Health, Empathy, and AI Learning (HEAL)
Vol. 2, No. 1, November, 2026, pp. 12−22
E-ISSN: 3163-267X | P-ISSN: 3163-2661

Advancing Mental Health Support Through Orange Technology and Empathetic Artificial Intelligence

Dwi Julianingsih[1*], Tasya Syafira[2], Sheila Aulia Anjani[3]
[1]Faculty of Science and Technology, Universitas Raharja, Indonesia
[2]Department of Computer Science, Ilearning Research Center, Colombia
[3]Faculty of Economics and Business, Universitas Raharja, Indonesia
dwi.julianingsih@raharja.info, syafiratasya11@ilearning.co, sheila@raharja.info
***Corresponding Author**

Article Info

Article history:

Submission August 12, 2026
Revised September 9, 2026
Accepted October 21, 2026
Published October 30, 2026

Keywords:

Conversational AI
Empathetic Artificial Intelligence
Emotional Wellbeing
Mental Health System
Orange Technology

ABSTRACT

Mental health systems face a persistent gap in accessible, continuous, and person-centred support. Conversational AI can expand access to psychoeducation, emotional check-ins, self-management, and service navigation, but may also introduce risks of digital exclusion, cultural mismatch, privacy concerns, unsafe advice, and overreliance on automation. **This paper develops** a conceptual framework integrating Orange Technology, which promotes emotional wellbeing, cultural sensitivity, social connectedness, and resilience, with empathetic AI for mental health support. **An integrative literature review** examined peer-reviewed systematic reviews and WHO and UN guidance published from 2021 to 2026. **Evidence suggests** that conversational AI can improve depression, distress, anxiety, and engagement, although effects are heterogeneous and may decline over time. Personalization and empathic interaction may enhance engagement, while the framework's effects on community resilience, cultural sensitivity, and community resource integration require further empirical validation. **The framework** therefore positions AI as a supportive, human-in-the-loop layer rather than a replacement for clinicians, families, peers, or communities. It connects emotional sensing, empathetic dialogue, personalized support, escalation pathways, community connection, inclusive access, and governance, while aligning mental health innovation with SDG 3, 9, and 17.

This is an open access article under the CC BY 4.0 license.

DOI: https://doi.org/10.68012/heal.v2i1.272
This is an open-access article under the CC-BY license (https://creativecommons.org/licenses/by/4.0/)
©Authors retain all copyrights

1. INTRODUCTION

Mental health is a global public-health priority, yet health systems remain inadequate in providing timely, continuous, and person-centred support. The WHO highlights persistent gaps in mental-health services, while the COVID-19 pandemic intensified anxiety and depressive symptoms and disrupted care [1, 2]. Artificial Intelligence (AI), particularly conversational agents, can extend psychoeducation, self-management, emotional check-ins, behavioural prompts, and service navigation beyond conventional clinical settings. Systematic reviews report potential improvements in depression, distress, anxiety, and engagement, with personalization and empathic responses identified as important facilitators [3, 4]. However, effects remain heterogeneous and may decline over time, while AI deployment raises concerns regarding digital exclusion, cultural mismatch, privacy, unsafe advice, and overreliance on automation. AI should therefore complement rather than replace

professional mental-health care and human relationships [5, 6].

In this study, Orange Technology is defined as a human-centred design orientation that directs technology toward emotional wellbeing, cultural sensitivity, social connectedness, and resilience. Unlike human-centred AI, which emphasizes human needs, agency, usability, and oversight; affective computing, which focuses on recognizing and responding to affective signals; and empathetic AI, which emphasizes supportive and context-sensitive interaction, Orange Technology provides a broader wellbeing and social-resilience perspective. It is not a specific technology, algorithm, or clinical intervention, but a conceptual lens for connecting AI capabilities with inclusive access, cultural responsiveness, community support, and human wellbeing. The research gap lies in the limited integration of AI-based mental-health interventions, empathetic conversational-agent design, and community resilience and digital inclusion [7, 8]. To address this gap, the study proposes a human-in-the-loop framework comprising six connected layers: Inclusive Access, Emotional and Contextual Understanding, Empathetic Dialogue, Personalized Support, Safety and Human Escalation, and Community Resilience and Governance. These layers synthesize recurring requirements from the literature and position empathy as a functional design requirement involving context-sensitive responses, respect for autonomy, uncertainty awareness, and timely human escalation [9].

The framework positions AI as a supportive layer rather than a replacement for clinicians, families, peers, or communities. Its contributions are threefold, including synthesizing evidence on conversational AI and empathetic mental health support, proposing an integrated architecture that links emotional understanding, personalization, safety, human oversight, community connection, and inclusive access, and providing a conceptual design and evaluation structure for future empirical research. Its novelty lies in integrating these dimensions through an Orange Technology orientation toward emotional wellbeing and social resilience, rather than introducing a new algorithm or autonomous clinical capability. The framework also aligns mental health innovation with SDG 3 (health and wellbeing), SDG 9 (inclusive and resilient innovation and digital infrastructure), and SDG 17 (partnerships). As an evidence-informed conceptual model, it remains a design hypothesis that requires empirical validation rather than a clinically validated intervention [10].

