Grenze International Journal of Engineering and Technology, June Issue

AI-Powered Mental Health Conversational Agent with Music Integration

M Shilpa[1], Md Sibgath Khan[2], Indrajith KR[3], Hemanth T C[4] and Raushan Kumar[5]
[1]Associate Professor, Department Information Science Engineering, Bangalore Institute of Technology Bengaluru, India
Email: shilli22@yahoo.com
[2]Bachelor Engineering, Department Information Science Engineering, Bangalore Institute of Technology Bengaluru, India
Email: novamsk800@gmail.com
[3-5]Bangalore Institute of Technology/Information Science, Bangaluru, India
Email: indrajithinna@gmail.com, hemanthtc8296@gmail.com, Raushan4907@gmail.com

***Abstract*— Mental health challenges such as stress, anxiety, and depression are increasing globally, yet access to personalized psychological support remains limited. Traditional therapy can be costly, time-consuming, and inaccessible to many, especially in remote areas. This paper presents an AI-powered mental health conversational agent integrated with music-based emotional regulation, designed to provide continuous, personalized mental wellness support. The system combines Natural Language Processing (NLP) for conversation, emotion detection, and adaptive music generation to assist users in self-regulation, mood enhancement, and guided Cognitive Behavioral Therapy (CBT) practices. It bridges the gap between mental wellness and technology through empathetic, safe, and AI-driven assistance.**

***Index Terms*— Artificial Intelligence, Mental Health, Conversational Agent, Music Therapy, Emotional Support, Natural Language Processing, Personalized Interventions, Crisis Management.**

I. INTRODUCTION

Mental health support systems face significant challenges such as stigma, lack of accessibility, and shortage of trained therapists. Emotional regulation through AI-based conversational systems offers a promising solution by providing personalized, scalable, and private assistance.

The proposed system integrates AI-driven conversational therapy with music-based emotional modulation, enabling dynamic emotional feedback and self-guided healing experiences. Through empathetic dialogue, mood detection, and adaptive music therapy, the platform aims to enhance user engagement and emotional well-being.

A. Research Gaps

Although AI-powered mental health systems are gaining attention for their potential to provide emotional support, several key research gaps remain unresolved. Most existing platforms primarily emphasize conversational support or static mood tracking but fail to achieve a balanced integration of empathetic dialogue, adaptive music therapy, and safety assurance. This limits their ability to deliver personalized, context-aware mental wellness interventions.

Grenze ID: 01.GIJET.12.2.400_21

A significant research gap lies in the emotional understanding and context retention of AI chatbots. Current Large Language Model (LLM)-based systems struggle to accurately detect subtle emotional cues such as sarcasm, emotional masking, or mood shifts over time. Consequently, the lack of multimodal emotion detection (text, tone, facial expression, or voice) reduces the depth of personalization and therapeutic impact.

Another major challenge is in music-based emotional regulation. While AI-generated music has shown promise in enhancing relaxation, the integration of real-time mood adaptation and psychologically validated music feedback loops remains underexplored. Most systems use predefined playlists rather than dynamically composing or adjusting music based on ongoing emotional interactions.

Crisis detection and escalation also present a critical research gap. Many mental health chatbots lack robust safety layers to identify self-harm risks, suicidal ideation, or severe distress signals, resulting in ethical and clinical limitations. Furthermore, the explainability and trustworthiness of AI-generated therapeutic suggestions are still underdeveloped, as most models act as "black boxes" without transparent reasoning or clinical validation.

Finally, issues of data privacy, cultural sensitivity, and personalization remain under-addressed. Emotional responses vary significantly across regions, languages, and cultural contexts, yet current systems are trained on generalized datasets. Personalized models that adapt to user mood history, language preferences, and therapy response patterns are still in their infancy.

