Preksha K, 2025, 13:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

MANAS: Personalized AI Mental Health Chatbot with Crisis Detection and Reports

Preksha K[1], Pragna Y S[2], Raksha H M[3], Smriti Ravichandra[4], Dr. Rajashekar M B[5]

[1,2,3,4]Department of Computer Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru, Affiliated to VTU, Belagavi, Karnataka, India

[5]Associate Professor, Department of Computer Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru, Affiliated to VTU, Belagavi, Karnataka, India

Abstract- Traditional mental health interventions struggle with accessibility, continuous monitoring, and personalized care, while current digital systems lack real-time crisis detection. The MANAS (Mental Health Assistant System) platform is an AI-supported ecosystem that addresses this gap, providing personalized psychological guidance and continuous emotional tracking. MANAS utilizes a large language model (LLM) for empathetic therapeutic dialogue and incorporates a Trait Detector to analyse user input for high-risk indicators, such as explicit suicidal intent. Safety is ensured via a "human-in-the-loop" Alert Management workflow, which automatically dispatches urgent, encrypted email notifications to registered guardians upon crisis detection. The platform also integrates personality-based therapist matching and a Report Generator for downloadable PDF session summaries. Successfully implemented and tested, MANAS is a reliable, secure, and impactful solution that significantly enhances the accessibility and safety of digital mental healthcare.

Keywords: Mental Health, AI Chatbot, Crisis Detection, Personalized Support, LLM, Telehealth.

I. INTRODUCTION

The rising global prevalence of mental health challenges necessitates a paradigm shift towards accessible, proactive, and technology-driven support systems. Traditional psychological interventions are often constrained by geographical barriers, social stigma, high costs, and intermittent support cycles. This common setup can leave individuals without assistance during moments of urgent emotional need. Moreover, conventional digital platforms are often limited, incapable of deep emotional understanding, contextual interpretation, or personalized conversational flow. They rely on manual documentation and fail to actively monitor a patient's emotional state or intervene when early signs of distress emerge.

Problem Statement

Existing mental health support systems suffer from a critical deficiency in integrating personalized care with real-time crisis intervention protocols. Current conversational models frequently provide generic advice, lacking the nuanced interaction required for impactful therapeutic support tailored to individual personality and evolving emotional traits. The absence of automated, real-time crisis detection and guardian alert mechanisms significantly limits the ability to initiate timely human intervention in critical, high-risk situations involving self-harm or suicidal ideation. Finally, the inability of many existing platforms to generate detailed, structured medical reports hinders the continuity of care and comprehensive assessment by healthcare professionals.

Proposed System

The MANAS platform is designed as a multi-component, AI-driven mental health support system that addresses the shortcomings of existing solutions. Its architecture integrates advanced emotional analysis, real-time monitoring, therapeutic interaction, and professional oversight.

- **Personalized AI Dialogue**

MANAS uses a large language model capable of generating empathetic, context-aware responses. By incorporating the user's MBTI personality type, the system adapts its communication style and tone, enhancing comfort, relatability, and user engagement.

- **Emotional Trait Detection**

A real-time trait detection engine analyses user input to identify psychological markers such as anxiety, stress, depressive patterns, and trauma-related cues. This component enables early recognition of emotional fluctuations and long-term monitoring of a user's mental state.

- **Crisis Alert Mechanism**

To ensure user safety, the system includes a multi-level crisis detection workflow. When the Trait Detector identifies suicidal intent or harmful expressions, the system flags the message for administrative review. After verification, automated email alerts are sent to registered guardians or emergency contacts, ensuring timely human intervention.

- **Comprehensive Patient Management**

The platform stores user information, emotional history, personality profiles, and conversation logs in a secure SQLite database. An integrated Admin Module allows therapists and authorized personnel to review user progress, monitor emotional trends, and analyse behavioural patterns.

- **Automated Report Generation**

MANAS generates detailed PDF session reports containing psychological summaries, chat transcripts, and detected emotional traits. These documents support therapist evaluation, clinical review, and long-term mental health tracking.