2. LITERATURE REVIEW

2.1. Mental Health, Access, and Community-Based Support

WHO emphasizes that mental health transformation requires more than increasing the availability of specialized clinical services. Community-based, person-centred, and rights-based approaches are central because mental health is influenced by social environments, relationships, housing, employment, education, discrimination, and access to care [11, 12]. This perspective is important for the design of AI-based mental health interventions. If a digital intervention is optimized only for symptom reduction, it may overlook broader social determinants and the protective role of community. Conversely, a support system can be designed to identify whether a user needs information, emotional validation, or connection to a trusted person or professional service. Such a system positions technology as an intermediary within a broader care network rather than as a substitute for that network [10].

2.2. Conversational AI and Mental Health Outcomes

The empirical literature provides evidence that conversational agents can be beneficial, especially for low-intensity and adjunctive support [9, 11, 13]. synthesized randomized controlled trials and reported that personalization and empathic responses were important facilitators [14–16] found statistically significant effects on depression and distress across randomized trials [17–19]. however, reported more modest effects in a larger and newer meta-analysis and observed that improvements were most evident after approximately eight weeks, with no clear benefit at three-month follow up. These findings support a nuanced interpretation including AI can help, but the magnitude and durability of benefit depend on intervention design, population, duration, comparator, and user engagement. Research on young people is particularly relevant to digital inclusion because adolescents and young adults are frequent users of digital communication [20–22]. A recent systematic review and meta-analysis found that chatbot-delivered interventions could reduce psychological distress among young people, while effects on psychological wellbeing were not consistently significant [23, 24]. Another review focused on young people reported potential effectiveness but emphasized the importance of moderators such as delivery platform and interaction mode [25]. Together, these findings caution against assuming that access to a chatbot automatically creates wellbeing. The quality of the interaction, the appropriateness of the intervention, and the surrounding support system remain decisive [26].

2.3. Empathy as a Design Requirement

Empathy in AI should be distinguished from the claim that a machine literally experiences human emotion. In a design and interaction sense, empathetic AI refers to the capacity of a system to detect or infer relevant emotional and contextual signals and to respond in ways that are perceived as understanding, respectful, supportive, and appropriate. The distinction is important because users may anthropomorphize conversational systems even when the underlying system has no subjective experience. An ethical system should therefore communicate its nature and limitations while still using interaction patterns that reduce friction and emotional invalidation.

A systematic review of empathic conversational-agent platforms found substantial variation in how empathy is implemented and evaluated [27]. Evidence also suggests that interaction features such as emotional responsiveness, structured feedback, and interaction frequency may influence adherence, although clinical-outcome relationships are less consistent [26]. Research on online mental-health communities further suggests that AI-assisted emotional-support processes can help people clarify emotions and formulate supportive responses [28]. The implication for Orange Technology is that empathy should be treated as a measurable interaction quality rather than a marketing label. Evaluation should include perceived empathy, emotional appropriateness, user agency, trust calibration, safety, and the ability to connect users with human support.

2.4. Generative AI Capability and Limitations

Generative AI expands the expressive capacity of conversational systems because large language models can produce flexible, context-sensitive language. Systematic reviews of generative AI in psychiatry and mental health report applications in assessment, psychoeducation, therapeutic support, clinical documentation, and decision support [16, 29]. Yet this flexibility introduces new risks. Models can generate inaccurate information, overconfident statements, culturally inappropriate responses, or unsafe advice. A recent systematic review found that current evaluations frequently 3 focus on isolated tasks and may fail to capture contextual reasoning, cultural sensitivity, and real-world interaction complexity [30]. Another review identified a rapid increase in mental-health GenAI research but noted limited real-world deployment and unresolved safety assurance [31]. Therefore, greater linguistic fluency should not be interpreted as clinical reliability or genuine empathy.

2.5. Digital Inclusion, Equity, and Governance

Digital mental-health innovation can unintentionally widen inequalities. Access depends on devices, connectivity, affordability, digital literacy, language, disability accessibility, and confidence in technology. SDG 9 explicitly emphasizes inclusive and resilient infrastructure and universal, affordable access to information and communication technologies [32]. For mental health, inclusion also requires culturally appropriate language and community participation. WHO's guidance on AI in health states that ethics and human rights should remain at the centre of AI design, deployment, and use [33]. This supports a governance model in which privacy, transparency, accountability, safety, fairness, and human oversight are not optional add-ons but architectural components [33].

