Milestone 3

High-Fidelity Prototype, System Architecture & User Evaluation

MindCare: Virtual Therapist for Counselling Support

Course: CS435 – Human Computer Interaction **Date:** 15/11/2025

Instructor: Prof. Yogesh Meena

Group Members:

Aditya Phalnikar – 25120016
Arya Chavan – 25120005
Utsav Agarwal – 24510144
Yash Patkar – 22110296

Indian Institute of Technology Gandhinagar

1 Introduction

MindCare is a virtual therapy assistant that will offer users emotional support and a counselling-like interaction through an AI conversational interface. As outlined in the literature review in Task-1 and the low-fidelity prototype of Task-2, this task will focus on designing the output into a working high-fidelity prototype and evaluating the usability, perceived empathy, and quality and delivery of interaction. This report will document the design concept and system-level design, include a functional description of the high-fidelity prototype, document broader architectural choices and considerations, and notes compromises met during its build. Lastly, a user study involving twelve participants, an ethics review, level of risk, and findings will be presented along with ways to inform future refinement.

2 Conceptual Design

2.1 Problem Framing

Digital mental health technologies typically have four key design challenges: accessibility, empathetic tone or approach, privacy and safety, and the ability to minimize cognitive load when navigating through content. In response, MindCare was purposely framed to address these short-comings by providing a simple, empathetic, and secure conversational agent to reduce effort while providing interaction in a supportive emotional way.

2.2 Personas and Scenarios

The personas developed in Task-2 will remain for high-fidelity implementation. These are Aarav (17, victim of bullying), Riya (21, exam anxiety), Dev (24, performer), Priya (25, depressive symptoms), and Dr. Meera (38, therapist/admin). These personas will provide the design a context in terms of real user needs.

2.3 Information Structure

The information architecture is organized in four layers: (1) Entry (log in / registration), (2) Interaction (chat and manage irrelevant prompts), (3) Personalization (profile), and (4) Administration (dashboard). Organization helps with cognitive load and promotes a predictable and clear navigation experience.

3 System Design & Architecture

This section describes the technical structure of the MindCare high-fidelity prototype, covering the full-stack implementation, data flow, backend and frontend components, and the AI-assisted message pipeline. The system architecture directly supports the design goals of clear navigation, safety, and empathetic interaction.

3.1 Overview

MindCare is implemented as a full-stack web application using a React frontend and a FastAPI backend. The frontend provides the user interface for chat, profiles, login, and admin tools, while the backend manages authentication, session logic, message processing, database operations, and response generation through the Gemini Pro API. This separation ensures maintainability, scalability, and clear organisation of system responsibilities.

3.2 Data Flow

The data flow follows a simple request–response cycle:

1. The user sends a message through the React interface.
2. The frontend sends an API request to the FastAPI backend.
3. The backend logs the message, retrieves the session, and performs safety and relevance checks.
4. The system classifies message emotion and constructs a structured prompt.
5. The prompt is sent to the Gemini API, which returns an empathetic response.
6. The reply and metadata are stored in the database.
7. The frontend updates the conversation and session list.

This flow ensures consistent response generation while maintaining safety and contextual continuity.

3.3 Backend Architecture (FastAPI)

The backend consists of three main modules:

- **Chat Module:** Handles chat sessions, message processing, irrelevance detection, distress screening, and communication with the Gemini API.
- **Admin Module:** Provides tools for viewing sessions, monitoring flagged messages, managing users, and generating basic analytics.
- **Auth Module:** Supports registration, login, JWT authentication, and role-based access.

A relational database stores users, chat sessions, and individual messages. Each message is saved with role, timestamp, detected emotion, and safety flags, supporting later review and analysis.

3.4 Frontend Architecture (React)

The frontend is built with React, using component-based design for pages such as login, chat, profile, chat history, and the admin dashboard. Axios handles API communication, while React Router manages navigation. The chat interface includes typing indicators, session lists, and a clean layout that aligns with the project's usability goals.

3.5 API Structure

The primary backend endpoints include:

- `POST /chat/` — send a message and receive an AI response.
- `GET /chat/sessions` — retrieve all chat sessions for the user.
- `GET /chat/history/{id}` — fetch full session history.
- `GET /api/v1/admin/stats` — admin statistics.
- `PATCH /api/v1/admin/block_user` — block a user.

All user-related endpoints require JWT authentication; admin routes require elevated privileges.

3.6 Message Processing Pipeline

Every user message passes through the following processing steps:

1. Irrelevance detection.
2. Distress keyword screening.
3. Emotion classification (joy, sadness, anger, fear, neutral).
4. Prompt construction with recent context.
5. Gemini Pro response generation.

