# ACTION PLAN

# CAREMATE : AN AGENTIC AI – POWERED HOSPITAL ASSISTANT FOR PATIENT- CENTERED CARE

## 1. ACTIVITY COMPONENTS

CareMate consists of the following major activities, each of which has a well-defined role in the overall architecture of the system:

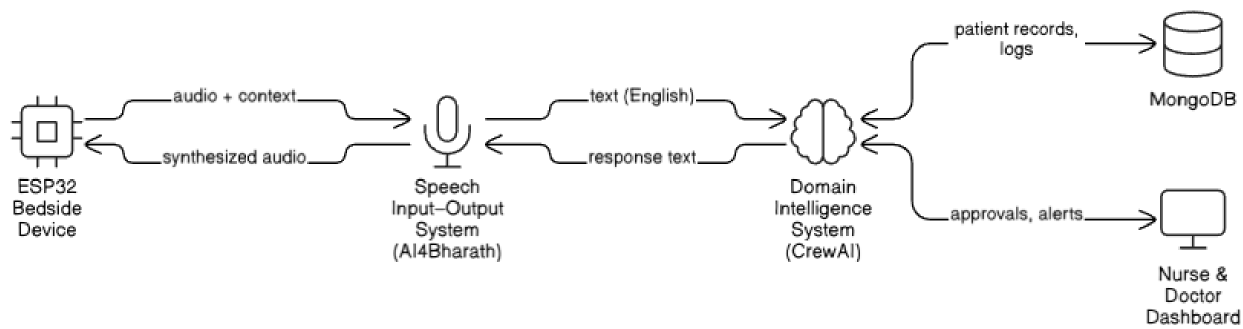


*Figure 1: High-level system architecture of CareMate and its core components.*

### 1.1 Bedside Interaction Device

- It is implemented using the ESP32 microcontroller accompanied by a microphone and a speaker.
- Allows the patients to communicate with the system through natural voices from the hospital bed.
- Records audio of patients and adds context such as hospital number and room number.
- Send pre-recorded audio to the hospital boundary system and a synthesized voice response playback.
- It is intended to be a stateless interface that carries out no speech processing, data access, or decision-making operations.

### 1.2 Speech Input-Output Processing System

- Employing AI4Bharath speech technologies for:
  - Automatic Speech Recognition (ASR).
  - Translation involving multiple languages (Indic languages and English).
  - Synthesis of Text-to-Speech.

- Translates patient audio to normalized text and then changes normalized text to speech.
- Acts as a speech transformation layer which has the ability to comprehend and convert Indic languages into English as well as English into Indic languages. It can also perform TTS & STT.

### 1.3 Domain Intelligence and Agentic AI System

- It is the main decision-making module in CareMate.
- The answers given to any of the questions posed to the patients, as per their report.
- Implemented with CrewAI multi-agent orchestration and an open-source domain language model.
- Pulls patient files depending on context parameters.
- Interprets intent of patient, urgency of intent, and achieves said intent in prompt
- Involves humans wherever necessary, as in ordering medicine, and so on.

### 1.4 Data Management and System Integration

- Patient data, interaction history, and audit trails are maintained through MongoDB (Community Edition).
- Communication between components is done with the aid of RESTful APIs created with FastAPI.

### 1.5 Clinical Staff Interface-RSAT

- An online nurse and doctor interface developed using Lovable.
- Offers the healthcare professionals real-time visibility of patient requests and system operations.
- Supports approval or rejection of medical requests and shows emergency notifications.
- It ensures human oversight, accountability, and transparency in processes with the aid of artificial intelligence.

## 2. PROBLEM SELECTION

- A continuous intelligent patient interaction system is absent in hospitals, which can help patients in between visits by the doctor and the nurse.
- Patients often come with questions that remain unanswered regarding:
  - Medication Regimens.
  - Diet & Non-Medical Needs.
  - Doctor visit times.
  - Emergency or distress circumstances.
- The existing nurse call systems only incorporate basic alerts and lack comprehension of patients' intent, urgency, and context.