Together, these components form a unified mental with structured professional oversight, enhancing safety, personalization, and therapeutic continuity.

II. RELATED WORK

Recent developments in artificial intelligence have encouraged the emergence of digital mental-health tools that aim to support individuals through automated conversation, emotional analysis, and continuous monitoring. Much of the existing work in this area highlights the growing potential of AI-powered systems to supplement traditional therapy, especially in situations where professional help is not immediately available. Researchers have shown that large language models (LLMs) are capable of engaging users in meaningful emotional dialogue, recognizing subtle verbal cues, and adapting responses based on conversational context. These abilities make them valuable tools for early mental-health screening and emotional support.

Several studies emphasize the importance of proactive detection in mental-health interventions. Instead of waiting for users to explicitly express distress, modern systems increasingly focus on identifying patterns such as tone shifts, negative sentiment, repeated expressions of hopelessness, or behavioural withdrawal during conversation. The ability to track such signals across time enables a more holistic understanding of a user's emotional state. This motivates the inclusion of components such as MANAS's Trait Detector, which analyses user text for emerging psychological markers.

However, the literature also warns about the ethical challenges associated with AI-based mental health tools. While they can offer empathetic responses, LLMs may unintentionally provide statements that lack professional sensitivity or fall outside ethical boundaries. This concern has led many researchers to advocate for hybrid models where automated support is combined with human supervision. Such observations strongly influenced MANAS's design, particularly its human-in-the-loop crisis-alert workflow, which ensures that no high-risk alert is communicated to guardians without professional review.

Another theme in related research is the importance of user trust and transparency. Mental-health practitioners prefer systems that provide interpretable outcomes, allowing them to understand why certain emotional indicators were detected or why a crisis flag was raised. Studies also highlight the value of converting unstructured user conversations into structured insights, enabling therapists to review progress more efficiently. These findings align with MANAS's reporting and administrative modules, which generate clear PDF summaries and visual trends that can support clinical evaluation.

Overall, the existing body of work underlines the need for systems that balance AI-based intelligence with responsible oversight, personalization, and meaningful documentation. MANAS builds on these lessons by combining emotional monitoring, context-aware dialogue, crisis detection, and therapist-friendly tools into a unified, user-centered platform.

III. SYSTEM ARCHITECTURE AND METHODOLOGY

The MANAS platform is built on a clear, modular architecture that allows each part of the system to function independently while still working together as a unified mental-health support tool. Instead of relying on a single block of code, MANAS follows a layered structure that separates the user interface, system logic, and data management. This approach improves flexibility, makes updates easier, and ensures that the system remains stable as new features are added.

Overall Architectural Design
The architecture can be understood through three main layers:

- **Presentation Layer**

This is the part of MANAS that users interact with. The patient, therapist, and admin dashboards are built using Streamlit, which allows the platform to display information clearly and update the interface instantly based on user input. Streamlit also supports simple navigation, making the system accessible even for users unfamiliar with technology.

- **Application Layer**

This layer acts as the "brain" of the platform. It processes all user messages, handles authentication, triggers AI responses, and manages crisis alerts. The Application Layer hosts the AI Orchestrator, the Trait Detector, the scheduling system, and all logic required to interpret user behaviour. It is responsible for communicating with the OpenAI API, interpreting emotional patterns, and ensuring the right actions are taken at the right time.

- **Data Layer**

All important information such as user profiles, emotional traits, session transcripts, therapist schedules, and crisis alerts is stored securely in an SQLite database. Using SQLite makes the system lightweight yet reliable, ensuring fast access to user records while maintaining strong data integrity. The Data Layer ensures that the system can be used across multiple sessions and still retain a full history of user interactions.

- **Core Functional Modules**

To carry out its features smoothly, MANAS is divided into several smaller modules, each responsible for a specific task.

AI Orchestrator:
This component handles all communication with the large language model. It constructs prompts based on the user's inputs, emotional patterns, and personality type. The Orchestrator ensures that the chatbot responds in an empathetic manner rather than sounding generic or robotic.

