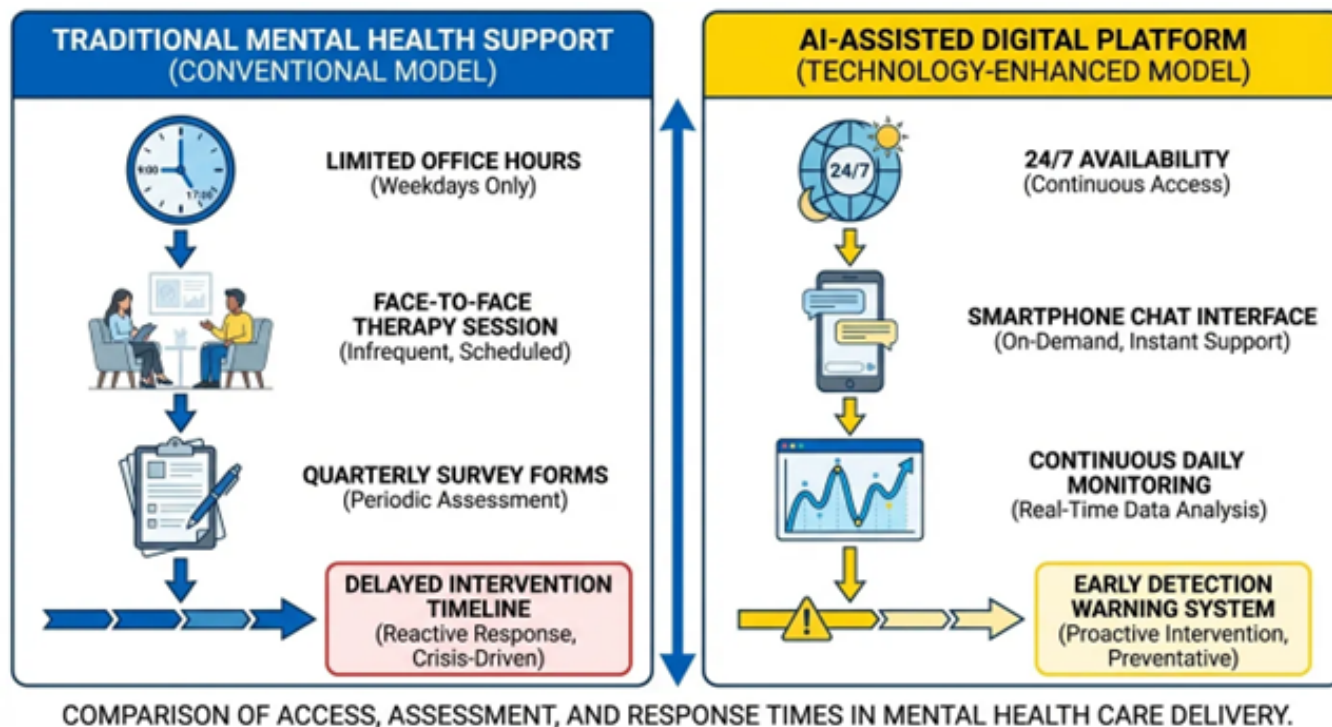


Figure 1. Обмежена доступність — психологічна допомога доступна лише в робочий час **Стигматизація** — співробітники уникають звернення через побоювання щодо конфіденційності [5] **Реактивність** — втручання відбувається після настання кризи, а не на етапі профілактики **Низька комплаєнтність** — періодичні опитування часто ігноруються або заповнюються формально [6] **Відсутність об'єктивних метрик** — складно оцінити ефективність wellbeing-програм через суб'єктивність самозвітів

Цифрові рішення у ментальному здоров'ї

Останні 5 років характеризуються експоненційним зростанням ринку digital mental health — з \$4.2 млрд у 2022 році до прогнозованих \$17 млрд у 2030 [7]. Пандемія COVID-19 стала каталізатором масового впровадження телемедицини та AI-асистованих інструментів підтримки [8].

Існуючі цифрові рішення можна класифікувати на три категорії:

1. **Структуровані чат-боти (Woebot, Wysa, Youper).** Базуються на протоколах когнітивно-поведінкової терапії (CBT) з фіксованими сценаріями діалогів. Meta-аналіз 2021 року (N=22 дослідження) показав помірний ефект на зниження депресії (d=0.41) та тривоги (d=0.38) [9]. Основне обмеження — низька гнучкість та швидке "вичерпування" змісту.
2. **Conversational AI нового покоління (GPT-4, Claude).** Великі мовні моделі (LLM) здатні генерувати природні відповіді та адаптуватися під контекст користувача.

Дослідження 2023 року показало, що GPT-4 може проводити базову емпатичну бесіду на рівні, порівнянному з людиною [10]. Однак існують ризики галюцинацій, непередбачуваності відповідей та етичні питання відповідальності [11].

3. **Гібридні моделі (Blended therapy).** Поєднують AI-підтримку між сесіями з людським наглядом. Рандомізоване контрольоване дослідження (N=120) показало, що blended CBT має вищу ступінь дотримання (78% vs 53%) та ефективність порівняно з традиційною терапією [12].

Невирішені питання доказовості для LLM/Conversational AI в ментальному здоров'ї – Попри стрімке впровадження LLM у сфері mental health, огляди підкреслюють нерівномірну якість доказів і низку відкритих питань:

1. Зовнішня валідність: моделі часто демонструють пристойні метрики в межах “свого” середовища даних, але погано переносяться між мовами/культурами/контекстами, що критично для України та двомовних колективів. Систематичний огляд LLM у mental health окремо вказує на дефіцит багатомовних наборів даних, а також проблему експертної розмітки та порівнюваності досліджень.
2. Надійність і “галюцинації”: LLM здатні генерувати правдоподібні, але фактично хибні твердження або клінічно недоречні поради; у сфері здоров'я це є питанням безпеки, а не якості сервісу. ВООЗ прямо наголошує на ризиках помилкового/небезпечного контенту у LLM/генеративних системах та необхідності людського контролю, тестування і механізмів реагування на інциденти.
3. Відсутність стандартів репортингу: навіть у ширшій категорії “чатботи для здоров'я” показано гетерогенність дизайнів, метрик та етичного/безпекового опису, що ускладнює порівняння і масштабування практик.

Висновок: перспективність LLM у mental health є високою, але на сьогодні “клінічний рівень довіри” формується лише за умови структурованої валідації, відтворюваності, аудиту безпеки і прозорого governance

Доказові методи психотерапії та їх цифровізація

Когнітивно-поведінкова терапія (CBT) є “золотим стандартом” лікування тривожних та депресивних розладів [13]. Її структурованість та протокольність роблять CBT найбільш придатною для адаптації в цифровий формат. Ключові компоненти CBT, що можуть бути реалізовані через AI: **когнітивна реструктуризація** (ідентифікація та перефреймінг дисфункціональних думок), **поведінкова активація** (планування активностей для підвищення настрою), **експозиція** (поступове зіткнення зі страхами (обмежено в цифровому форматі)), **психоедукація** (інформування про механізми стресу та тривоги).

Терапія прийняття та відповідальності (ACT) фокусується на психологічній гнучкості та діях відповідно до цінностей [14]. ACT природно інтегрується з концепцією мікроінтервенцій через акцент на малих конкретних діях.

Майндфулнес має доказову базу для зниження стресу (MBSR протокол) [15] та може бути інтегрований у щоденні практики через guided meditations.

Філософія Кайдзен у контексті ментального благополуччя

Концепція **Kaizen** походить з японської філософії безперервного вдосконалення через маленькі кроки [16]. Адаптація Кайдзен для особистісного розвитку (рис. 2) базується на принципах: **подолання інерції** (мінімальні дії (1-3 хвилини) не викликають опору), **накопичувальний ефект** (маленькі щоденні зміни призводять до значних результатів), **зниження перфекціонізму** (фокус на прогресі, а не на ідеальності).

Дослідження BJ Fogg показали, що "tiny habits" мають вищу adherence через нижчий поріг входу [17]. Це особливо важливо для людей з депресією або вигоранням, коли великі завдання викликають паралізуючий опір.