- Repetitive and lower priority requests clog the workload of nurses and doctors, thus delaying responses to critical situations.
- There is no integrated system which provides voice interaction, multi-language support, and AI orchestration capabilities under the vicinity of patient queries in the room by supervisors.
- This project will cover the issue of inefficient patient-staff communications in hospital environments by implementing a voice-first agentic AI-powered patient assistance system.

## 3. PROBLEM UNDERSTANDING

- Patients admitted to hospital wards may go for a remarkably long period without coming into contact with either a doctor or a nurse.
- There are patient-related demands which are routine or even non-clinical in nature, but these require manual intervention by nurses.
- Patients may not be able to convey their urgent or distressed messages appropriately through the existing calling system.
- Language barriers are another obstacle to communication, especially for older patients and patients who speak different languages.

## 4. INITIAL ASSESSMENT

- The project is intended to be technically feasible using low-cost hardware and open-source software platforms.
- The choice of a voice-first design is based on ensuring that the system is accessible to patients regardless of their ages and levels of literacy.
- The required resources are:
    - An ESP32-based bedside device with microphone and speaker.
    - A local system running the speech processing and backend services.
    - Open-source frameworks for AI orchestration and reasoning.
- The duration of this project is estimated to be five months and requires a daily average effort of 3 to 4 hours along with extra efforts on holidays and weekends.
- The initial risks of accuracy of spoken words, latency, and clear input audio quality are recognized and mitigated by means of module design, artificial intelligence behavior, and human-in-the-loop strategies.

## 5. PROBLEM DEFINITION

- Patient assistance support needs to be done in a continuous, dependable, and intelligent manner in hospital wards in addition to visits from doctors and nurses.
- Taking care of all the patient requests manually increases the workload and delays the response process in critical situations.
- There isn't a channel for patients to access assistance in a natural fashion and in multiple languages.
- The solution must address patient safety, compliance issues, and human intervention while preventing the AI system from making any medical decisions.
- The challenge thus posed is the design of a voice-first and policy-driven agentic AI that aids efficient patient-staff communication and directs requests intelligently.

## 6. PROBLEM RECOGNITION (SIGNIFICANCE AND RELEVANCE)

- Hospitals function in an environment where effective communication and fast response are of essence in terms of patients' safety and satisfaction.
- There is pressure on nurses and doctors with too many patients, especially because of non-clinical requirements by patients.
- The consequences of postponing the resolution of concerns expressed by the patient may result in:
    - Rising patient anxiety.
    - Loss of trust in health care delivery.
    - Safety risks involved during emergencies.
- Current communication infrastructure in hospitals is lacking in the following ways
    - Context awareness of patient medical records.
    - Capable of prioritizing requests based on urgency.
    - Multilingual and Voice-based Accessibility.
- The most affected sections of patients would include senior citizens, post-operative patients, and patients who lack mobility and literacy skills.
- Increasing demands on technology-enabled patient care solutions that can improve the delivery of healthcare while preserving a strong safety margin and human control.
- It is particularly relevant in the modern healthcare context, as it affects patient experience, efficiency of healthcare providers, and ultimately healthcare quality.

## 7. OUTCOME IDENTIFICATION

- The direct result of this project will be the design and implementation of a System of a voice-first, agentic AI-driven patient support system.
- The system will show complete functionality from patient voice input to the correct response and notification of health personnel.
- The result comprises the incorporation of:
    - A bedside voice interaction device.
    - A speech processing system.
    - Centralized domain intelligence system.
    - Clinical Dashboard for Staff.
- The project will mimic a realistic deployment environment for a hospital, where patient inquiries will be demonstrated to be handled and escalated in a safe and controlled fashion.
- It will be possible to clearly differentiate between:
    - Non-medical requests.
    - Medical requests that are required to be approved by the staff.
    - Life-threatening emergencies demanding immediate escalation.
- The system will be able to respond to all the questions asked by the patients as per their report and also completes tasks.