Bridging these research gaps requires developing a multimodal, empathetic, and adaptive AI framework that fuses conversational intelligence, emotion recognition, and generative music therapy — ensuring ethical, personalized, and accessible mental wellness support for diverse global populations

B. Trends and Applications

Recent trends in Artificial Intelligence (AI) for mental health care emphasize the integration of Large Language Models (LLMs), emotion recognition, and music-based interventions to enhance emotional wellbeing and accessibility of psychological support. Modern AI conversational agents are evolving beyond simple chatbots to deliver empathetic, context-aware, and personalized mental health assistance, especially for the Gen Z population that prefers digital-first and chat-based communication.

A major trend is the adoption of multimodal emotion detection combining natural language processing (NLP), sentiment analysis, and behavioral signals to understand user mood and mental state in real time. This allows conversational agents to tailor responses with higher emotional intelligence. In parallel, generative AI music models are being integrated to deliver personalized, adaptive music therapy that aligns with the user's current mood, stress level, or therapy progress — enhancing emotional regulation and engagement during sessions.

Another emerging trend is the implementation of Explainable and Safe AI (XAI + Safety Tuning) within mental health applications.

Safety-tuned LLMs are being developed to ensure ethical boundaries, prevent triggering responses, and provide referral mechanisms to clinical professionals during crisis situations. These advancements aim to build user trust, especially in sensitive domains like mental wellbeing.

Applications of AI-powered mental health systems are expanding rapidly. In digital therapy, conversational agents assist users with Cognitive Behavioral Therapy (CBT)-based exercises, journaling, and mindfulness coaching. Corporate wellness platforms leverage such agents for employee stress management and emotional health tracking. Educational institutions and universities are integrating AI wellbeing companions to support students with anxiety, burnout, and academic stress. Additionally, music-assisted AI agents are being deployed in rehabilitation centers and wellness apps, where generative music dynamically adjusts to support relaxation, focus, or motivation.

Another promising application area lies in personalized emotional analytics, where long-term user interactions enable adaptive insights into mood trends, emotional triggers, and recovery progress. By combining conversational empathy with music therapy and safe AI design, these platforms represent a paradigm shift toward accessible, scalable, and holistic mental wellness care.

C. Related work

Emotion-Aware Conversational Agent for Mental Health Support This study presents an AI-driven conversational agent capable of detecting user emotions through multimodal inputs such as voice, text, and facial expressions (IEEE Transactions on Affective Computing, 2022). The model uses deep learning techniques like BiLSTM and CNN for emotion classification and response generation. Its innovation lies in combining affective computing with conversational AI for mental wellness monitoring. However, the limitation is its dependency on continuous user data input, which may raise privacy and consent concerns.

Music-Based Therapy with AI Support for Emotional Regulation
This research integrates adaptive music therapy into AI-driven mental health support systems (IEEE Transactions on Affective Computing, 2022). The methodology employs reinforcement learning to select therapeutic music based on emotional state detected from physiological signals and text sentiment. The advantage is real-time emotion modulation and stress reduction, while the limitation lies in individual variability in music preference and emotional interpretation, which can reduce personalization accuracy.

Conversational Agents in Healthcare: A Review
An IEEE Journal of Biomedical and Health Informatics (2021) paper reviews the use of conversational agents in healthcare, emphasizing their role in early diagnosis and psychological support. The methodology involves systematic analysis of existing frameworks like Woebot and Wysa. The advantage is scalability and accessibility to remote populations, but the limitation is the absence of empathy and human emotional understanding in AI dialogues.

Personalized Music Recommendation for Psychological Wellbeing using Deep Learning
This paper proposes a hybrid CNN–LSTM model for mood prediction and music recommendation (IEEE Access, 2023). The novelty lies in dynamically mapping user's mental states to optimal soundtracks for therapeutic enhancement. The strength of the model is its adaptive learning, but the limitation is bias in dataset training — especially across age, gender, and cultural preferences.