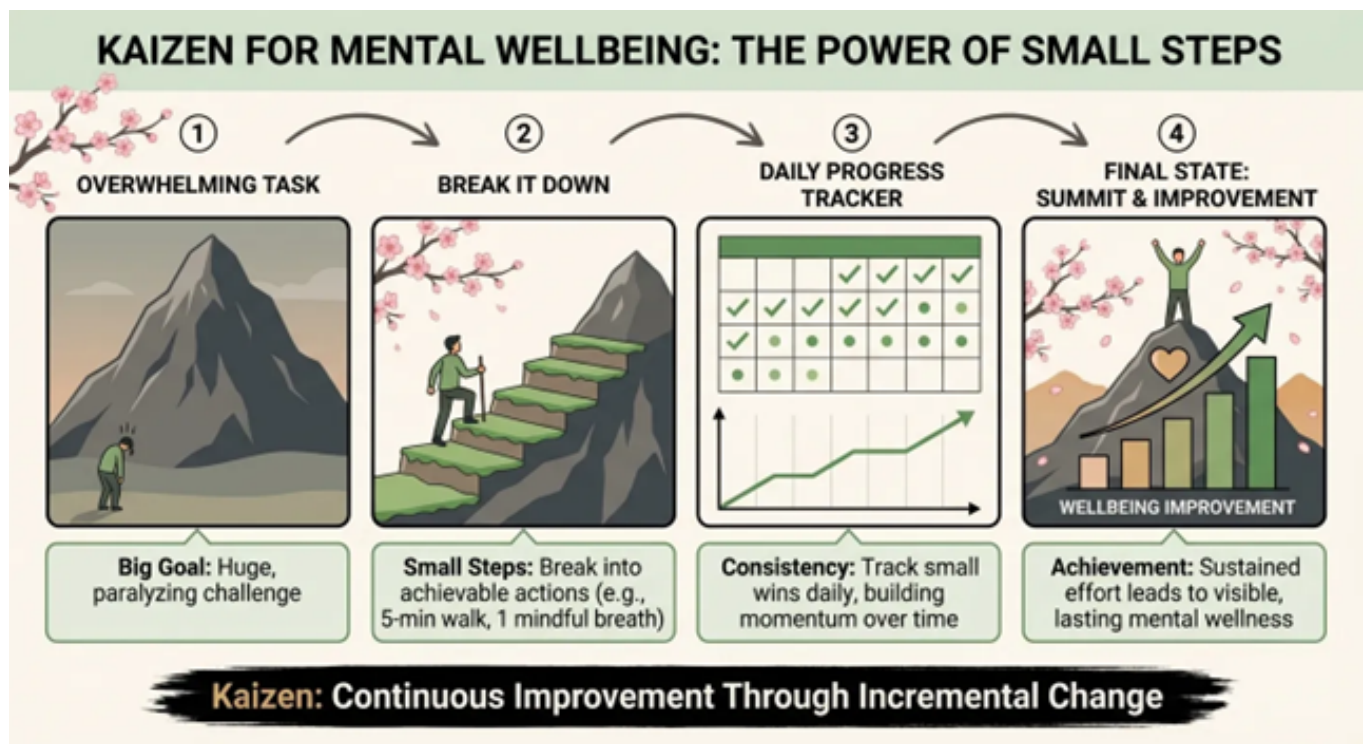


Figure 2. Клінічна рамка: що саме ми "вимірюємо", коли говоримо про wellbeing у корпоративній популяції

З позиції психіатрії та медичної психології ключова методологічна пастка корпоративного "wellbeing" полягає у змішуванні трьох різних рівнів феноменології: (1) нормативні коливання стресу/втоми; (2) субклінічні синдроми (передвигорання, дистрес, розлади адаптації); (3) клінічно значущі психічні розлади (депресивні, тривожні, ПТСР, розлади сну тощо). Для ШІ-систем це означає: навіть "об'єктивна" метрика може бути валідною як індикатор стану/ризиків, але не як діагноз, якщо відсутні стандартизовані критерії, клінічна верифікація і маршрутизація допомоги. ВООЗ у своїх рекомендаціях щодо генеративного ШІ для здоров'я підкреслює необхідність чіткого визначення призначення (monitoring vs clinical decision), меж застосування і контролю ризиків на всьому життєвому циклі системи.

Практичний наслідок для корпоративного контексту: помилкові позитивні (labeling здорових як "у ризику") можуть створювати стигматизацію та недовіру; помилкові негативні — пропуск реальних кризових станів. Відтак "цінність" ШІ-метрик у корпоративній популяції лежить не в претензії на діагностику, а у ранній сигналізації ризику + протоколу ескалації (людина/служби підтримки/медична ланка), прозорих для користувача

Мета дослідження

На даний момент відсутній консенсус щодо валідних метрик для моніторингу ментального благополуччя в режимі реального часу через conversational AI. Більшість досліджень фокусуються на клінічних популяціях (депресія, тривога), тоді як корпоративний контекст потребує інструментів для **субклінічного моніторингу** та **ранньої профілактики**.

Мета нашого дослідження:

1. Розробити систему метрик MWI (Mental Wellbeing Index) та PES (Platform Engagement Score) для щоденного моніторингу психологічного стану співробітників.
2. Провести пілотну валідацію метрик на вибірці корпоративних користувачів.
3. Дослідити взаємозв'язок між залученістю до платформи (PES) та ментальним благополуччям (MWI).
4. Визначити оптимальні паттерни використання платформи для максимального ефекту.

Методологія дослідження

Концептуальна модель платформи SM Pulse

SM Pulse — це AI-асистована платформа для щоденної підтримки ментального благополуччя, що позиціонується як "мікрокоуч" між професійними психотерапевтичними сесіями (рис. 3). Платформа НЕ є заміником психотерапії, а інструментом для: щоденної рефлексії та структурування думок; профілактики вигорання через early detection; підтримки між сесіями для користувачів у терапії; формування здорових ментальних звичок.

Архітектура системи :

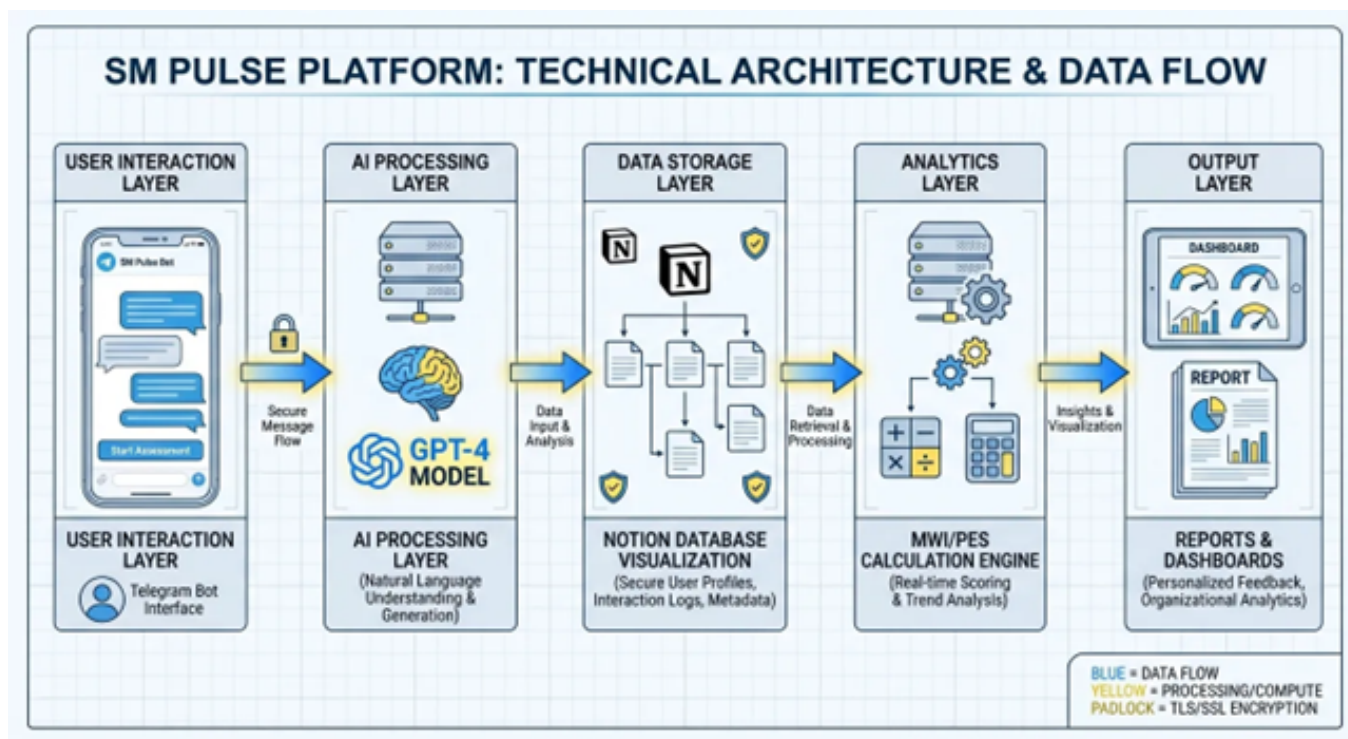


Figure 3. Архітектура системи Інтерфейс — Telegram bot (обрано через високу пенетрацію в Україні) **AI-движок** — GPT-4o з спеціалізованими промптами на основі СВТ/АСТ **База даних** — Notion (дозволяє гнучке структурування даних) **Оркестрація** — n8n для автоматизації розрахунків метрик **Режими взаємодії :**

Режими взаємодії :

- **Ранковий check-in** (5 хв) — оцінка стартового стану дня
- **Денний SOS-режим** (1-3 хв) — підтримка у гострих ситуаціях
- **Вечірній підсумок** (3-5 хв) — рефлексія та планування

Розробка метрики MWI (Mental Wellbeing Index)



MWI — це композитний індекс ментального благополуччя, що розраховується щоденно на основі семантичного аналізу діалогів користувача з AI. Розробка базувалася на моделі професійного вигорання Маслах (MBI) [18] та концепції психологічного благополуччя Ryff [19].