3. RESEARCH METHOD

This study used an integrative literature-review approach to develop a conceptual framework rather than estimate a pooled clinical effect. The review examined peer-reviewed systematic reviews, meta-analyses, research syntheses, and authoritative health-policy documents, primarily from 2021–2026, identified through targeted searches of PubMed-indexed literature and WHO and UN sources. Search terms covered AI and conversational systems, including "artificial intelligence," "conversational agent," "chatbot," and "empathetic AI," as well as "mental health," "emotional wellbeing," "digital mental health," "community resilience," and "digital inclusion."

Sources were included when they addressed mental-health outcomes, empathetic interaction, accessibility and equity, safety and human oversight, community support, or ethical governance. The synthesis involved three stages: organizing evidence into technological capability, psychological and health outcomes, human interaction, and governance; comparing recurring design requirements; and mapping these requirements to six framework components [34, 35].

The six components were Inclusive Access, Emotional and Contextual Understanding, Empathetic Dialogue, Personalized Support, Safety and Human Escalation, and Community Resilience and Governance.

Evidence on accessibility, emotional context, empathy, personalization, safety, community support, and governance informed each component. The framework follows the principle that AI should augment rather than replace human and community support [36, 37]. As an integrative conceptual review rather than a registered systematic review or meta-analysis, the framework represents a structured evidence-informed synthesis intended to guide future empirical research, not establish clinical effectiveness or provide a definitive evidence map.

Table 1. Synthesis of Evidence and Implications for an Orange Technology Approach.

Evidence domain	**Key finding**	**Design implication**
Clinical and psychological outcomes	Conversational agents can reduce depression, distress, and anxiety in some studies, but effects vary and may not persist.	Use AI as an adjunctive, low-intensity support layer with outcome monitoring.
Empathy and personalization	Empathic response, personalization, emotional responsiveness, and interaction quality can improve engagement and perceived support.	Make emotional context and user agency explicit design requirements.
Generative AI	Language flexibility is promising, but accuracy, cultural competence, and safety remain inconsistent.	Use bounded knowledge, uncertainty signaling, safety policies, and human escalation.
Community and social support	Mental health is embedded in relationships and social environments; digital interventions can facilitate social connection.	Connect users to peers, family, community resources, and professionals where appropriate.
Equity and governance	Access barriers, privacy, bias, accountability, and human-rights risks can undermine benefits.	Design for low bandwidth, accessibility, language diversity, privacy, and transparent governance.

Table 1 demonstrates that the integration of Orange Technology in mental healthcare must be holistically designed by aligning five key pillars emotional support, personalization, safety, community connection, and equitable access as augmented intelligence that enhances rather than replaces professional clinical judgment. Within this framework, AI operates under strict human oversight to ensure that a clinician's empathy and intuition remain the primary drivers of intervention, while robust data encryption safeguards patient privacy and trust. Furthermore, the system incorporates proactive safety mechanisms capable of detecting emotional crisis signals in real time to immediately escalate care to emergency medical personnel, thereby establishing a care ecosystem that is not only adaptive and inclusive across all demographics, but also ethical and clinically sound.

3.1. Proposed Orange Technology–Empathetic AI Framework

The proposed framework translates the literature into a layered mental-health support ecosystem. Its core premise is that emotional wellbeing cannot be produced by an algorithm alone. Instead, AI should perform specific supportive functions while preserving user autonomy and maintaining pathways to human and community resources. The framework contains six interconnected layers: (1) inclusive access, (2) emotional and contextual understanding, (3) empathetic interaction, (4) personalized support, (5) safety and human escalation, and (6) community resilience and governance. These layers should not be interpreted as a strictly linear sequence. Instead, they form a dynamic support architecture in which some components provide foundational conditions while others operate through continuous interaction and feedback. Inclusive access establishes the conditions for equitable participation; emotional and contextual understanding informs empathetic interaction and appropriate personalization; personalized support provides low-intensity assistance; and safety and human escalation constrain these functions through risk-sensitive boundaries [38, 39]. Community resilience and governance provide the wider social and institutional context and can also feed information, resources, and human

support back into the individual-level layers. Thus, the framework represents a dynamic human-in-the-loop ecosystem rather than a one-directional technical pipeline.



Figure 1. Conceptual architecture for Orange Technology and empathetic AI in mental-health support

Figure 1 illustrates the conceptual architecture of Orange Technology and empathetic AI for mental-health support. The framework begins with inclusive access and digital equity, followed by emotional and contextual understanding that enables empathetic dialogue. These elements support personalized self-management while maintaining safety, human oversight, and escalation mechanisms. At the highest level, the framework connects individual support with community and resilience, emphasizing collaboration between AI, humans, and the wider community ecosystem.