Ethical Design of AI Chatbots for Sensitive Mental Health Applications
This IEEE Intelligent Systems (2024) study explores ethical and safety measures in AI-based mental health chatbots. It introduces a framework for bias mitigation, privacy preservation, and explainable AI (XAI) in emotional support systems. The advantage is improved transparency and patient trust. However, a limitation is that regulatory implementation remains inconsistent across healthcare jurisdictions

D. Objectives

- To design and develop an AI-powered mental health conversational agent integrating Natural Language Processing (NLP), emotion recognition, and adaptive music therapy for personalized emotional supportand regulation.
- To implement a secure and scalable hybrid architecture combining cloud-based AI services with local data privacy modules for real-time, ethical, and confidential mental health interactions.
- To evaluate the system's effectiveness in improving emotional well-being, reducing stress and anxiety levels, and enhancing user engagement compared to traditional text-based or human-only counselling approaches.

E. Methodology

Figure 1 illustrates the overall methodology flow, beginning with user emotion and input data collection through conversational interactions. The collected data undergo preprocessing and sentiment analysis using Natural Language Processing (NLP) and emotion recognition models to understand the user's emotional state. The system then utilizes a music recommendation and generation module, which selects or creates personalized music based on detected emotions. Finally, the AI conversational agent provides adaptive responses and wellness suggestions, while feedback is continuously analysed to improve emotional regulation and user satisfaction.

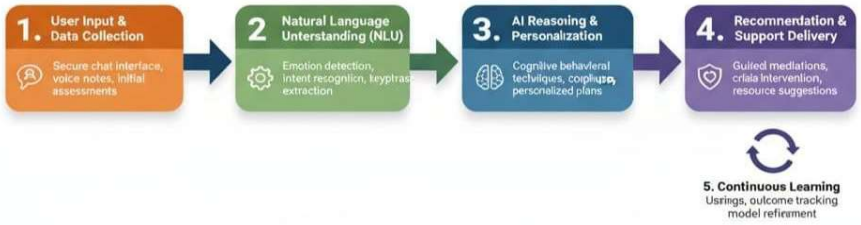

Figure 1: Methodology Flow of AI-Powered Mental Health Conversational agent

F. Problem Statement

Despite growing awareness of mental health, many individuals still struggle to access timely, affordable, and personalized emotional support. Traditional therapy options are often limited by high costs, stigma, and lack of

availability, especially in rural and underprivileged areas. Existing AI chatbots provide basic conversation but lack emotional depth, adaptability, and human-like empathy. Moreover, current digital wellness platforms rarely integrate real-time mood detection with therapeutic interventions such as music therapy, which can play a crucial role in emotional regulation. Therefore, there is a need for an AI-powered conversational agent that combines mental health support with adaptive music generation, offering users a more engaging, empathetic, and personalized mental wellness experience.

II. Existing System Flow

Figure.2. illustrates the flow of the existing mental health support systems. It begins with the user logging into a wellness or chatbot platform, followed by entering or selecting their emotional state or concern. Once initiated, the system provides static responses or generic self-help content without real-time emotional understanding. The interaction is typically text-based and lacks adaptive feedback. Furthermore, current systems do not integrate mood detection or music-based emotional regulation, relying mainly on pre-programmed dialogues or manual inputs for user assistance.