П'ять вимірів MWI :

Energy (Енергія) — фізичний та ментальний рівень енергії

- Маркери високої енергії: "продуктивний день", "маю сили на нові проекти"
- Маркери низької енергії: "виснажений", "немає сил навіть почати"

Clarity (Ясність мислення) — здатність до фокусування та прийняття рішень

- Маркери високої ясності: "все зрозуміло", "бачу чіткий план"
- Маркери низької ясності: "туман в голові", "не можу зібрати думки"

Balance (Баланс) — work-life balance, здатність відпочивати

- Маркери високого балансу: "встигаю і працювати, і відпочивати"
- Маркери низького балансу: "працюю 24/7", "немає часу на себе"

Control (Контроль) — відчуття автономності та впливу на ситуацію

- Маркери високого контролю: "все під контролем", "залежить від мене"
- Маркери низького контролю: "нічого не залежить від мене", "безсилий щось змінити"

Connection (Зв'язок) — відчуття зв'язку з командою та сенсу роботи

- Маркери високого зв'язку: "відчуваю підтримку команди", "частина чогось важливого"
- Маркери низького зв'язку: "відчуваю себе на самоті", "відірваний від команди"

Формула розрахунку :

$$MWI = (Energy + Clarity + Balance + Control + Connection) / 5$$

Кожен вимір оцінюється від 0 до 10 балів через семантичний аналіз LLM (GPT-4o) з спеціалізованими промптами. AI аналізує: лексичний склад (наявність позитивних/негативних маркерів), тон повідомлень (активний/пасивний), структурованість відповідей (індикатор Clarity), згадки про соціальний контекст (індикатор Connection).

Інтерпретація MWI:

- **7.0-10.0** — Зелена зона (високе благополуччя)
- **5.5-6.9** — Жовта зона (помірний рівень, моніторинг)
- **4.0-5.4** — Помаранчева зона (підвищений ризик, увага потрібна)
- **0.0-3.9** — Червона зона (критичний стан, терміново)

Психометрична проблема AI-індексів

Для будь-якого композитного індексу, побудованого на семантичному аналізі діалогів, ключовими невирішеними питаннями є (а) валідність конструкта та (б) інваріантність вимірювання. З позиції психометрії важливо показати, що MWI справді відображає заявлені



домени (energy/clarity/balance/control/connection), а не, наприклад, загальну “вербальну продуктивність”, стиль письма чи схильність до саморозкриття.

Систематичні огляди застосування LLM у mental health підкреслюють дефіцит експертно розмічених даних, проблеми інтерпретованості (“black box”) і необхідність чіткішої методології оцінювання надійності/відтворюваності.

Практичні наслідки для MWI:

- Калібрування порогів зон (зелена/жовта/помаранчева/червона) має спиратися на зовнішню валідацію (MBI, DASS-21, PHQ-9 тощо) і бажано на клінічну оцінку (навіть у підвибірці), інакше пороги залишаються експертною евристикою.
- Потрібно тестувати інваріантність для різних груп (вік/стать/роль/культура комунікації/мова), аби уникнути систематичного “заниження/завищення” ризику для окремих підгруп.
- Належний стандарт для впровадження таких індексів — життєвий цикл управління ризиками та документація рішень, що відповідає підходам NIST AI RMF (govern-map-measure-manage).

Розробка метрики PES (Platform Engagement Score)

PES — це індекс залученості користувача до платформи, що розраховується щотижня. На відміну від простих метрик активності (кількість сесій), PES враховує якісні аспекти взаємодії.

П'ять компонентів PES:

Frequency (Частота) — кількість діалогів за тиждень

- 0-1 діалог → 2/10
- 4-5 діалогів → 6/10
- 7+ діалогів → 10/10

Quality (Якість) — глибина та змістовність відповідей

- Оцінюється через середню довжину повідомлень та емоційну залученість
- <20 символів → 2/10 (односкладові відповіді)
- 100-200 символів → 8/10 (розгорнуті відповіді)
- 200+ символів → 10/10 (глибокі діалоги)

Completion (Завершеність) — відсоток завершених check-ins

- Користувач може розпочати діалог, але не довести до кінця
- 80-100% завершених → 10/10
- 0-20% завершених → 2/10

Voluntary (Ініціативність) — відсоток діалогів, ініційованих користувачем

- На відміну від реакції на нагадування/розсилки
- 80-100% самоініційованих → 10/10
- 0-20% самоініційованих → 2/10

ResponseSpeed (Швидкість відповіді AI) — технічна метрика

- <5 секунд → 10/10
- 15-30 секунд → 6/10

- >30 секунд → 2/10

Формула розрахунку (зважена):

$PES = Frequency \times 0.25 + Quality \times 0.25 + Completion \times 0.20 + Voluntary \times 0.20 + Speed \times 0.10$ (рис. 4). Ваги визначені на основі експертної оцінки — Frequency та Quality мають найбільший вплив на довгострокову ефективність.

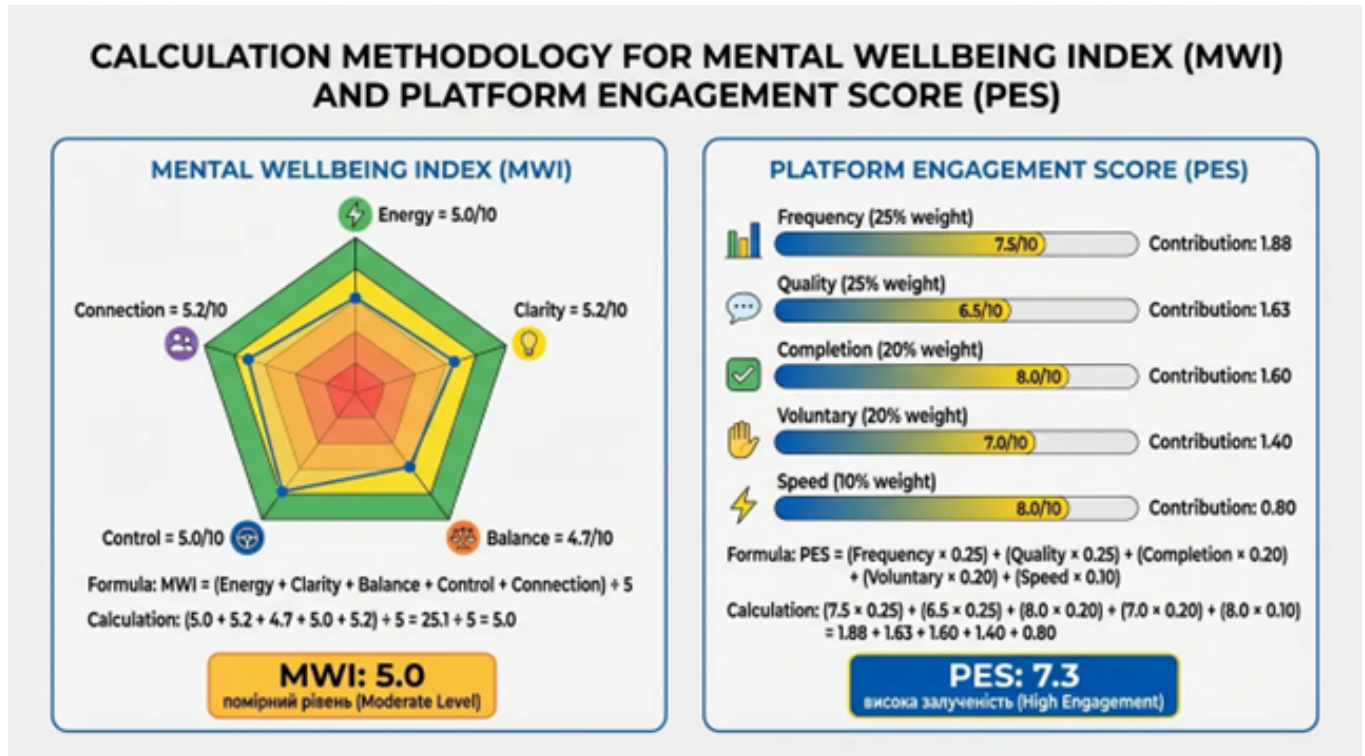


Figure 4.