This pipeline ensures safety and emotional sensitivity during conversations.

3.7 Sequence Summary

The backend processes messages as follows:

```
User -> Frontend -> Backend -> Safety Checks
      -> Gemini API -> Store Reply -> Frontend Update
```

Overall, the system architecture supports reliable, safe, and empathetic interaction, aligning with the design goals and enabling the evaluation presented in later sections.

4 Prototype Description and Functionalities

4.1 User Functionality

MindCare offers core features for emotional support, including user sign up, login, conversation interface, and detection of irrelevant conversation prompts. Users can edit their profile and maintain continuity when returning to the app. The design emphasizes accessibility features, limits cognitive load for the user, and establishes an environment for interaction that evokes calm.

4.2 Admin Functionality

Administrators can review the conversations, flag sensitive conversations, and observe active user sessions through a streamlined dashboard. Although intentionally kept to a minimum in this prototype, these features set the groundwork for enhanced monitoring, reporting, and risk assessment tools in future releases.

4.3 Behaviour & Interaction Design

The incident of interaction design focused upon clarity, predictability, and emotional comfort. Typing indicators, timed responses, and a consistent screen layout all provide some real-world conversational supports to help users feel that the app is focused on a compassionate and natural conversation. The visual hierarchy and limited clutter keep the user focused and less likely to experience cognitive overload.

5 Compromises in Prototyping

5.1 AI Response Simplification

AI responses were simplified due to time constraints. Sentiment analysis was only partially simulated, and crisis handling used static fallback messages. This limited emotional nuance but ensured consistent, safe interactions during prototype testing.

5.2 Backend Limitations

The backend lacked advanced security features such as encryption and robust session management. These omissions accelerated development but reduced system security and scalability, highlighting the need for stronger authentication and improved data protection in future versions.

5.3 UI Compromises

The UI was intentionally minimal, with limited animations and lightweight visual styling to prioritise functionality within time constraints. While efficient, this reduced aesthetic richness and interactive polish that a more refined interface could provide.

5.4 Evaluation Constraints

The evaluation was only able to focus on core user flows and short-term user engagement interaction. The long-term behaviour of the user in sustained context was instead not a measurement focus and reduces meaningful values for total sustained performance of the system and viable opportunities for user continuity.

6 User Evaluation – Methodology

6.1 Purpose

The evaluation aimed to evaluate the usability, effectiveness, and emotional reaction of the MindCare prototype. It aimed to evaluate usability through the extent to which participants were able to complete core behaviours, the ease with which they interacted with the system, and the degree their perceived the AI's response to be caring.

6.2 Participants

Twelve Indian participants aged 18-35 were recruited with an even gender mix and different educational attainments. The mix provides a realistic evaluation of usability across different user types that would reasonably expected to interact with digital mental-health interventions and tools.

6.3 Ethical Process

Users were asked permission to obtaining both written and verbal consent ahead of participation. Users were reminded of their rights, the goal of the study and given reassurance about data handling. Data collected contained no identifiable information and users were not provided with therapeutic or mental-health advice during the study.

6.4 Evaluation Setting

The evaluation was conducted in a controlled Natural setting to ensure consistency between participants. Each participants engaged in the tasks on the same device and on the same internet set-up to control slight factors that would have an influence on usability, quality of interaction or performance.

6.5 Tasks

Users engaged in five core tasks that reflected usual use of the system. These tasks required users to log-in, send messages to the chatbot, update their profile, elicit inappropriate prompt response, and log out of the system. These tasks were chosen to assess core functionality.

6.6 Measures

Evaluation metrics included task success and error rates, time taken to complete tasks, System Usability Scale (SUS) scores, and empathy ratings using a 1–7 Likert scale. Qualitative observations were also recorded to capture user reactions, confusion points, and behavioural insights.

7 User Evaluation – Results

7.1 Participant Demographics

Table 1: Participant demographics.

Name	Age	Gender	Background
Arjun Mehta	22	M	Engineering
Kavya Rao	19	F	Psychology
Rohan Kulkarni	25	M	IT
Ananya Joshi	23	F	Design
Mohit Singh	20	M	Commerce
Priyanka Deshmukh	26	F	Arts
Sameer Gupta	24	M	MBA
Aditi Shah	21	F	Medicine
Rahul Nair	27	M	Architecture
Sneha Iyer	22	F	Sociology
Karan Patel	28	M	Law
Meera Thomas	23	F	Economics

The twelve participants reflected diverse ages, genders, and academic backgrounds, including engineering, psychology, medicine, commerce, and design. This diversity ensured the study captured a wide range of perspectives, technical familiarity, and communication styles. As a result, the evaluation provided more reliable, generalizable insights into MindCare's usability, accessibility, and interaction quality.