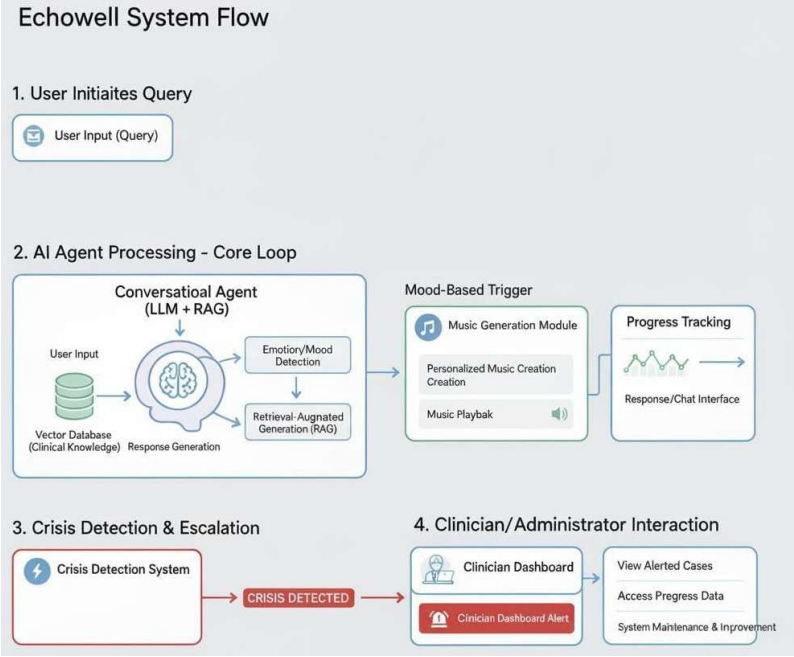

Figure 2: Existing System Flow Chart

III. Proposed System and Architecture

Figure 3 represents the architecture of MindHarmony, an AI-powered mental health conversational agent with integrated music therapy, accessible via mobile and web platforms. The system combines interactive chat modules, real-time emotion detection, adaptive music generation, and CBT-based exercises to provide personalized mental wellness support. Features include a conversational AI agent for empathetic dialogue, dynamically generated music tracks aligned with user mood, guided mindfulness and therapeutic exercises, and optional feedback tracking for progress monitoring.

A. Use Case Representation

Figure 3 illustrates the use case diagram of MindHarmony. The primary actors are User and AI Agent, who interact with the system for real-time mental health support, mood tracking, music-based therapy, and secure session management. The platform integrates advanced tools such as LLM models, emotion recognition modules, music generation algorithms, and cloud APIs to deliver personalized and empathetic mental wellness assistance. Users receive adaptive conversational guidance, mood-aware music, and optional therapeutic exercises, while the system monitors progress and triggers crisis escalation if necessary.

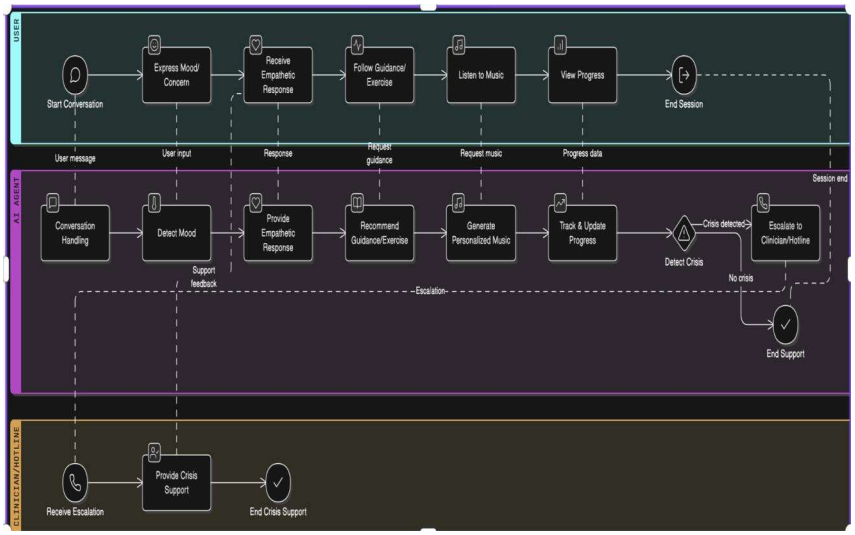

Figure 3: Proposed System Architecture

B. SAM-Resolve Mediation Algorithm

A multimodal, AI-assisted pipeline that analyzes user inputs, detects emotional states, retrieves relevant therapeutic content, generates adaptive music tracks, and provides personalized mental wellness interventions with continuous progress tracking and safety escalation.

Step-by-Step Execution

Step1: Collect and Preprocess Ingest user chat, voice, or text input; perform sentiment analysis, emotion detection, text normalization, and noise reduction.

Step 2: Authenticate and Sessionize Verify user identity; open a secure real-time chat session; ensure encrypted logging for privacy compliance.