## 8. OUTCOME CHARACTERISTICS

The expected result of the CareMate project will have the following main features:

- **Voice-First Interaction:**
  They can communicate in a natural way through speech without having to use physical interfaces. Communication does not require them to be literate.

- **Multilingual Accessibility:**
  This system supports various Indian languages and thus provides equal interaction capabilities for different language-speaking patients.

- **Context-Aware Responses:**
  Patient questions are answered using available patient information and hospital information to ensure that the relevant and meaningful answer is derived for the question asked.

- **Human-in-the-Loop**
  The nursing and medical staff are still in charge of medical aspects through approvals and alerts.

- **Emergency Sensitivity:**
  The system is capable of recognizing situations of distress/emergency and initiating an urgent escalation.

- **Auditability and Transparency:**
  Every engagement, decision, and escalation is recorded for tracking and analysis.

- **Demonstrable Performance:**
  The design aims to provide acceptable response latency, reliable speech recognition, and system consistency in the simulated environment within the hospital setting.

## 9. REQUIREMENT ANALYSIS

The following is an outline of theoretical bases, data needs, and expertise required for successful completion of the CareMate project.

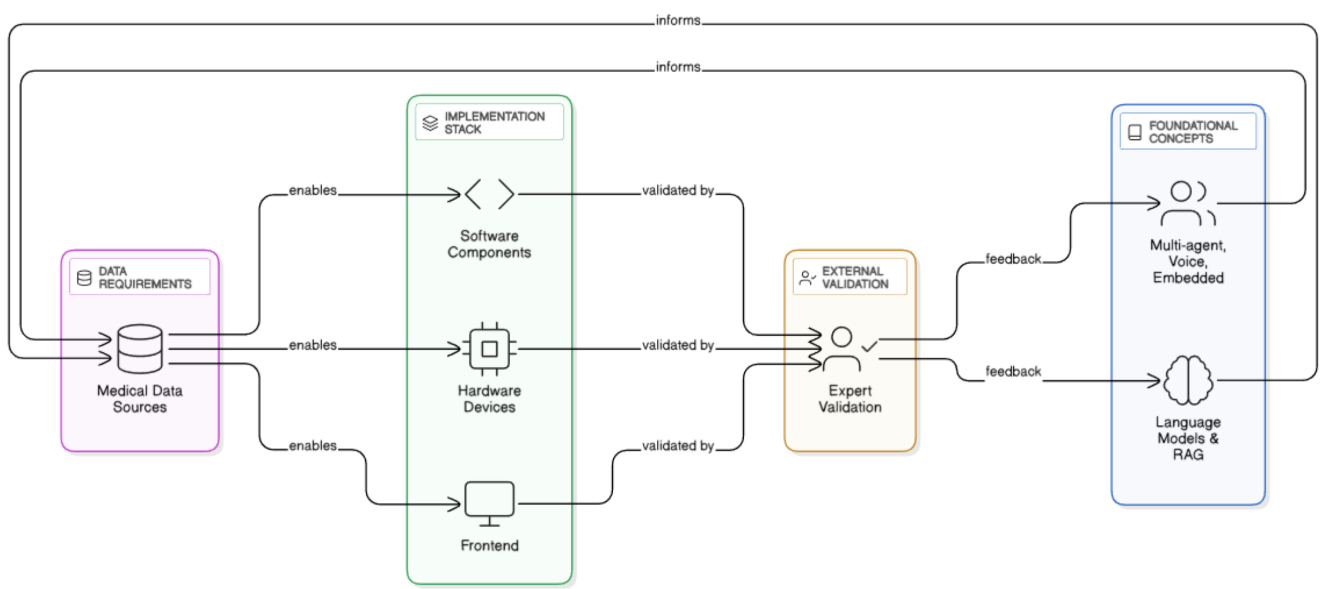


*Figure 2: Overview of the foundational concepts, data requirements, implementation stack, and validation inputs of the CareMate system.*

### 9.1 Existing Theory and Devices

The following established theories, architectures, and systems will serve as the basis for this project:

- **Retrieval-Augmented Generation (RAG):**
  Used to make sure answers are embedded within recalled patient-specific medical records rather than simply depending on generative knowledge.