Trait Detector:
The Trait Detector continuously analyses the user's messages to identify emotional indicators. It looks for keywords, tone, and patterns that may suggest stress, anxiety, depression, or other concerns. When it notices potential danger signals like self-harm language it initiates the crisis workflow.

Alert Manager:
Instead of sending automated warnings directly, MANAS uses a human-supervised alert flow. If high-risk behaviour is detected, the Alert Manager logs it in the system as a "pending alert." This is reviewed by an administrator, who decides whether a guardian should be notified. This ensures safety without unnecessary escalation.

Patient and Therapist Management Module:
This module keeps all user information organized. It stores personal data, personality assessments, emotional histories, and therapist details. It also supports therapist matching and appointment handling, making the platform suitable for structured mental-health support.

Report Generator:
For clinical and academic purposes, MANAS generates downloadable PDF reports containing chat summaries, emotional trends, and detected psychological traits. These reports help professionals understand long-term user behaviour and allow individuals to track their own progress.

IV. IMPLEMENTATION DETAILS

The implementation of the MANAS system follows a structured workflow that ensures reliable emotional monitoring, safe crisis detection, efficient therapist matching, and smooth user interaction. Each component was developed using Python and integrated to function as a cohesive mental-health platform.

Crisis Detection and Alert Workflow
The crisis detection mechanism was designed with extreme care because it handles sensitive and high-risk scenarios. The workflow involves several sequential steps to guarantee accuracy, safety, and human supervision.

Step 1: User Input Processing
Every time the user enters a message, the system captures the text and forwards it to the Trait Detector. This happens instantly so that emotional cues are never missed.

Step 2: LLM-Based Emotional Screening
The Trait Detector uses a prompt specifically engineered to identify high-risk expressions, including phrases indicating suicidal thoughts such as:

- "I want to die"
- "I can't continue anymore"
- "There is no point in living"

The LLM analyses not only keywords but also emotional tone and context, ensuring higher detection accuracy.

Step 3: Crisis Flag Generation
If the model detects suicidal intent or severe emotional distress, the system does not take immediate action. Instead, it logs the incident by creating a pending alert in the database. This preserves user safety while avoiding false alarms.

Step 4: Administrator Review
Pending alerts appear on the Admin Dashboard in real time. A licensed professional or designated admin reviews the flagged message to confirm whether it genuinely represents a crisis situation.

Step 5: Guardian Notification
If the admin approves the alert:

- The system triggers the SMTP email module
- A pre-formatted crisis message is generated
- Notifications are sent to the registered guardian or parent email addresses

This human-supervised approach ensures responsibility, accuracy, and ethical escalation.

Therapist Matching and Scheduling
Beyond crisis support, MANAS also allows patients to connect with mental-health professionals based on compatibility and availability.

Step 1: User Profile and Personality Extraction
Upon registration, each user submits their MBTI personality type and basic demographic details. These parameters are stored securely in the SQLite database.
Step 2: Therapist Database Matching Therapist profiles include:

- Name
- Specialization
- MBTI or preferred client personality types
- Available time slots

The system compares the user's MBTI type with the therapist's profile to find the best possible match.

Step 3: Displaying Matching Options
Users are presented with a list of therapists that best align with their personality and needs. This improves comfort and rapport in later sessions.

Step 4: Appointment Scheduling
When a user selects a therapist and slot, the system performs validation:

- The slot must be in the future
- The therapist must not already be booked

- No two users can occupy the same slot for a single therapist

If valid, the appointment is stored in the appointments table.

Step 5: Session Management
Users can return to the dashboard to:

- Reschedule an appointment
- Cancel a session
- View upcoming bookings

Therapists can also review their schedules through their admin interface.

V. RESULTS AND DISCUSSION

The MANAS platform was evaluated across multiple functional components to understand its effectiveness as a digital mental-health system and to identify how it differs from commonly available AI wellness tools. The assessment covered areas such as conversational quality, emotional-state identification, crisis-response accuracy, therapist coordination, and automated documentation. Findings indicate that MANAS delivers a more comprehensive and organized support framework compared to platforms that focus mainly on casual dialogue. During testing, the system also showed strong continuity in tracking emotional cues across sessions, which strengthened the consistency of user support. In addition, MANAS provided clearer oversight features for administrators and therapists, making it more suitable for structured mental-health care settings.