3.2. Inclusive Access and Digital Equity

The bottom layer establishes inclusion as a prerequisite rather than an outcome. A mental-health AI service should support affordable and commonly available devices, low-bandwidth interaction, text-first alternatives, accessible interfaces, multilingual content, and culturally appropriate communication. Accessibility should include people with visual, auditory, cognitive, or motor disabilities. This layer also requires digital-literacy support and clear explanations of what the system can and cannot do. Without these measures, the users who might benefit most from additional support can be excluded by the technology itself

3.3. Emotional and Contextual Understanding

Empathetic AI requires more than sentiment classification; it should combine emotional and contextual cues with appropriate, transparent, and safety-aware responses. Empathy does not mean that AI experiences human emotions, but refers to interaction quality evaluated through perceived appropriateness, respect, cultural sensitivity, user agency, and safety [40, 41]. Because users may seek information, validation, practical solutions, or human connection, systems should consider linguistic cues, conversation history, user preferences, and explicit check-ins while minimizing unnecessary data collection. Inferred emotional states should be treated as uncertain hypotheses rather than facts. When signals are ambiguous, systems should acknowledge uncertainty, seek clarification, and allow users to correct the interpretation. Cultural and linguistic variation should also be considered because similar expressions may have different meanings across communities. This approach preserves user agency and reduces algorithmic misinterpretation [42].

3.4. Empathetic Dialogue and Emotional Support

The dialogue layer should implement several principles: acknowledge emotion before offering solutions; avoid judgemental language; reflect the user's concern without pretending to possess 5 human feelings; ask permission before giving advice when appropriate; encourage self-efficacy; and maintain conversational continuity. Empathy should also include cultural sensitivity. Expressions of distress differ across languages and communities, and literal translation may fail to capture culturally meaningful ways of communicating suffering. Evaluation should therefore involve diverse users and domain experts rather than relying only on automated language metrics [43].

3.5. Personalized Support and Self-Management

Once the system has established context, it can offer low-intensity interventions such as breathing exercises, grounding techniques, journaling prompts, psychoeducation, sleep or routine reminders, problem-solving steps, and links to validated resources. Personalization should be constrained by evidence and user

preferences. The goal is not maximal personalization but appropriate personalization: enough adaptation to make support relevant without creating intrusive profiling. The system should also distinguish between general wellbeing support and clinical treatment. When symptoms are severe, persistent, or associated with safety concerns, self-management content should not become a substitute for professional assessment.

3.6. Safety, Human Oversight, and Escalation

Safety is the central boundary condition for mental-health AI. Systems should use risk-sensitive policies, uncertainty detection, clear escalation rules, and pathways to human support. High-risk situations should be identified through predefined criteria, including self-harm or harm to others, immediate danger, severe deterioration, inability to maintain personal safety, or other crisis indicators. When risk is detected or cannot be reliably determined, the interaction should be escalated to appropriate human or emergency support, without prolonged autonomous crisis management. Escalation should provide clear handoff information, communicate system limitations, and enable human review where available. AI should not independently diagnose, prescribe medication, or make high-stakes clinical decisions outside a properly governed clinical workflow, consistent with WHO principles on human rights and accountability.

3.7. Community Resilience and Social Connection

The top layer extends the framework from individual wellbeing to community resilience. AI may indirectly support resilience by reducing barriers to information, services, and social resources through community services, peer-support groups, educational programmes, referrals, and trusted public-health information. It can also help community organizations provide multilingual educational materials and non-clinical information. These mechanisms may strengthen social connection, resource access, and coordination, although their effects remain empirical questions. Importantly, AI should not replace community relationships, but reduce barriers between individuals and the resources that support wellbeing.

4. RESULTS AND DISCUSSION

The value of empathetic AI should be evaluated at three complementary levels. At the individual level, outcomes include emotional wellbeing, distress, perceived support, self-efficacy, engagement, and help-seeking behaviour. At the interaction level, evaluation should assess perceived empathy, response appropriateness, cultural sensitivity, transparency, trust, user agency, and harmful responses. At the system level, assessment should consider accessibility, equity, referral completion, human workload, safety, privacy, and effects on community services. This approach distinguishes meaningful mental-health outcomes from simple chatbot engagement.

Second, personalization must be balanced with privacy. While contextual data may improve responses, excessive collection increases privacy and governance risks. Therefore, the framework emphasizes data minimization, user-controlled memory, transparent retention policies, local processing where feasible, and explicit consent for sensitive information.

Third, empathy is relational and context-dependent rather than a single model capability. Responses that support one user may be inappropriate for another. Participatory design involving patients, mental-health professionals, caregivers, community organizations, and digitally excluded groups can identify risks that benchmark datasets may miss. The goal is responsible AI within clearly defined boundaries rather than AI that simply appears human.