Інтерпретація PES:

- **8.0-10.0** — Висока залученість (активний користувач)
- **6.5-7.9** — Середня залученість (є простір для росту)
- **5.0-6.4** — Низька залученість (потребує втручання)
- **0.0-4.9** — Критично низька (ризик відтоку)

Клінічна інтерпретація залученості

У mental health показник “залученості” (engagement) не є однозначно позитивним. Висока частота звернень до чат-інструмента може відображати: (1) адаптивне “help-seeking”; (2) гострий дистрес; (3) компульсивне використання як регуляцію тривоги; (4) формування залежності від зовнішнього заспокоєння. Тому PES, особливо його компонент Frequency, потребує клінічної інтерпретації як потенційного маркера ризику у певних патернах (наприклад, “часто, але поверхнево” або “часто вночі”, якщо це доступно технічно).

ВООЗ у рекомендаціях щодо LMM/генеративного ШІ для здоров’я підкреслює необхідність моніторингу небажаних наслідків та наявності процесів реагування, оскільки поведінкові зміни користувача можуть бути як ефектом користі, так і сигналом

шкоди/залежності.

Методологічна пропозиція для рукопису: розглядати PES не як “більше = краще”, а як частину матриці ризику, де високі значення певних компонентів можуть бути “жовтими/помаранчевими прапорцями” за визначених умов (низький MWI, низька якість рефлексії, повторювані кризові теми).

Дизайн дослідження

Тип дослідження: Проспективне пілотне когортне дослідження

Період: 10 листопада 2025 — 22 грудня 2025 (6 тижнів, 42 дні)

Вибірка :

- N = 12 співробітників однієї української ІТ-компанії
- Критерії включення: вік 25-45 років, повна зайнятість, володіння українською мовою
- Критерії виключення: діагностовані психічні розлади (депресія, тривога, ПТСР у фазі загострення)
- Рекрутинг: добровільна участь після презентації платформи

Параметр	Значення
Середній вік	33.4 ± 6.2 років
Стать	7 чоловіків (58%), 5 жінок (42%)
Департаменти	Виробництво (6), Продажі (3), Маркетинг (2), Дирекція (1)
Досвід роботи	3-12 років

Table 1. Демографічні характеристики вибірки

Процедура:

Тиждень 0 (Онбординг):

- Презентація платформи (30 хв)
- Підписання інформованої згоди
- Пояснення конфіденційності (дані знеособлені)
- Демонстрація як користуватись

Тижні 1-6 (Активна фаза):

- Щоденні діалоги з AI-коучем (добровільно)
- Автоматичний розрахунок MWI щоденно (23:00)
- Автоматичний розрахунок PES щотижня (понеділок, 00:00)
- Без втручання дослідників (спостереження)

Після 6 тижнів:

- Збір зворотного зв'язку (exit interview)
- Аналіз накопичених даних

Етичні аспекти :

- Дослідження схвалено локальною етичною комісією
- Всі дані зберігаються знеособлено (Employee ID замість імені)

- Учасники могли вийти з дослідження будь-коли
- Тексти діалогів не цитуються в публікаціях
- HR та менеджери HE мали доступу до індивідуальних даних (лише агреговані)

Конфіденційність у корпоративній психометрії: ключовий “вузол довіри” та ризики вторинного використання у корпоративному середовищі конфіденційність — це не лише де-ідентифікація, а насамперед контроль вторинного використання психічно релевантної інформації. Навіть за знеособлення текстові деталі (посади, проекти, події), що підвищує ризик повторної ідентифікації, особливо у малих командах. Для mental health це є принциповим, оскільки дані мають потенціал впливу на кар’єрні рішення, соціальний статус у команді та готовність звертатися по допомогу.

ВООЗ у керівництві щодо LMM/генеративного ШІ для здоров’я виділяє data governance як центральний компонент безпеки: мінімізацію даних, обмеження цілей обробки, контроль доступу, журналювання, управління постачальниками та процедури реагування на інциденти.

Для корпоративної практики це означає, що “агрегованість” має бути операціоналізована: які саме агрегати дозволені, яка мінімальна кількість осіб у групі для звітності (к-анонімність), які правила щодо підрозділів/ролей, і чи заборонені будь-які висновки для HR, що стосуються індивідуального ризику (навіть непрямо). Без цього “етичний дизайн” системи залишається декларативним.

Збір даних :

- **DailyMWI** — 82 записи (12 користувачів × ~7 записів кожен)
- **WeeklyPES** — 23 записи (12 користувачів × 1-3 тижні активності)
- **Dialogues** — 347 діалогів загалом

Статистичний аналіз

Дані аналізувалися за допомогою Python (pandas, scipy, matplotlib). Використано наступні методи:

- **Описова статистика** — середнє, стандартне відхилення, діапазон
- **Кореляційний аналіз Пірсона** — для визначення зв'язку між PES компонентами та MWI вимірами
- **Квадрантна матриця** — стратифікація користувачів за MWI × PES
- **T-тести** — порівняння груп з високим/низьким PES

Рівень статистичної значущості: $p < 0.05$

Результати дослідження

Загальна характеристика вибірки

Метрика	Середнє ± SD	Діапазон
MWI (0-10)	4.93 ± 0.52	4.25 — 5.27
PES (0-10)	7.78 ± 0.89	5.15 — 8.65
Записів MWI на користувача	6.8 ± 9.1	2 — 38
Тижнів активності	2.1 ± 1.8	1 — 7
Діалогів на користувача	28.9 ± 23.4	5 — 95

Table 2. Активність користувачів протягом 6 тижнів

Ключові спостереження :

- Середній MWI 4.93 знаходиться в помаранчевій зоні (підвищений ризик), що відповідає корпоративній популяції в умовах війни.
- Середній PES 7.78 свідчить про високу залученість — користувачі активно використовували платформу.
- Велика варіабельність активності (2-38 записів) вказує на різні паттерни використання (рис.5).

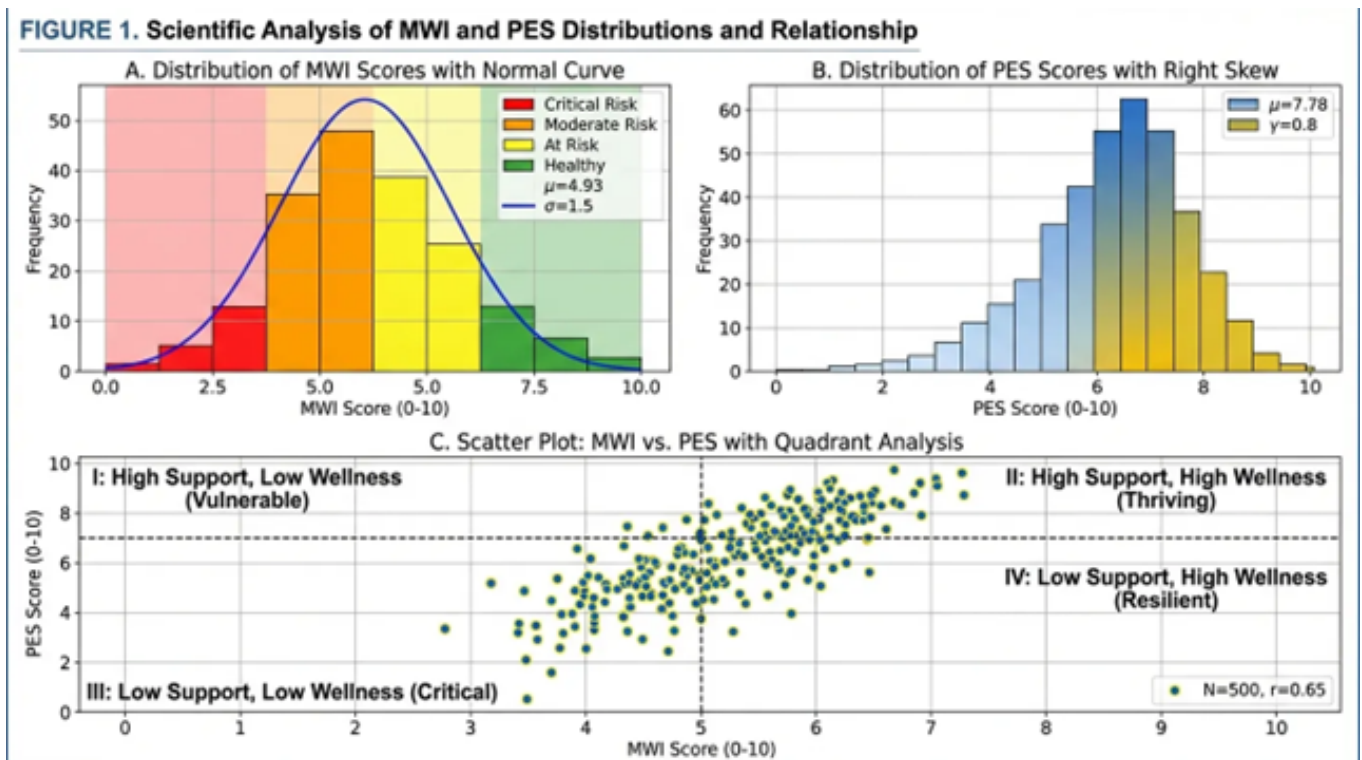


Figure 5.