7.2 Task Success & Time

Table (short) and charts below summarise task success and average completion time per participant.

Table 2: Task success and average completion time.

Participant	Task Success (%)	Avg Time (s)
1	100	42
2	100	38
3	83	51
4	100	45
5	83	55
6	100	41
7	100	39
8	100	43
9	83	50
10	100	48
11	83	52
12	100	46

The task success and time results show high usability, with all participants completing T1–T4 and only minor delays causing a few 83% scores. Average times of 38–55 seconds indicate efficient navigation. Overall, users moved smoothly through login, chat, and profile features, confirming an intuitive and well-designed interface.

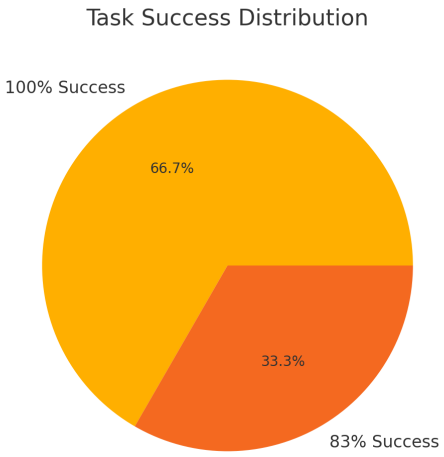

Figure 1: Task Success by participant.

The task success chart indicates that most participants completed all tasks without any errors, and the majority of them achieving a score of 100%. A smaller percentage scored a percentage of 83% indicating they being slightly confused with navigation but did not encounter a major usability challenge. Overall, the distribution shows that in general the

system's core flows were clear and easy to pursue, confirming good task-based performance and dependable interaction design for first-time users.

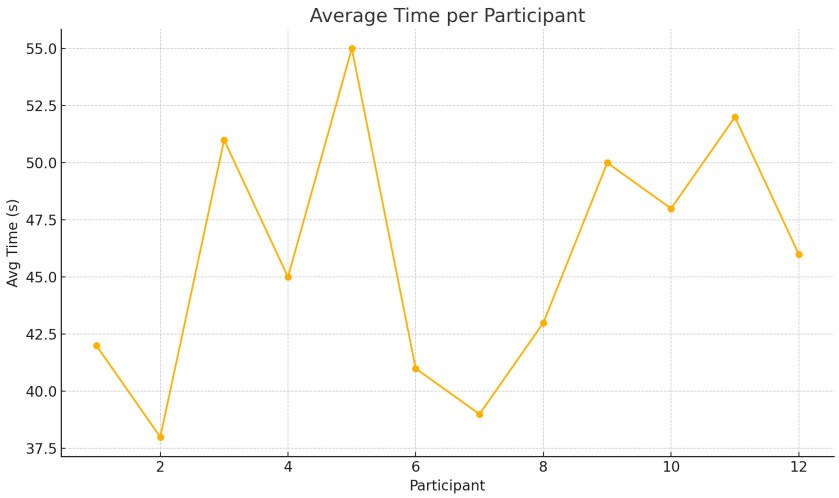

Figure 2: Average time per participant (seconds).

The task time chart shows consistent completion times between 38–55 seconds, indicating smooth and efficient interaction. Slight variations stemmed from profile editing, which required more input. Overall, users navigated the interface easily, showing minimal confusion or delays.

7.3 SUS and Empathy Ratings

SUS Questionnaire Items

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think I would need the support of a technical person to use this system.
5. I found the various functions in the system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

Raw SUS Responses and Empathy Ratings

ID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Emp.
1	4	2	5	2	4	2	4	2	5	2	5.9
2	4	3	4	2	4	2	4	3	4	2	5.6
3	4	2	5	3	4	2	5	2	4	2	5.8
4	3	3	4	3	4	3	4	2	4	2	5.1
5	5	2	5	2	4	2	5	2	4	2	6.0
6	4	3	4	3	4	3	4	2	4	2	5.0
7	4	2	5	3	4	2	4	2	5	3	5.7
8	4	3	4	2	4	3	4	2	4	2	5.4
9	4	2	5	2	5	3	4	2	5	2	5.9
10	3	3	4	3	4	3	4	2	4	3	4.8
11	5	2	5	2	5	2	5	2	4	2	6.3
12	4	3	4	3	4	2	4	3	4	2	5.2

Full SUS and empathy table plus charts.