Fourth, the framework supports SDG 3, SDG 9, and SDG 17. SDG 3 is addressed through Emotional and Contextual Understanding, Empathetic Dialogue, and Personalized Support to promote wellbeing, self-management, and appropriate help-seeking. SDG 9 is supported through Inclusive Access, low-bandwidth interaction, accessibility, multilingual communication, and equitable infrastructure. SDG 17 is strengthened through Safety and Human Escalation and Community Resilience and Governance, requiring collaboration among health professionals, technology developers, governments, researchers, communities, and civil-society organizations. Thus, Orange Technology links AI innovation with inclusive infrastructure, participatory governance, community partnership, and equitable social impact.

4.1. Comparison with Conventional AI Mental-Health Approaches

This study compares conventional AI applications with the proposed approach based on five key dimensions: primary orientation, empathy and personalization, safety and human involvement, community inclusion and equity, and expected outcomes. These dimensions highlight differences in addressing functional

service delivery and the broader emotional and social needs of individuals. The comparison is presented in Table 2.

Table 2. Analytical comparison between automation-oriented mental-health AI and the Orange Technology–empathetic AI approach

Dimension	Automation-Oriented AI	Orange Technology–Empathetic AI
Primary Orientation	Efficiency and automation	Emotional wellbeing and human-centred support
Empathy & Personalization	Limited or functional personalization	Context-sensitive, empathetic, and personalized interaction
Safety & Human Role	Primarily system-driven	Human oversight, risk-sensitive escalation, and referral
Community & Equity	Access mainly determined by technology availability	Inclusive access, cultural sensitivity, and community connection
Expected Outcome	Task completion and engagement	Emotional wellbeing, self-management, help-seeking, and resilience

Table 2 shows that Orange Technology–Empathetic AI provides a more human-centered approach than conventional AI mental-health systems. While conventional approaches mainly focus on efficiency and automation, Orange Technology emphasizes emotional wellbeing, personalized support, human oversight, cultural sensitivity, and community inclusion. Therefore, its expected outcomes extend beyond task completion to include self-help, emotional resilience, and improved wellbeing.

4.2. Risks and Mitigation Priorities

Implementing empathetic AI in mental-health support requires addressing risks related to safety, inclusiveness, and effectiveness. Key concerns include information reliability, user dependence, cultural appropriateness, data protection, accessibility, and clinical coordination. These challenges highlight the need for safeguards that maintain human involvement and ensure responsible AI-based support. The corresponding risks and mitigation priorities are presented in Table 3.

Table 3. Priority risks and mitigation requirements for empathetic AI in mental-health support

Risk	Potential consequence	Mitigation priority
Hallucinated or unsafe advice	Users may act on inaccurate or inappropriate guidance.	Evidence-bounded responses, uncertainty disclosure, clinical review, escalation.
Over-reliance and anthropomorphism	Users may treat AI as a replacement for people or professionals.	Transparent identity, relational boundaries, prompts toward human connection.
Bias and cultural mismatch	Responses may be less appropriate for minority or underserved groups.	Diverse evaluation, participatory design, multilingual and cultural testing.
Privacy and secondary use	Sensitive emotional disclosures may be exposed or reused improperly.	Data minimization, consent, retention controls, security, accountability.
Digital exclusion	People with limited connectivity, devices, or access may be left behind.	Low-bandwidth design, accessible interfaces, affordable access, community channels.
Weak clinical integration	AI support may become disconnected from appropriate care.	Referral protocols, human-in-the-loop workflows, interoperable service pathways.

Table 3 highlights the main risks that may arise when implementing empathetic AI for mental-health support and the corresponding mitigation priorities. The key risks include unsafe or inaccurate advice, user

over-reliance on AI, bias and cultural mismatch, privacy concerns, digital exclusion, and weak clinical integration. To address these risks, the system should prioritize evidence-based responses, human oversight, cultural sensitivity, strong data protection, accessible design, and clear referral pathways to professional care. Overall, the table emphasizes that empathetic AI should complement not replace human and clinical support.

5. IMPLICATIONS FOR RESEARCH, PRACTICE, AND POLICY

A practical implementation pathway can be organized into six stages: participatory co-design, prototype development based on the six framework layers, safety and usability assessment, small-scale pilot testing, longitudinal evaluation, and broader implementation with continuous monitoring and governance. Future research should validate these layers through qualitative, prototype-based, controlled, and longitudinal studies, focusing on usability, empathy, personalization, accessibility, safety, wellbeing, help-seeking, engagement, and sustainability. Subgroup and implementation analyses should consider age, language, socioeconomic position, disability, cultural context, privacy, governance, referral completion, and equity. Key questions should assess whether empathetic interaction improves help-seeking, personalization works without excessive data collection, escalation reliably identifies cases requiring human support, and AI strengthens rather than replaces community support.