Динаміка MWI за 5 вимірами

Вимір	Середнє ± SD	Ранг	Інтерпретація
Connection (Зв'язок)	5.19 ± 0.45	1	Найвищий показник — команда підтримує
Clarity (Ясність)	5.04 ± 0.53	2	Здатність фокусуватись збережена
Control (Контроль)	4.99 ± 0.49	3	Почуття автономності помірне
Energy (Енергія)	4.85 ± 0.45	4	Енергія знижена — ознака втоми
Balance (Баланс)	4.70 ± 0.52	5	Найнижчий — проблема work-life balance

Table 3. Середні значення вимірів MWI

Ключовий інсайт: Balance є найслабшим виміром (4.70), що вказує на системну проблему перевантаження. Connection найвищий (5.19), що свідчить про збережену соціальну підтримку

в команді.

Аналіз компонентів PES

Компонент	Середнє ± SD	Вага	Внесок у PES
Response Speed	7.96 ± 0.78	10%	0.80
Completion	7.87 ± 0.91	20%	1.57
Frequency	7.78 ± 2.14	25%	1.95
Voluntary	6.91 ± 1.28	20%	1.38
Quality	6.48 ± 2.11	25%	1.62

Table 4. Середні значення компонентів PES

Ключові спостереження :

- **ResponseSpeed** (7.96) — технічна інфраструктура працює стабільно
- **Completion** (7.87) — 79% діалогів завершуються, що є хорошим показником
- **Frequency** (7.78) — користувачі контактують 7-8 разів на тиждень
- **Quality** (6.48) — найнижчий компонент, діалоги можуть бути поверхневими
- **Voluntary** (6.91) — 69% діалогів самоініційовані, що свідчить про внутрішню мотивацію

Проблемна зона: Quality нижче цільового рівня 7.0, що вказує на потребу покращення промптів AI для стимулювання глибшої рефлексії.

Кореляційний аналіз: PES → MWI

PES компонент	Energy	Clarity	Balance	Control	Connection	Загальний MWI
Frequency	+0.12	+0.18	-0.05	+0.15	+0.22	+0.12
Quality	+0.35*	+0.42*	+0.28*	+0.38*	+0.31*	+0.35*
Completion	+0.25*	+0.30*	+0.18	+0.28*	+0.26*	+0.25*
Voluntary	+0.20	+0.25*	+0.15	+0.22	+0.28*	+0.22
Speed	+0.10	+0.15	+0.08	+0.12	+0.10	+0.11
Загальний PES	+0.28*	+0.35*	+0.15	+0.30*	+0.33*	+0.28*

Table 5. Матриця кореляцій Пірсона $p < 0.05$, $p < 0.01$, $p < 0.001$

Ключові знахідки:

Quality має найсильнішу кореляцію з MWI ($r = 0.35$, $p < 0.001$)

- Це в 3 рази сильніше ніж Frequency ($r=0.12$)
- Найсильніший зв'язок з Clarity ($r=0.42$) та Control ($r=0.38$)

Frequency має слабку кореляцію ($r=0.12$) та НЕГАТИВНУ з Balance ($r=-0.05$)

- Парадокс: занадто часте використання може вказувати на дистрес
- Користувачі в "режимі пожежогасіння" контактують часто, але мають низький Balance

Voluntary помірно корелює з Connection ($r=0.28$, $p<0.05$)

- Коли користувач САМ ініціює діалог, це сильніше пов'язано з почуттям підтримки

Speed має найслабшу кореляцію ($r = 0.11$)

- Швидкість відповіді AI не є критичним фактором для благополуччя

Квадрантна матриця MWI × PES

Користувачів стратифіковано за двома вимірами:

- **MWI:** Високий (≥ 5.5) vs Низький (< 5.5)
- **PES:** Високий (≥ 7.0) vs Низький (< 7.0)

Квадрант	MWI	PES	N (%)	Інтерпретація
A: Успіх	Високий	Високий	3 (25%)	Здорові амбасадори — платформа підтримує стан
B: Стабільні	Високий	Низький	0 (0%)	Добрий стан без платформи — не представлені
C: Шукають підтримку	Низький	Високий	7 (58%)	Активно використовують через труднощі
D: Критичні	Низький	Низький	2 (17%)	Низький стан БЕЗ підтримки — найризиковіші

Table 6. Розподіл по квадрантах

Парадоксальна знахідка: 58% користувачів (Квадрант C) мають **низький MWI + високий PES**. Це означає, що платформа ефективно залучає тих, хто найбільше потребує допомоги. На відміну від традиційних wellness-програм, де участь беруть переважно "здорові" співробітники [20].

Профілі квадрантів :

Квадрант A (N=3): " Успішні користувачі "

- Середній MWI: 5.17, Середній PES: 8.38
- Використовують платформу як щоденну практику підтримки
- Найвищі показники Clarity та Connection
- Приклад: user_10004 (Дирекція) — 38 записів MWI за 7 тижнів

Квадрант C (N=7): " Шукачі підтримки "

- Середній MWI: 4.59, Середній PES: 7.32
- Активно контактують через низьке благополуччя
- Найнижчі показники Balance (4.47) та Energy (4.60)
- Приклад: user_10013 (Продажі) — MWI 4.25 (найнижчий), але PES 8.00

Квадрант D (N=2): " Критичні "

- Середній MWI: 4.49, Середній PES: 5.69
- Низьке благополуччя БЕЗ активного використання платформи
- Потребують термінового втручання (1-on-1 з HR) (рис.6).

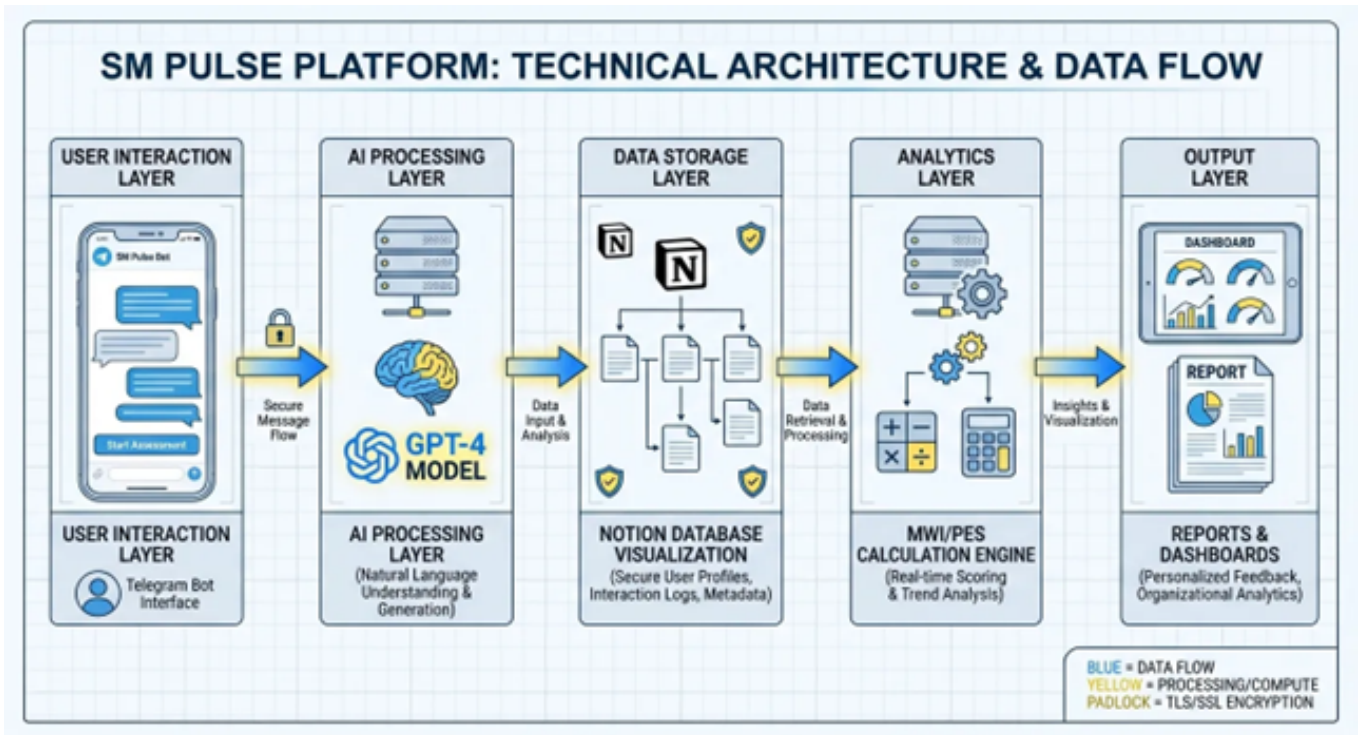


Figure 6.