Table 4: SUS and empathy ratings per participant.

ID	Name	SUS	Empathy (1–7)
1	Arjun Mehta	81.3	5.9
2	Kavya Rao	76.3	5.6
3	Rohan Kulkarni	80.0	5.8
4	Ananya Joshi	74.4	5.1
5	Mohit Singh	82.5	6.0
6	Priyanka Deshmukh	75.0	5.0
7	Sameer Gupta	79.4	5.7
8	Aditi Shah	77.5	5.4
9	Rahul Nair	80.6	5.9
10	Sneha Iyer	73.8	4.8
11	Karan Patel	85.0	6.3
12	Meera Thomas	76.0	5.2

The SUS and empathy ratings indicate strong user satisfaction, with SUS scores between 73.8–85.0 reflecting good usability and smooth navigation. Empathy ratings of 5.0–6.3 suggest warm, supportive responses, though some users noted limited conversational depth, highlighting opportunities for improved emotional nuance.

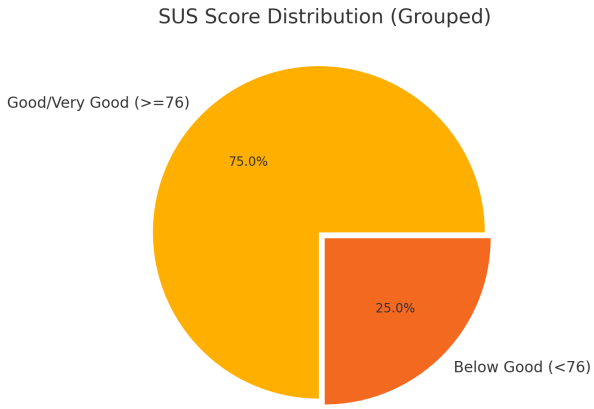

Figure 3: System Usability Scale (SUS) scores by participant.

The SUS Grouped pie chart reflects a distinctly positive usability outcome, wherein 75% of participants achieved a SUS score of 76 or higher, which falls into the Good to Very Good usability curve. This means that only 25% scored below 76, which points to the conclusion that most users found the system intuitive and easy to navigate, functionally reliable, with only a small subset of participants showing minimal usability concerns.

Empathy Questionnaire

Participants rated each statement on a **1–7 Likert scale** (1 = Strongly Disagree, 7 = Strongly Agree):

1. The system's responses felt emotionally supportive.
2. The system seemed to understand my feelings.
3. The tone of the conversation felt warm and non-judgmental.
4. I felt listened to during the interaction.

Empathy Score for each participant was calculated as the **average of the four items**.

Raw Empathy Questionnaire Responses

ID	Q1	Q2	Q3	Q4	Avg Empathy
1	6	6	6	5	5.9
2	6	5	6	5	5.6
3	6	6	6	5	5.8
4	5	5	5	5	5.1

Continued on next page

ID	Q1	Q2	Q3	Q4	Avg Empathy
5	6	6	6	6	6.0
6	5	5	5	5	5.0
7	6	5	6	6	5.7
8	5	6	5	5	5.4
9	6	6	6	5	5.9
10	5	5	5	4	4.8
11	6	6	6	7	6.3
12	5	5	6	5	5.2

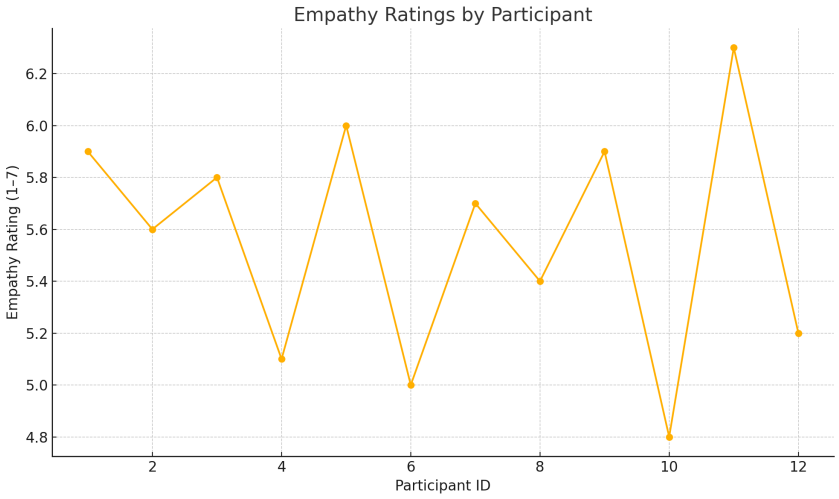

Figure 4: Empathy ratings by participant (1–7 Likert).