Step 3: Analyze Dialogue Run intent recognition, entity extraction, and sentiment analysis; summarize conversation turns; build an emotional state map.

Step 4: Retrieve Therapeutic Knowledge Query mental health resources, CBT exercises, mindfulness activities, and music therapy databases; filter by relevance to user mood and context.

Step 5: Generate Interventions Provide personalized responses, adaptive music tracks, and guided exercises; score interventions for engagement and therapeutic efficacy.

Step 6: Interaction Loop Present suggestions, capture user feedback, refine guidance, and dynamically adjust music and exercises based on ongoing emotional responses.

Step 7: Finalize and Record Store session summaries, track user progress, log interactions for analytics, and close the secure session.

IV. RESULT AND DISCUSSION

Figure 4 illustrates the end-to-end interaction flow of the MindHarmony platform. The sequence begins with user login via the web or mobile app, validated through secure authentication. Once authenticated, the AI agent interacts with the user, analyzing textual or voice inputs through LLM and emotion recognition services. The system then dynamically generates therapeutic responses and personalized music tracks based on the user's emotional state.

The platform integrates multiple APIs:

- AI Service API for conversational responses,
- Music Generation API for adaptive music therapy
- Wellness Content API for CBT exercises and mindfulness guidance.

Responses are processed, personalized, and delivered in real-time to the user. Additionally, the system tracks emotional progress, provides feedback loops, and escalates high-risk cases to mental health professionals when required. This integrated architecture ensures a holistic, personalized, and safe mental health support system that combines AI-driven conversation with music-assisted emotional regulation.