For practitioners, AI should function as an adjunctive support tool with clear boundaries. Mental-health professionals should establish escalation thresholds, validate content, monitor failures, and develop referral pathways, while community organizations provide culturally relevant resources and trusted human contact. This allows AI to support low-intensity needs while professionals focus on complex assessment, therapy, crisis management, and relational care. For policymakers and institutions, governance should address safety, data protection, accessibility, accountability, and evidence requirements, consistent with WHO principles on ethics, human rights, and public benefit in AI for health. Public investment should also prevent unequal access caused by expensive devices, limited connectivity, or paid services, making digital inclusion an essential part of Orange Technology.

6. LIMITATIONS

This study has several limitations. First, it is an integrative conceptual review rather than a registered systematic review or meta-analysis; therefore, the search and selection process was targeted rather than exhaustive. Second, much of the current evidence evaluates conversational agents developed before the newest generation of large language models, so findings cannot automatically be generalized to modern generative AI systems. Third, definitions and measurements of empathy vary across studies, making direct comparison difficult. Fourth, community resilience is a broader construct than individual mental-health symptom reduction, and empirical evidence directly connecting AI conversational support to community resilience remains less mature. Finally, the proposed Orange Technology framework is conceptual and requires empirical validation with users, clinicians, and community stakeholders.

7. CONCLUSION

Orange Technology offers a useful way to reframe empathetic AI from an automation technology into a humane mental-health support infrastructure. However, an important gap remains in translating evidence on conversational AI into a broader framework that addresses cultural sensitivity, community connection, inclusive access, safety, privacy, and human oversight simultaneously. Current evidence indicates that conversational AI can reduce some symptoms of depression, anxiety, and distress and improve access to low-intensity support, but effects remain heterogeneous and should not be overstated. The central challenge is therefore not to make AI appear indistinguishable from a human therapist, but to ensure that it operates within clearly defined supportive boundaries.

The novelty of this study lies in the proposed Orange Technology framework, which integrates inclusive access, emotional and contextual understanding, empathetic dialogue, personalized self-management, safety and human escalation, and community resilience within a single mental-health support architecture. This framework positions AI as a supportive, human-in-the-loop layer rather than a replacement for therapists, caregivers, peers, or communities. It also connects mental-health innovation with SDG 3, SDG 9, and SDG 17 by linking emotional wellbeing with equitable digital infrastructure, social connectedness, and collaborative

governance.

Future research should validate the framework through a staged empirical programme. Initial studies should use participatory design involving patients, clinicians, caregivers, and community stakeholders to assess usability, perceived empathy, cultural appropriateness, and user agency. Subsequent longitudinal studies should evaluate emotional and behavioural outcomes, subgroup differences, help-seeking behaviour, and sustainability. Safety audits and implementation studies should further examine escalation accuracy, privacy protection, referral pathways, human workload, and effects on community services. This staged validation can advance the framework from an evidence-informed architecture toward an empirically tested mental-health support model and clarify its contribution to a more connected, inclusive, and resilient mental-health ecosystem.

8. DECLARATIONS

8.1. About Authors

Dwi Julianingsih (DJ) https://orcid.org/0000-0002-6257-4881

Tasya Syafira (TS) -

Sheila Aulia Anjani (SA) https://orcid.org/0009-0007-9121-1151

8.2. Author Contributions

Conceptualization: DJ; Methodology: TS; Software: SA; Validation: DJ and TS; Formal Analysis: DJ and TS; Investigation: TS; Resources: TS; Data Curation: SA; Writing – Original Draft Preparation: DJ and SA; Writing – Review and Editing: DJ and SA; Visualization: SA; All authors, DJ, TS, and SA, have read and agreed to the published version of the manuscript.

8.3. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

8.4. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

8.5. Declaration of Generative AI and AI-assisted Technologies in The Writing Process

During the preparation of this work, the author(s) used ChatGPT (OpenAI) in order to support the drafting and editing of summaries, keywords and highlights, and to correct language throughout the text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

8.6. Declaration of Conflicting Interest

The authors declare that they have no conflicts of interest, known competing financial interests, or personal relationships that could have influenced the work reported in this paper.

REFERENCES

[1] M. Mardiana, O. Ferli, H. Henderi, E. Anderson, and A. Cahyani, "The role of perceived emotional support in ai mental health chatbots among generation z," *Health, Empathy, and AI Learning (HEAL)*, vol. 1, no. 2, pp. 163–173, 2026.

[2] N. Mikaeili, M. Naeim, and M. Narimani, "Reimagining mental health with artificial intelligence: early detection, personalized care, and a preventive ecosystem," *Journal of Multidisciplinary Healthcare*, pp. 7355–7373, 2025.

[3] W. H. Organization, *World mental health report: transforming mental health for all. Executive summary*. World Health Organization, 2022.