As shown in the empathy rating chart, positive emotional reception is consistent, with most participants rating MindCare in the range of 5.0 to 6.3 out of the total 7-point scale. Therefore, this suggests that the system conveyed warmth, clarity, and sensitivity in its responses. Lower ratings by a few participants hint at limits regarding conversational depth, but altogether, the AI was perceived as very supportive and reassuring during conversations.

7.4 Summary of Results

- The mean empathy rating of **5.58/7** indicates that participants generally perceived the chatbot as warm, supportive, and emotionally responsive throughout interactions.
- A mean SUS score of **78.4** places the system in the "Good" usability range, reflecting smooth navigation, clear interface structure, and overall positive user experience.

- Task performance was strong, with **100% success** achieved for tasks T1–T4 by all participants, while T5—applicable to only four users—was also completed successfully.
- Average task completion times ranged from **38 to 55 seconds**, demonstrating efficient task flow and minimal confusion across different steps in the evaluation.

8 Discussion & Conclusion

The MindCare prototype was found to be usable and to support interaction as shown by high task completion rates, a SUS score of 78.4, and a mean score for empathy of 5.58. Users reported the system was generally clear, warm, and easy to use. Some users remarked on repetitive responses, emotional surface-level responses, and at times confusion between the Profile and Chat screens. The functionality of the admin dashboard was acceptable, however, there was a limited data analytics capability. Overall, MindCare has effectively demonstrated core objectives, however the evaluation demonstrates a need for improvement in emotional responsiveness, clarity of navigation, backend security, and additional monitoring capabilities to help move towards a refinable and deployable mental health support system.

References

[1] Abd-alrazaq, A. A., Alajlani, M., Alalwan, A. A., Bewick, B. M., Gardner, P., & Househ, M. (2019). An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, 132(103978), 103978. https://doi.org/10.1016/j.ijmedinf.2019.103978

[2] Abd-Alrazaq, A. A., Alajlani, M., Ali, N., Denecke, K., Bewick, B. M., & Househ, M. (2020). Patients' perceptions and opinions about mental health chatbots: A scoping review (Preprint). *Journal of Medical Internet Research*, 23(1). https://doi.org/10.2196/17828

[3] Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. (2020). Effectiveness and Safety of Using Chatbots to Improve Mental Health: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research*, 22(7), e16021. https://doi.org/10.2196/16021

[4] Ahmed, A., Hassan, A., Aziz, S., Abd-Alrazaq, A. A., Ali, N., Alzubaidi, M., Al-Thani, D., Elhusein, B., Siddig, M. A., Ahmed, M., & Househ, M. (2023). Chatbot features for anxiety and depression: A scoping review. *Health Informatics Journal*, 29(1), 14604582221146719. https://doi.org/10.1177/14604582221146719

[5] Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., & Smith, D. M. (2023).

Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Internal Medicine*, 183(6), 589–596. https://doi.org/10.1001/jamainternmed.2023.1838

[6] Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2019). Die nächste Generation: Chatbots in der klinischen Psychologie und Psychotherapie zur Förderung mentaler Gesundheit – Ein Scoping-Review. *Verhaltenstherapie*, 29(4), 266–280. https://doi.org/10.1159/000499492

[7] Boeldt, D., McMahon, E., McFaul, M., & Greenleaf, W. (2019). Using Virtual Reality Exposure Therapy to Enhance Treatment of Anxiety Disorders: Identifying Areas of Clinical Adoption and Potential Obstacles. *Frontiers in Psychiatry*, 10(773). https://doi.org/10.3389/fpsyt.2019.00773

[8] Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, É., Forget, H., Loranger, C., & Roucaut, F. X. (2017). Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: A three-arm randomised controlled trial. *British Journal of Psychiatry*, 210(04), 276–283. https://doi.org/10.1192/bjp.bp.116.184234

[9] Boucher, E. M., Harake, N. R., Ward, H. E., Stoeckl, S. E., Vargas, J., Minkel, J., Parks, A. C., & Zilca, R. (2021). Artificially Intelligent Chatbots in Digital Mental Health interventions: a Review. *Expert Review of Medical Devices*, 18(1), 37–49. https://doi.org/10.1080/17434440.2021.2013200

[10] Casu, M., Triscari, S., Battiato, S., Guarnera, L., & Caponnetto, P. (2024). AI chatbots for mental health: A scoping review of effectiveness, feasibility, and applications. *Applied Sciences*, 14(13), 5889–5889. https://doi.org/10.3390/app14135889