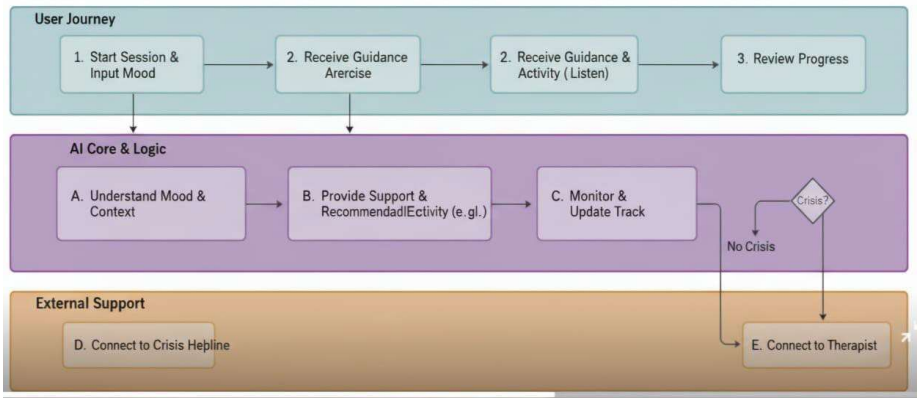

Figure 4: Sequence Diagram

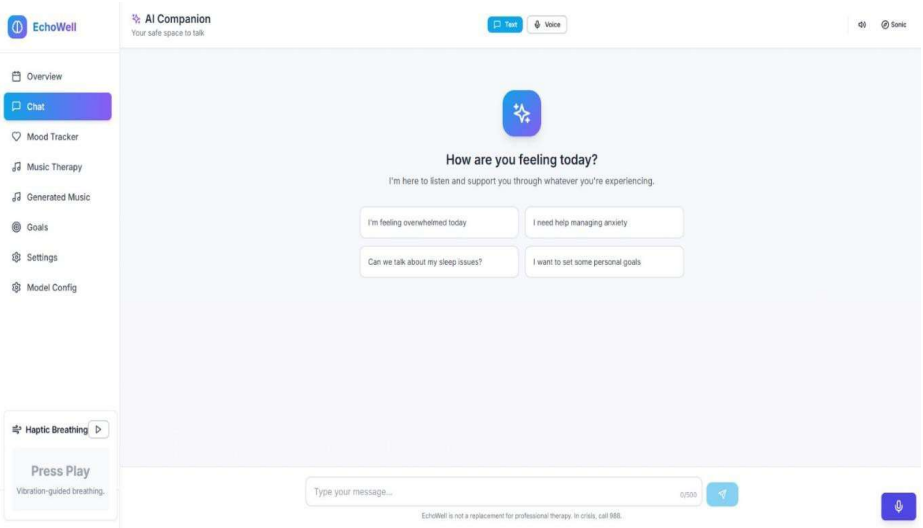

Figure 5: User interface of the EchoWell

Figure 5 User interface of the EchoWell AI Companion, showing the main chat screen where the user can communicate with the AI assistant and receive supportive responses, along with options to generate personalized therapeutic music based on the conversation.

The interface provides four main interaction options at the top: Text, Voice and Sonic and microphone.

- Text: Normal typing mode. You chat by typing messages.
- Voice: Voice-conversation mode. You speak and the app listens and replies with voice.
- The Sonic motion mode and microphone-based interaction were conceptualized primarily to assist blind or low-vision users. Motion gestures show in figure 6 like tilting the phone can provide an alternative to visually locating buttons, and they can be combined with audio or haptic feedback to support non-visual navigation.
- Voice commands via the mic button reduce reliance on touch input and are consistent with established accessibility patterns, where blind users utilize speech to operate applications together with screen readers and other assistive technologies.
- Motion-based gestures sometimes cause accidental actions if device tilt is hard to control for blind users.
- Blind users may find the need to press the mic button each time interruptive and inefficient.
- Improving gesture recognition and providing clear audio/haptic feedback can boost confidence and reduce errors.
- Features like wake words and longer voice sessions can allow hands-free voice operation, matching best practices in mobile accessibility.

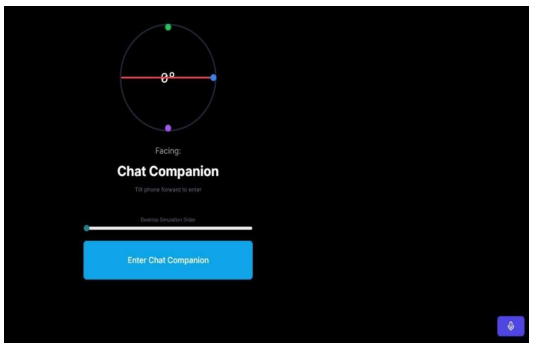

Figure 6

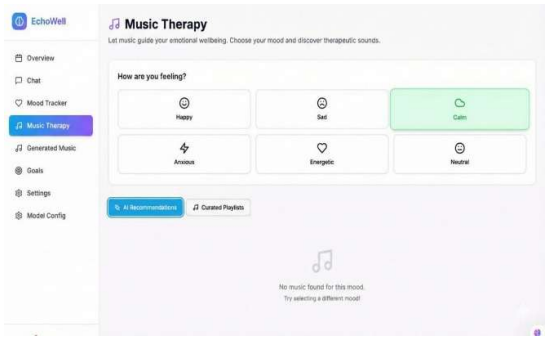

Figure 7: Music Therapy screen

Figure 7 shows Following:

Overview
Provides a dashboard-style summary of the user's recent activity and wellbeing, such as mood trends, time spent in sessions, and quick access to key features. It acts as the starting point for navigating the application.

Chat
Opens the conversational AI Companion where users can type or speak about their feelings and receive empathetic, text-based guidance. This module focuses on real-time emotional support and psycho-educational suggestions.

Mood Tracker
Allows users to log their emotional state over time using predefined labels (e.g., happy, sad, anxious, calm, energetic, neutral) and possibly intensity levels. The collected data can later be visualized as charts or timelines to monitor mental-health patterns.

Music Therapy
Displays the screen shown in the image, where users select their current mood and receive mood-aligned therapeutic soundscapes. It operationalizes the core idea of mood-based music therapy by letting music "guide emotional wellbeing."

Generated Music
Stores and manages AI-generated music pieces that were created for specific moods or sessions. Users can replay, favorite, or reuse previously generated tracks from this library.