[4] A. Felix and G. D. Rembulan, "Analysis of key factors for improved customer experience, engagement, and loyalty in the e-commerce industry in indonesia," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 5, no. 2sp, pp. 196–208, 2023.

[5] W. H. Organization, *Achieving well-being: A global framework for integrating well-being into public health utilizing a health promotion approach*. World Health Organization, 2024.

[6] S. Syahyono, D. Apriliasari, N. A. Santoso, R. Rahardja, and O. Sauntos, "Empathetic ai for preventive healthcare in smart air quality monitoring," *Health, Empathy, and AI Learning (HEAL)*, vol. 1, no. 2, pp. 113–125, 2026.

[7] A. Chaturvedi, "Exploring empathy in artificial intelligence: synthesis and paths for future research," *Information Discovery and Delivery*, vol. 53, no. 3, pp. 406–424, 2025.

[8] A. Babu and A. P. Joseph, "Artificial intelligence in mental healthcare: transformative potential vs. the necessity of human interaction," *Frontiers in Psychology*, vol. 15, p. 1378904, 2024.

[9] Y. He, L. Yang, C. Qian, T. Li, Z. Su, Q. Zhang, and X. Hou, "Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials," *Journal of medical Internet research*, vol. 25, p. e43862, 2023.

[10] D. S. S. Wuisan, R. A. Sunardjo, Q. Aini, N. A. Yusuf, and U. Rahardja, "Integrating artificial intelligence in human resource management: A smartpls approach for entrepreneurial success," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 5, no. 3, pp. 334–345, 2023.

[11] H. Li, R. Zhang, Y.-C. Lee, R. E. Kraut, and D. C. Mohr, "Systematic review and meta-analysis of ai-based conversational agents for promoting mental health and well-being," *NPJ Digital Medicine*, vol. 6, no. 1, p. 236, 2023.

[12] E. Overton, D. Tobes, and A. Lee, "Preeclampsia diagnosis and management," *Best practice & research Clinical anaesthesiology*, vol. 36, no. 1, pp. 107–121, 2022.

[13] Y. Yuan, Y. Gao, H. Moen, E. Isometsä, P. Marttinen, and T. Aledavood, "Digital phenotyping with large language models to detect depressive state changes in patients," *npj Digital Medicine*, 2026.

[14] Y. M. Cho, S. Rai, L. Ungar, J. Sedoc, and S. Guntuku, "An integrative survey on mental health conversational agents to bridge computer science and medical perspectives," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 2023, pp. 11 346–11 369.

[15] R. Martin and S. Richmond, "Conversational agents for children's mental health and mental disorders: A scoping review," *Computers in Human Behavior: Artificial Humans*, vol. 1, no. 2, p. 100028, 2023.

[16] R. Sanjeewa, R. Iyer, P. Apputhurai, N. Wickramasinghe, and D. Meyer, "Empathic conversational agent platform designs and their evaluation in the context of mental health: systematic review," *JMIR Mental Health*, vol. 11, p. e58974, 2024.

[17] C. Wani, L. McCann, M. Lennon, and C. Radu, "Digital mental health interventions for adolescents in low-and middle-income countries: scoping review," *Journal of Medical Internet Research*, vol. 26, p. e51376, 2024.

[18] M. E. Taylor, M. Liu, S. Abelson, D. Eisenberg, S. K. Lipson, and S. M. Schueller, "The reach, effectiveness, adoption, implementation, and maintenance of digital mental health interventions for college students: a systematic review," *Current Psychiatry Reports*, vol. 26, no. 12, pp. 683–693, 2024.

[19] T. Tavory, "Regulating ai in mental health: ethics of care perspective," *JMIR Mental Health*, vol. 11, no. 1, p. e58493, 2024.

[20] Z. Guo, A. Lai, J. H. Thygesen, J. Farrington, T. Keen, and K. Li, "Large language models for mental health applications: systematic review," *JMIR mental health*, vol. 11, no. 1, p. e57400, 2024.

[21] O. A. Stein and A. Prost, "Exploring the societal implications of digital mental health technologies: A critical review," *SSM-mental health*, vol. 6, p. 100373, 2024.

[22] D. B. Olawade, O. Z. Wada, A. Odetayo, A. C. David-Olawade, F. Asaolu, and J. Eberhardt, "Enhancing mental health with artificial intelligence: Current trends and future prospects," *Journal of medicine, surgery, and public health*, vol. 3, p. 100099, 2024.

[23] W. Zhong, J. Luo, and H. Zhang, "The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: a systematic review and meta-analysis," *Journal of affective disorders*, vol. 356, pp. 459–469, 2024.

[24] S. Lestari, S. Watini, and D. E. Rose, "Impact of self-efficacy and work discipline on employee performance in sociopreneur initiatives," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 2, pp. 270–284, 2024.