Functional Evaluation

- **Conversational Responsiveness and Personalization**

During testing, the chatbot module provided responses that were emotionally sensitive, coherent, and adapted to the user's personality type. Unlike commercial chatbots such as Woebot or Wysa which rely heavily on predefined dialogue templates MANAS uses an LLM-based orchestration system that adjusts language tone, empathy level, and guidance style according to the user's MBTI profile. This personalized interaction produced higher engagement and a more natural conversational flow. Users reported that the system felt more like a companion that understood their emotional tendencies, whereas existing chatbots often feel repetitive or scripted. This demonstrates MANAS's advantage in tailoring therapeutic communication to individual needs rather than using a one-size-fits-all conversational pattern.

- **Crisis Detection Reliability**

A major gap in many current mental-health chatbots is the lack of reliable crisis escalation. Most existing systems either provide general hotline information or perform shallow keyword checks without human supervision. MANAS, however, combines contextual emotional analysis with a human-review stage before notifying guardians.

Testing showed that the system correctly identified high-risk expressions, analysed their emotional tone, and generated a pending alert for administrator verification. This significantly reduces false alarms an issue seen in platforms that send automated warnings without validation. The inclusion of professional review makes MANAS safer and more ethically aligned compared to autonomous chatbot interventions.

- **Therapist Matching and Administrative Functionality**

Another distinguishing component of MANAS is the integration of therapist matching and scheduling. Existing mental-health chatbots rarely offer direct professional involvement; they function mostly as self-help tools. MANAS tested successfully in identifying suitable therapist matches based on personality compatibility and availability. Appointment bookings, modifications, and therapist dashboards worked smoothly, confirming that the system can support hybrid digital-therapy models.

- **Report Generation and Data Structuring**

The Report Generator produced clear and well-organized PDF summaries that included emotional-trait observations, conversation transcripts, and timestamps. Such structured documentation is typically absent in commercial AI chatbots, which

often store conversations without formatting or analytical interpretation.

Therapists who evaluated the reports noted that the summaries could be used for session planning, progress monitoring, and long-term behavioural assessment an added layer of professionalism that enhances the usefulness of MANAS in clinical contexts.

- **Observed System Trends**
- Emotional-Trait Trends and User Behaviour The system's emotional-analytics module tracked variations in stress, anxiety, and mood across sessions. These trends provided insight into how user emotions fluctuated over time. In contrast, many existing chatbots require manual mood check-ins and cannot derive psychological patterns automatically from conversation data.

Even with a limited number of sample users, MANAS demonstrated an ability to detect recurring emotional signals and shifts in tone, indicating potential for large-scale emotional monitoring in institutional or clinical settings.

- User Engagement and Trust Factors Participants expressed higher trust in MANAS compared to generic LLM-based chatbots because of two main factors:
    1. Transparency of stored emotional history, and
    2. Presence of therapist oversight

This addresses a common concern in many existing AI tools where users are unsure how their data is used or whether a professional monitors their emotional state. The hybrid human-AI structure of MANAS created a sense of security without reducing the conversational comfort offered by AI.

- **System Stability and Scalability**

The database and scheduling modules were tested under multiple simultaneous sessions and performed consistently without data loss or performance degradation. Many existing chatbots store interactions temporarily or without structured indexing, limiting their usefulness for therapy continuity. MANAS's persistent data storage ensures that the system can grow to support larger user groups while maintaining reliability.

CONCLUSION

The MANAS platform represents a comprehensive effort to bridge the growing gap between traditional mental-health services and the increasing demand for accessible, personalized psychological support. Through its integration of AI-driven conversational capabilities, emotional trait detection, crisis monitoring, and structured data management, MANAS demonstrates the potential of intelligent digital systems to enhance early mental-health intervention. The design choices particularly the combination of automated emotional screening with human-supervised crisis verification reflect a balanced approach that prioritizes both efficiency and ethical responsibility.