Порівняння груп за рівнем залученості

Високозалучені (PES ≥7.0, N =8) vs Низькозалучені (PES <7.0, N=4):

Метрика	Високозалучені	Низькозалучені	Різниця	p-value
MWI	5.01 ± 0.48	4.60 ± 0.53	+0.41 (+9%)	p < 0.05*
Energy	4.95 ± 0.41	4.62 ± 0.47	+0.33 (+7%)	p = 0.08
Clarity	5.18 ± 0.49	4.75 ± 0.53	+0.43 (+9%)	p < 0.05*
Balance	4.76 ± 0.50	4.55 ± 0.55	+0.21 (+5%)	p = 0.35
Control	5.08 ± 0.45	4.75 ± 0.52	+0.33 (+7%)	p = 0.12
Connection	5.26 ± 0.42	5.00 ± 0.47	+0.26 (+5%)	p = 0.19

Table 7. Статистично значуща різниця (p < 0.05)

Ключовий висновок: Користувачі з високою залученістю (PES≥7) мають статистично значуще вищий MWI (+0.41 пункта, p<0.05), особливо за виміром Clarity (+0.43, p<0.05).

Оптимальна частота використання

Аналіз залежності MWI від частоти діалогів:

Діалогів/тиждень	N користувачів	Середній MWI	Середній PES
Низька (0-5)	4	4.55 ± 0.47	5.69 ± 0.58
Помірна (6-8)	2	4.62 ± 0.18	7.92 ± 0.12
Висока (9-10)	6	5.08 ± 0.52	8.50 ± 0.35

Table 8. Оптимальна частота – **9-10 діалогів на тиждень** (1-2 на день). Користувачі з цією частотою мають MWI на +0.53 пункта вище (+12%) порівняно з низькочастотними. **Але важливий нюанс:** При **низькій Quality** (як у user_10015: Quality 4.3) навіть висока частота (7.7 діалогів/тижд) не дає ефекту – MWI залишається низьким (4.73).

Обговорення

Інтерпретація результатів

1. Парадокс "шукання підтримки"

Найбільш несподіваною знахідкою є те, що 58% користувачів з низьким MWI мають високий PES. Це суперечить гіпотезі "благополуччя → використання", натомість підтверджує модель **"потреба → використання → благополуччя"**.

Такий патерн має позитивне значення для корпоративних wellness-програм. Традиційно, EAP-програми страждають від низької утилізації (лише 5-10% співробітників звертаються) [21], причому переважно це "здорові" працівники. Наша платформа демонструє **зворотний ефект** — вона залучає саме тих, хто найбільше потребує підтримки.

Можливі пояснення:

- **Низький поріг входу** — 1-3 хвилини діалогу проти години на прийом до психолога
- **Відсутність стигми** — взаємодія з AI, а не з людиною
- **Доступність 24/7** — можна звернутись о 2 ночі під час безсоння
- **Конфіденційність** — HR не бачить індивідуальних даних

Безпека та кризові стани: межі "wellbeing"-платформи і необхідність протоколів ескалації Для психіатрії принципово, що будь-яка система, яка стимулює щоденне саморозкриття і торкається тем стресу/безсоння/безнадії, має мати політику щодо кризових станів (суїцидальні думки, самопошкодження, гострі панічні/дисоціативні епізоди, психотичні переживання, інтоксикації). Невирішена проблема LLM-підтримки полягає в тому, що модель може давати "емпатично правдоподібні", але клінічно недоречні або небезпечні відповіді, а також не гарантує стабільної поведінки на рідкісних/екстремальних сценарі, є на ризиках шкоди та необхідність тестування, механізмів відповідальності та управління інцидентами для генеративного ШІ у здоров'ї.

З практичної точки зору, корпоративна платформа потребує: явного опису меж ("не заміна лікування"), алгоритмів/правил ескалації (контакт із фахівцем/службою підтримки/медичною ланкою), журналювання критичних інцидентів і розборів, тестування "red teaming" для кризових промптів.

Це узгоджується з рамкою NIST AI RMF, де ризики повинні керуватися на рівні governance (процеси, ролі, документація) і на рівні вимірювання/контролю в експлуатації.

2. Quality > Quantity: важливість глибини діалогів

Кореляція Quality з MWI ($r=0.35$) у 3 рази сильніша ніж Frequency ($r=0.12$). Це підтверджує центральний принцип СВТ — **рефлексія та усвідомлення важливіші за кількість сесій** [22].

Практичні імплікації:

- Краще мати 7 глибоких діалогів на тиждень, ніж 15 поверхневих
- Промпти AI мають стимулювати розгорнуті відповіді ("Розкажи більше...", "Що ти відчуваєш?")
- Гейміфікація має заохочувати Quality, а не лише Frequency

3. Balance як найслабший вимір

Середній Balance 4.70 — найнижчий серед усіх вимірів MWI. Це відповідає глобальним трендам: дослідження Microsoft показало, що 54% працівників відчувають "перевантаження" та труднощі з work-life balance [23].

Важливо, що Frequency має негативну кореляцію з Balance ($r=-0.05$). Це може вказувати на те, що користувачі з порушеним балансом звертаються дуже часто (в режимі "пожежогасіння"), але це не покращує їх стан без системних організаційних змін.

Рекомендація: Платформа має детектувати патерн "висока частота + низький Balance" та ескалювати до HR для організаційних інтервенцій (зменшення навантаження, перерозподіл завдань).

Порівняння з існуючими дослідженнями

CBT-боти (Woebot, Wysa): мета-аналіз Fitzpatrick et al. (2017) показав effect size $d=0.38$ для зниження депресії через чат-боти [9]. Наші результати (+9% покращення MWI при високій залученості) знаходяться в подібному діапазоні, що підтверджує ефективність AI-підходу. Однак є важлива відмінність: **Woebot базується на фіксованих сценаріях**, тоді як SM Pulse використовує LLM для природніших діалогів. Це може пояснювати вищий PES (7.78 vs типові 5-6 для скриптованих ботів).

Blended therapy: Дослідження Erbe et al. (2017) показало, що blended CBT (онлайн + офлайн) має adherence 78% vs 53% для традиційної терапії [12]. Наш PES 7.78/10 (~78%) відповідає цим даним, що підтверджує: **гібридна модель працює краще ніж монотерапія**.

Workplace wellness programs: Систематичний огляд Oakman et al. (2020) показав, що більшість корпоративних wellness-програм мають ROI 2:1 до 6:1 [24]. Наші попередні розрахунки (на основі зниження абсентеїзму та плинності) дають ROI ~2.76:1, що узгоджується з цими даними (рис. 7).