[25] U. Rusilowati, U. Narimawati, Y. R. Wijayanti, U. Rahardja, and O. A. Al-Kamari, "Optimizing human resource planning through advanced management information systems: A technological approach," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 1, pp. 72–83, 2024.

[26] W. K. C. Leung, S. C. Lam, B. C. L. Chan, J. N. L. Chow, Y. Y. Y. Wong, F. Ng, C. C. K. Yip, and A. C. K. Chan, "Chatbot interventions for improving mental health among people in asia: a systematic

review and meta-analysis of randomised controlled trials," *BMJ Health & Care Informatics*, vol. 33, no. 1, p. e101479, 2026.
[27] U. Rahardja, "The economic impact of cryptocurrencies in indonesia," *ADI Journal on Recent Innovation (AJRI)*, vol. 4, no. 2, pp. 194–200, 2023.
[28] N. Lutfiani, N. P. L. Santoso, R. Ahsanitaqwim, U. Rahardja, and A. R. A. Zahra, "Ai-based strategies to improve resource efficiency in urban infrastructure," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 121–127, 2024.
[29] M. H. R. Chakim, Q. Aini, P. A. Sunarya, N. P. L. Santoso, D. A. R. Kusumawardhani, and U. Rahardja, "Understanding factors influencing the adoption of ai-enhanced air quality systems: A utaut perspective," in *2023 Eighth International Conference on Informatics and Computing (ICIC)*. IEEE, 2023, pp. 1–6.
[30] S. Kolding, R. M. Lundin, L. Hansen, and S. D. Østergaard, "Use of generative artificial intelligence (ai) in psychiatry and mental health care: a systematic review," *Acta Neuropsychiatrica*, vol. 37, p. e37, 2025.
[31] J. Siswanto, V. A. Goeltom, I. N. Hikam, E. A. Lisangan, and A. Fitriani, "Market trend analysis and data-based decision making in increasing business competitiveness," *Smart Human-centered Emerging Research in Machine Intelligence*, vol. 1, no. 1, pp. 1–8, 2025.
[32] E. Fransson, "Retrieval-augmented generation in mental health: A systematic review," 2026.
[33] V. El Ardeliya, J. Taylor, and J. Wolfson, "Exploration of artificial intelligence in creative fields: Generative art, music, and design," *International Journal of Cyber and IT Service Management (IJCITSM)*, vol. 4, no. 1, pp. 40–46, 2024, https://doi.org/10.34306/ijcitsm.v4i1.149.
[34] A. Arjun, "Knowledge and attitude of counsellors towards ai counselling," 2025.
[35] M. Annas and S. N. Wahab, "Data mining methods: K-means clustering algorithms," *International Journal of Cyber and IT Service Management (IJCITSM)*, vol. 3, no. 1, pp. 40–47, 2023, https://doi.org/10.34306/ijcitsm.v3i1.122.
[36] Y. Feng, Y. Hang, W. Wu, X. Song, X. Xiao, F. Dong, and Z. Qiao, "Effectiveness of ai-driven conversational agents in improving mental health among young people: systematic review and meta-analysis," *Journal of medical Internet research*, vol. 27, p. e69639, 2025.
[37] A. Mutter, M. Bücheler, A. R. Idrees, M. Domhardt, and H. Baumeister, "Persuasive design principle of social support in digital interventions targeting mental health symptoms: a systematic review and meta-analysis," *BMJ open*, vol. 14, no. 11, p. e086728, 2024.
[38] D. Shin, S. Park, E. H. Kim, S. Kim, J. Seo, and H. Hong, "Exploring the effects of ai-assisted emotional support processes in online mental health community," in *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 2022, pp. 1–7.
[39] S. Agarwal, "Goal 9: industry, innovation, and infrastructure," in *Pathways and Barriers to Achieving the UN Sustainable Development Goals*. Edward Elgar Publishing, 2025, pp. 101–113.
[40] W. Guidance, "Ethics and governance of artificial intelligence for health," *World Health Organization*, vol. 1, 2021.
[41] M. K. Suganya, "Goal 3: Good health and well-being: Ensuring healthy lives and promoting well-being for all," *Data, Design, and Development Goals: Digital Transformation for Sustainability*, p. 26, 2026.
[42] M. Malgaroli, T. D. Hull, J. M. Zech, and T. Althoff, "Natural language processing for mental health interventions: a systematic review and research framework," *Translational psychiatry*, vol. 13, no. 1, p. 309, 2023.
[43] L. Wang, Z. Wan, C. Ni, Q. Song, Y. Li, E. Clayton, B. Malin, and Z. Yin, "Applications and concerns of chatgpt and other conversational large language models in health care: systematic review," *Journal of Medical Internet Research*, vol. 26, p. e22769, 2024.