The system's performance during functional testing confirms that MANAS can reliably analyse user sentiment, detect high-risk cues, and facilitate therapist involvement through detailed PDF reports and administrative tools. These findings highlight the platform's value not merely as a conversational assistant but as a structured mental-health ecosystem capable of supporting long-term psychological care. By enabling continuous emotional monitoring, the system helps identify patterns that might otherwise go unnoticed in conventional therapy sessions. Furthermore, its emphasis on personalization through personality-based response generation contributes to user comfort and sustained engagement two critical factors in digital mental-health adoption.

Overall, MANAS contributes to ongoing research in AI-assisted mental health by demonstrating how large language models, when properly guided and supervised, can function as supportive companions while also providing clinically meaningful insights. The platform lays a strong foundation for the development of digital mental-health systems that are safe, scalable, empathetic, and capable of complementing human therapists rather than replacing them.

REFERENCES

1. [1] K. Kasaudhan, "AI-Driven Tools for Detecting and Monitoring Mental Health Conditions Through Behaviour Patterns," International Journal of Preventive Medicine and Health (IJPMH), vol. 5, no. 3, pp. 14-19, 2025.
2. [2] A. C. De Oliveira, J. P. C. Azevedo, L. Ruback, R. Moreira,
3. S. S. Teixeira, and A. S. Teles, "Effect of Explainable Artificial Intelligence on Trust of Mental Health Professionals in an AI-Based System for Suicide Prevention," IEEE Access, 2025.
4. [3] Z. Alghamdi, T. Kumarage, G. Agrawal, M. Karami, I. Almuteb, and H. Liu, "RedditESS: A Mental Health Social Support Interaction Dataset Understanding Effective Social Support to Refine AI-Driven Support Tools," arXiv preprint arXiv:2503.21888, 2025.
5. [4] K. Jung, G. Lee, Y. Huang, and Y. Chen, "'I've talked to ChatGPT about my issues last night.': Examining Mental Health Conversations with Large Language Models through Reddit Analysis," Proceedings of the ACM on Human-Computer Interaction, vol. 9, no. 7, pp. 1-25, 2025.
6. [5] E. Aghakhani, L. Wang, K. T. Washington, G. Demiris, J. Huh-Yoo, and R. Rezapour, "From Conversation to Automation: Leveraging LLMs for Problem-Solving Therapy Analysis," arXiv preprint arXiv:2501.06101, 2025.
7. [6] C. Yin, F. Li, S. Zhang, Z. Wang, J. Shao, P. Li, and X. Jiang, "MDD-5k: A New Diagnostic Conversation Dataset for Mental Disorders Synthesized via Neuro-Symbolic LLM Agents," in Proc. AAAI Conf. Artificial Intelligence, vol. 39, no. 24, pp. 25715-25723, Apr. 2025.
8. [7] Z. Iftikhar, A. Xiao, S. Ransom, J. Huang, and H. Suresh, "How LLM Counselors Violate Ethical Standards in Mental Health Practice: A Practitioner-Informed Framework," in Proc. AAAI/ACM Conf. AI, Ethics, and Society, vol. 8, no. 2,
9. pp. 1311-1323, Oct. 2025.
10. [8] Y. Wang, Y. Wang, Y. Xiao, L. Escamilla, B. Augustine, K. Crace, and Y. Zhang, "Evaluating an LLM-Powered Chatbot for Cognitive Restructuring: Insights from Mental Health Professionals," arXiv preprint arXiv:2501.15599, 2025.
11. [9] S. Tu, A. Powers, S. Doogan, and J. D. Choi, "TRUST: An LLM-Based Dialogue System for Trauma Understanding and Structured Assessments," arXiv preprint arXiv:2504.21851, 2025.
12. [10] S. Wang, Y. Cheng, A. Song, S. Keedy, M. Berman, and
13. N. Feamster, "Can LLMs Address Mental Health Questions? A Comparison with Human Therapists," arXiv preprint arXiv:2509.12102, 2025.