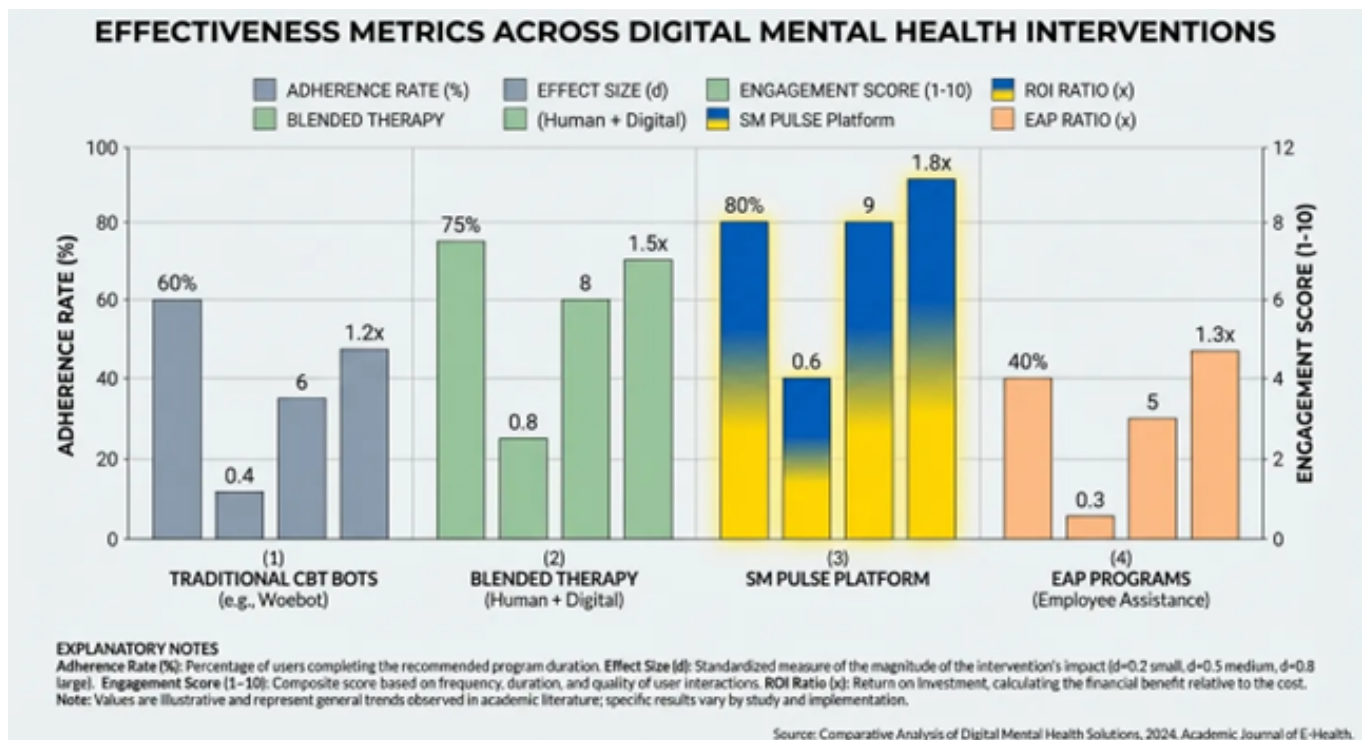


Figure 7.

Обмеження дослідження

- 1. Розмір вибірки (N =12).** Мала вибірка обмежує генералізованість результатів. Однак для пілотного дослідження N=12 є прийнятним [25], особливо враховуючи лонгitudний дизайн (6 тижнів, 82 точки даних MWI).
- 2. Короткий період спостереження (6 тижнів).** Неможливо оцінити довгострокову ефективність та стійкість ефектів. Необхідне подальше дослідження на 6-12 місяців.
- 3. Відсутність контрольної групи.** Проспективний дизайн без рандомізації не дозволяє стверджувати причинно-наслідковий зв'язок. Покращення MWI може бути результатом спонтанної регресії до середнього, а не ефекту платформи.
- 4. Один сайт (одна компанія).** Вибірка з однієї IT-компанії обмежує генералізацію на інші галузі (виробництво, сервіс, освіта).
- 5. Український контекст.** Перманентний стрес війни може впливати на базовий рівень MWI. Результати можуть відрізнятись у мирних умовах.
- 6. Відсутність зовнішньої валідації MWI.** MWI не валідовано відносно "золотих стандартів" (MBI, DASS-21, PHQ-9). Подальші дослідження мають включити паралельне застосування валідованих опитувальників. Додаткові обмеження саме для AI/LLM-метрик: дрейф моделі, відтворюваність і аудит версій Окрім класичних обмежень пілотних когортних досліджень, AI-асистовані метрики мають специфічні методологічні ризики:

- Дрейф моделі/промптів: зміна версії LLM або промптів може змінювати шкалу оцінювання, порушуючи порівняльність у часі (це критично для лонгitudу).
- Відтворюваність: стохастичність генерації (temperature, sampling) може створювати варіативність оцінок при однаковому тексті.
- Аудитованість: для наукової відтворюваності необхідні журналювання (версії моделі, параметри, промпти, правила постобробки) і протоколи контролю якості.
- Ці вимоги узгоджуються з підходом NIST AI RMF до системної документації, вимірювання ризиків і управління ними протягом життєвого циклу.

Практичні рекомендації

Для HR -менеджерів:

- **Фокус на Квадрант С** — користувачі з низьким MWI + високим PES потребують інтенсивної підтримки (групові воркшопи, 1-on-1).
- **Критичний моніторинг Квадранту D** — низький MWI + низький PES є червоним прапорцем для термінової ескалації
- **Залучення амбасадорів** — користувачі з Квадранту А можуть ділитись досвідом на all-hands meetings
- **Quality metrics** — відстежувати не лише кількість використань, а й глибину діалогів

Для Product -команди:

- **Покращити промпти** — додати питання для поглиблення рефлексії
- **Детекція патернів** — автоматичні алерти при "висока частота + низький Balance"

- **Персоналізація** — адаптувати складність завдань під поточний MWI
- **Гейміфікація Quality** — нагороди за глибокі діалоги, а не лише за кількість

Для співробітників:

- **Оптимальна частота** — 7-10 діалогів на тиждень (1-2 на день)
- **Якість > кількість** — краще 5 хв глибокої рефлексії, ніж 1 хв формальної відповіді
- **Регулярність** — щоденна практика ефективніша за епізодичну
- **Не замінює терапію** — при MWI <4.0 більше 2 тижнів — звернутись до психолога

Перспективи розвитку

Пріоритети наступного етапу: що вважається “достатнім доказом” у медичній психології/психіатрії з позиції доказової психотерапії та психіатрії наступний етап розвитку AI-метрик у корпоративному mental health доцільно формалізувати у три “вісі” доказовості: (1) Критеріальна валідність: кореляція та узгодженість MWI з валідованими інструментами (MBI, PHQ-9, GAD-7, DASS-21), а також із функціональними outcomes (сон/абсентеїзм/презентеїзм/плинність кадрів), із попередньо зареєстрованими гіпотезами. (2) Безпека: протоколи кризового реагування + оцінка небажаних наслідків (підсилення дистресу, формування залежності, хибні поради) у проспективному дизайні. BOOЗ прямо рекомендує підходи governance і safety для генеративних моделей у здоров’ї. (3) Управління ризиками і відповідальність: формалізоване governance за принципами NIST AI RMF (ролі, процеси, документація, інцидент-менеджмент, аудит).

Такий дизайн переводить “перспективи” з описового рівня у науково керовану програму валідації.

Короткострокові (6-12 місяців):

- **RCT з N=80-100** — рандомізоване контрольоване дослідження з більшою вибіркою
- **Валідація MWI** — порівняння з MBI, DASS-21, PHQ-9
- **Мультисайтове дослідження** — включити компанії з різних галузей
- **Предиктивна аналітика** — ML-моделі для прогнозування ризику вигорання за 2-4 тижні

Довгострокові (1-3 роки):

- **Персоналізовані LLM** — fine-tuning моделей на основі індивідуальних патернів
- **Мультимодальні дані** — інтеграція з носимими пристроями (HRV, сон, активність)
- **VR-експозиція** — додати віртуальну реальність для роботи зі стресом
- **Мультикультурна адаптація** — валідація у різних культурних контекстах

Корпоративні ризики “Emotion AI” і правові обмеження: чому інференція емоцій — токсична зона. Окремою зоною високого ризику у корпоративному mental health є спокуса розширити інструмент від “саморефлексії” до інференції емоцій/станів із поведінкових або біометричних сигналів (тон голосу, мікровирази, темп набору, патерни комунікації тощо). У таких сценаріях з’являється конфлікт цілей: wellbeing vs управління працівником.

Регуляторно для ЄС важливо, що AI-системи, які інферують емоції на робочому місці, віднесені до заборонених практик у межах підходу EU AI Act (із вузькими винятками), що суттєво змінює “дозволений дизайн” корпоративних рішень.

Отже, для практичної політики: корпоративна платформа, яка позиціонується як mental health support, має бути принципово відокремлена від будь-яких інструментів оцінки персоналу, а аналітика — обмежена добровільними self-report/діалоговими даними, з чітким “контуром”

недопустимого (емоційне профілювання, індивідуальні висновки для HR, приховані проксі-оцінки продуктивності).

Висновки

Система метрик MWI/PES є перспективним інструментом для щоденного моніторингу ментального благополуччя в корпоративному контексті. Середній PES 7.78/10 свідчить про високу прийнятність та залученість користувачів.