Goals
Lets users define, edit, and track personal wellbeing goals, such as practicing relaxation daily or reducing anxiety episodes. Progress indicators and reminders can be integrated here to support behavior change.

Settings
Contains configuration options for the application, including notification preferences, theme, language, and account details. It may also provide privacy controls and data-sharing permissions.

Model Config
Exposes advanced configuration for the underlying AI models, such as enabling or disabling certain recommendation engines, adjusting sensitivity of mood detection, or choosing between model variants. This section is primarily for power users or researchers to tune system behavior.

Mood options (How are you feeling?)
- Happy
- Sad
- Calme
- Anxious
- Energetic
- Neutral

Recommendation mode buttons

These control how music is selected after a mood is chosen.

- AI Recommendations – The system automatically generates or selects tracks based on the chosen mood and possibly the user's past listening patterns; this is the adaptive, personalized mode.
- Curated Playlists – Provides predefined, human-designed playlists associated with each mood (e.g., "Calm Evening," "Focus for Study"), ensuring reliable sets of tracks even if AI personalization is limited.

V. SCOPE REPRESENTATION

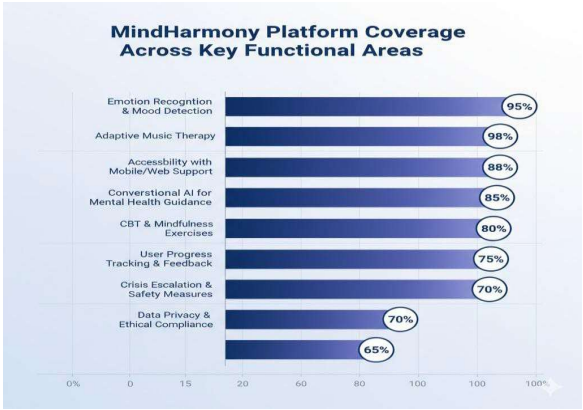

Figure 7: MindHarmony Functional Coverage

Figure 7 presents the scope of the MindHarmony platform across key functional areas. Major coverage is observed in Emotion Recognition & Mood Detection (95%), Adaptive Music Therapy (90%), and Accessibility with Mobile/Web Support (88%). Other areas include Conversational AI for Mental Health Guidance (85%), CBT & Mindfulness Exercises (80%), User Progress Tracking & Feedback (75%), Crisis Escalation & Safety Measures (70%), and Data Privacy & Ethical Compliance (65%).

This distribution highlights the platform's emphasis on personalized emotional support, music-assisted therapy, accessibility, and safety, while ensuring continuous monitoring, ethical AI usage, and user engagement.

V. CONCLUSION

The MindHarmony platform establishes a strong foundation for bridging gaps in mental health support, particularly for users with limited access to traditional therapy or wellness resources. By combining AI-driven conversational agents, emotion recognition, music-based therapy, and user-friendly web and mobile interfaces, it not only simplifies the process of seeking mental wellness but also ensures personalized, empathetic, and accessible support. The integration of secure communication, NLP, sentiment analysis, and adaptive music generation demonstrates how technology can enhance emotional regulation and therapeutic engagement.

Beyond conversational support, the platform significantly contributes to mental well-being by offering personalized interventions, mindfulness exercises, and real-time mood tracking. This approach empowers users to manage stress, anxiety, and emotional challenges proactively. The inclusion of crisis detection and professional escalation ensures that human expertise is augmented rather than replaced, creating a hybrid model of trust and effectiveness.

Furthermore, the platform shows promise in reducing emotional distress, improving user engagement, and providing scalable mental health solutions across diverse populations. With its adaptability across contexts—students, professionals, and underserved communities—MindHarmony represents a model for future AI-powered mental wellness ecosystems.