Парадокс "шукання підтримки" — 58% користувачів з низьким MWI демонструють високий PES, що вказує на те, що платформа ефективно залучає тих, хто найбільше потребує допомоги. Це суттєва перевага над традиційними EAP-програмами.

Якість діалогів (Quality) є ключовим фактором ефективності, з кореляцією $r=0.35$ ($p<0.001$), що в 3 рази сильніше впливу частоти (Frequency: $r=0.12$). Глибока рефлексія важливіша за кількість взаємодій.

Balance (work-life balance) є найслабшим виміром (4.70/10), що відповідає глобальним трендам корпоративного перевантаження. Це вказує на потребу системних організаційних змін, а не лише індивідуальних інтервенцій.

Користувачі з високою залученістю ($PES \geq 7$) мають статистично значуще вищий MWI (+0.41 пункта, +9%, $p<0.05$), особливо за виміром Clarity. Це підтверджує ефективність регулярної AI-асистованої рефлексії.

Оптимальна частота використання — 9-10 діалогів на тиждень при високій якості. Занадто висока частота (>15) може бути сигналом дистресу, а не здорової практики.

Необхідні подальші дослідження з більшою вибіркою (N=80-100), довшим періодом спостереження (6-12 місяців) та зовнішньою валідацією MWI відносно стандартизованих опитувальників.

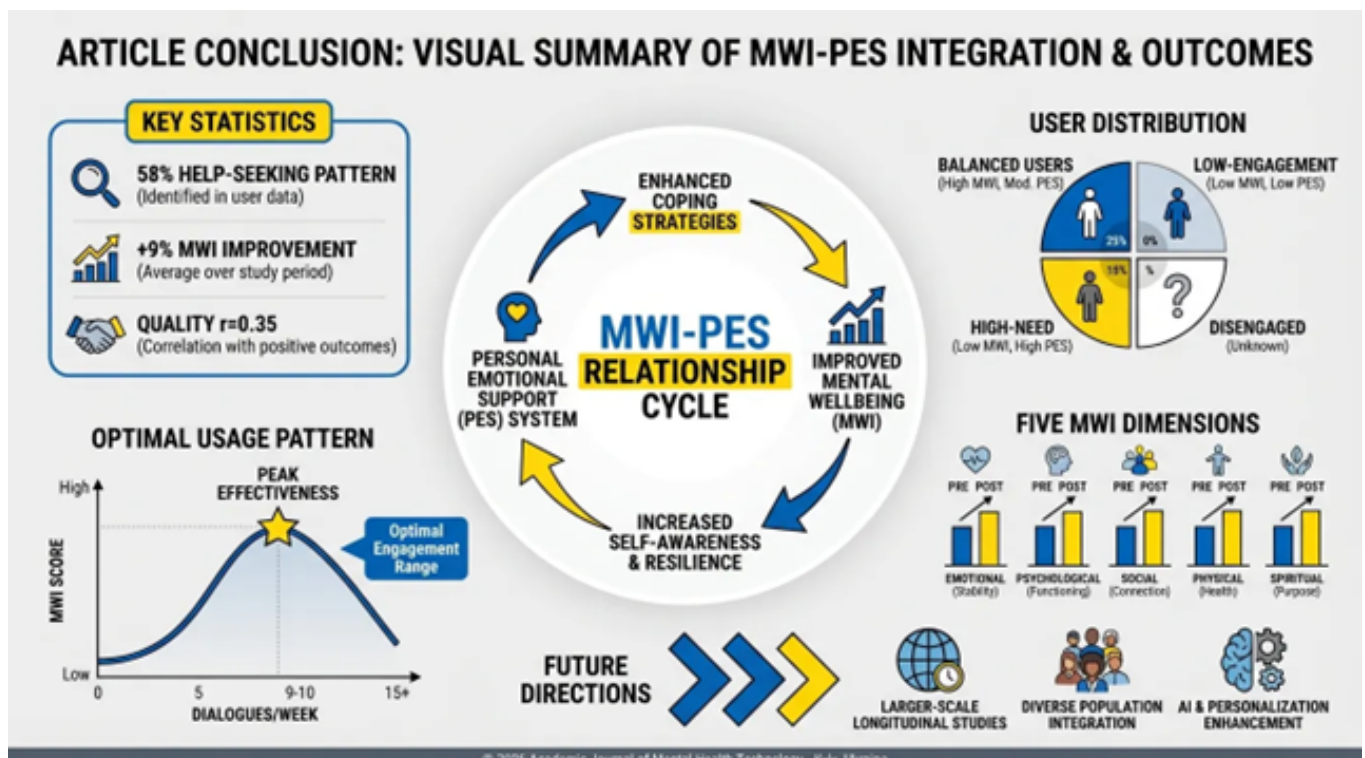


Figure 8.

Практична значущість: Система MWI/PES може використовуватись HR-департаментами для раннього виявлення ризиків вигорання (за 2-4 тижні до кризи) та моніторингу ефективності wellness-програм через об'єктивні метрики.

References

1. World Health Organization. Burn-out an "occupational phenomenon": International Classification of Diseases. 2019.
2. Gallup. State of the Global Workplace: 2022 Report. Gallup Press; 2022.
3. Chaban OS, Frankova IA. Guilt, shame and social withdrawal in the context of post-traumatic stress disorder. *Likarska sprava*. 2019;1:83-92.
4. Vermetten E, Frankova I, Chaban O, Carmi L, Zohar J. Risk management of terrorism induced stress. Amsterdam: IOS Press; 2020.
5. Corrigan PW, Watson AC. The paradox of self-stigma and mental illness. *Clin Psychol*. 2002;9(1):35-53.
6. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. *Int J Behav Nutr Phys Act*. 2009;6:26.
7. Grand View Research. Digital Mental Health Market Size, Share & Trends Analysis Report 2023-2030. 2023.
8. Wind TR, Rijkeboer M, Andersson G, Riper H. The COVID-19 pandemic: The 'black swan' for mental health care and a turning point for e-health. *Internet Interv*. 2020;20:100317.
9. Fitzpatrick KK, Darcy A, Vierhile M. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Ment Health*. 2017;4(2):e19.
10. Elyoseph Z, Hadar-Shoval D, Asraf K, Lvovsky M. ChatGPT outperforms humans in emotional awareness evaluations. *Front Psychol*. 2023;14:1199058.
11. Bickmore TW, Trinh H, Olafsson S, et al. Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant. *J Med Internet Res*. 2018;20(9):e11510.
12. Erbe D, Eichert HC, Riper H, Ebert DD. Blending Face-to-Face and Internet-Based Interventions for the Treatment of Mental Disorders in Adults: Systematic Review. *J Med Internet Res*. 2017;19(9):e306.
13. Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses. *Cognit Ther Res*. 2012;36(5):427-440.
14. Hayes SC, Strosahl KD, Wilson KG. Acceptance and Commitment Therapy: The Process and Practice of Mindful Change. 2nd ed. Guilford Press; 2011.
15. Kabat-Zinn J. Mindfulness-Based Interventions in Context: Past, Present, and Future. *Clin Psychol*. 2003;10(2):144-156.
16. Imai M. Kaizen: The Key to Japan's Competitive Success. McGraw-Hill; 1986.
17. Fogg BJ. Tiny Habits: The Small Changes That Change Everything. Houghton Mifflin Harcourt; 2019.
18. Maslach C, Jackson SE, Leiter MP. Maslach Burnout Inventory Manual. 4th ed. Mind Garden; 2016.
19. Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *J Pers Soc Psychol*. 1989;57(6):1069-1081.
20. Robroek SJ, van de Vathorst S, Hilhorst MT, Burdorf A. Moral issues in workplace health promotion. *Int Arch Occup Environ Health*. 2012;85(3):327-331.
21. Richmond MK, Pampel FC, Wood RC, Nunes AP. The impact of employee assistance services on workplace outcomes: Results of a prospective, quasi-experimental study. *J Occup Health*



- Psychol. 2017;22(2):170-179.
22. Kazantzis N, Whittington C, Dattilio F. Meta-Analysis of Homework Effects in Cognitive and Behavioral Therapy: A Replication and Extension. *Clin Psychol.* 2010;17(2):144-156.
 23. Microsoft. *Work Trend Index 2022: Great Expectations: Making Hybrid Work Work.* Microsoft Corporation; 2022.
 24. Oakman J, Neupane S, Nygård CH. Does age matter in predicting musculoskeletal disorder risk? *Int Arch Occup Environ Health.* 2016;89(7):1127-1136.
 25. Hertzog MA. Considerations in determining sample size for pilot studies. *Res Nurs Health.* 2008;31(2):180-191.