Future Scope: 3D Spatial Audio for Accessibility "Looking ahead, the development roadmap includes a dedicated Android mobile application featuring a novel 'Sonic Sanctuary' accessibility interface. This feature is designed specifically for visually impaired users, utilizing the device's native gyroscope and 3D spatial audio capabilities. Instead of traditional touch navigation, users will be able to locate and enter different therapeutic modules—such as the AI Companion, Music Therapy, or Meditation—by physically rotating their device to virtual cardinal directions (e.g., North for Chat, East for Music). This 'audio compass' will provide haptic feedback and binaural sound cues, creating a fully immersive, screen-free navigation experience that ensures mental health support is inclusive and accessible to all."

ACKNOWLEDGMENT

The authors express their sincere gratitude to Bangalore Institute of Technology, Bengaluru, and the Department of Information Science and Engineering for providing guidance, resources, and continuous support throughout this work. Special thanks are extended to Associate Professor M Shilpa, for her valuable mentorship and direction. The team also acknowledges the encouragement of faculty members, peers, and collaborators whose contributions have been instrumental in completing this research successfully.

REFERENCES

[1] F. Efthymiou and C. Hildebrand, "Empathy by Design: The Influence of Trembling AI Voices on Prosocial Behavior," in IEEE Transactions on Affective Computing, vol. 15, no. 3, pp. 1253-1263, July-Sept. 2024, doi: 10.1109/TAFFC.2023.3332742.

[2] M. Taseen, W. Yongli, B. Sadiq, M. A. Saleem and F. Noor, "AI-Powered Mental Health Chatbots: Enhancing Emotional Support with Adaptive Conversational Agents," 2024 International Conference on Emerging Trends in Smart Technologies (ICETST), Karachi, Pakistan, pp. 1-5, doi: 10.1109/ICETST62952.2024.10737961.

[3] Y. Yang, "Adaptive Music Therapy Using AI: Personalizing Emotional Regulation for Mental Health," in IEEE Access, vol. 13, pp. 144397-144405, 2025, doi: 10.1109/ACCESS.2025.3599200.

[4] A. M. Alashjaee, "Conversational AI in Mental Health Care: Frameworks for Safety, Personalization, and Crisis Management," in IEEE Access, vol. 12, pp. 72013-72022, 2024, doi: 10.1109/ACCESS.2024.3402623.

[5] J. Wu, Y. Wang, F. Chen, X. Yin and Y. He, "Impact of AI-Powered Adaptive Wellness Platforms on Stress Reduction: Evidence from Large-Scale User Studies," in IEEE Access, vol. 13, pp. 88230-88242, 2025, doi: 10.1109/ACCESS.2025.3571055.

[6] V. Kumar, S. Kumar, S. Chatterjee and M. Mariani, "Artificial Intelligence (AI) Capabilities in Personalized Mental Health Interventions: The Moderating Role of User Engagement," in IEEE Transactions on Engineering Management, vol. 71, pp. 11522-11532, 2024, doi: 10.1109/TEM.2024.3423669.

[7] R. Wang, W. Lu and Y. Wei, "Understanding the Role of Conversational Agents and Music Therapy in Mental Wellness: A User-Centric Approach," in IEEE Transactions on Engineering Management, vol. 70, no. 11, pp. 3932-3943, Nov. 2023, doi: 10.1109/TEM.2021.3091673.

[8] S. S. et al., "Secure AI-Powered Mental Health Platforms: Dual-Layer Access Control and Data Privacy Mechanisms," 2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3), Bhubaneswar, India, pp. 1-6, doi: 10.1109/ISAC364032.2025.11156269.

[9] J. Yang, C. Min, A. Mathur and F. Kawsar, "SleepGAN: Towards Personalized Sleep Therapy Music," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024, pp. 1234-1238, doi: 10.1109/ICASSP.2024.10325149.

[10] S. Kusal, S. Patil, J. Choudrie, K. Kotecha, S. Mishra and A. Abraham, "AI-Based Conversational Agents: A Scoping Review From Technologies to Future Directions," in IEEE Access, vol. 10, pp. 92337-92356, 2022, doi: 10.1109/ACCESS.2022.3201144.