- **Multi-Agent AI Architectures:**
  The CrewAI is used to implement task coordination, assessment of intent, policy compliance, and responses in specific agents.

- **Voice-Based Human-Computer Interaction:**
  This is the main interaction tool, allowing for natural communication among patients within the wards.

- **Embedded Systems For Hardware Devices:**
  Used in the design of a low-cost bedside communication interface based on the ESP32 microcontroller.

- **Large Language Models (LLMs):**
  Used in the domain intelligence system for reasoning, orchestration, and natural language response generation with policy constraints.

### 9.2 Data Collection Plan

The system uses only publicly available and synthetic data and is thus guaranteed to be privacy-compliant and of sound ethical standards.

- **Medical knowledge sources:**
  - Peer-reviewed journals and abstracts obtained from PubMed.
  - Obtained sample patient medical records from Kaggle health datasets.

- **Patient Query Data:**
  - Common inpatient questions derived from datasets such as: ChatDoctor and HealthCareMagic
  - Synthetical queries designed for simulating normal, medical, and emergency situations.

- **Emergency signals:**
  - List of emergency-related keywords and distress phrases that are curated and used in the detection of urgency.

- **Dataset size:**
  - About 1,000-2,000 samples of question-and-answer pairs, consisting of both curated and synthetic data.

### 9.3 Tools and Equipment

- **Artificial Intelligence & Language Processing**
  - AI4Bharath: Regional language speech recognition, translation, and speech synthesis.
  - Fine-tuning open LLMs using Hugging Face - Domain reasoning and response generation.

- **Agentic Framework**
  - CrewAI - Multi-agent Orchestration & Workflow Management.

- **Backend Technology**
  - Python - Core development language.
  - FastAPI – building REST APIs.
  - MongoDB (Community Edition) – Patient records, Agent memory, and Audit logs.

- **Hardware Elements**
  - Microcontrolador ESP32
  - Microphone module
  - Speaker module
  - Push button for manual operation/access or emergency use

- **Frontend**
  - React – Nurse and Doctor dashboard for web.

### 9.4 External Expertise

- Informal consultation with nurses and doctors to validate clinical workflows and patient interaction scenarios.
- Selection of experts based on their availability, field of expertise, and experience in the hospital setting.
- The feedback is factored in mostly at stages of design validation and workflow analysis

.

## 10. TEAM KNOWLEDGE AND LOGISTICS

- The team has the necessary fundamental knowledge in artificial intelligence and backend systems and design of agent-based systems essential for the implementation of a system's core modules.
- Experience in Python-based development, RESTful APIs, and elementary embedded systems is prior knowledge to integrate software services and hardware resources well.

- They also have an understanding of orchestrating and agent AI terminology on platforms such as Crew AI, which enables decision-making and task flow management.
- Development, integration, and test activities will remain primarily based on personal computer resources, with a laptop that will simulate the hospital edge server environment.
- However, for domain language modeling tasks and fine-tuning of models, local computing power is insufficient and hence the need to avail selective GPU power from the cloud.
- Internet connection and access to open-source tools and cloud services are adequate to perform speech processing, data preparation, training of models, and integration.
- Logistical problems which have been identified include
    - Effective integration between hardware and software components
    - Handling speech accuracy in real environments
    - Maintaining Homogeneous System Performance Across Various Situations
- The team will be able to allocate 3-4 hours on any given day on the project, including longer working hours on weekends and holidays so that they are aligned with the development timeline discussed.

## 11. CAPABILITY GAP ANALYSIS

- Although it has in-depth knowledge in AI systems and backend developers, some advanced areas of system design for healthcare applications still require improvement.
- The capability gaps that have been identified are
    - Formal modeling of hospital workflows and escalation hierarchies
    - Designing AI-safe behavior for application in healthcare settings
    - Guaranteed hardware-software integration for continuous voice communication
    - Evaluation of system functionality within a simulated operational environment, such as those of hospitals.
- These gaps are filled by:
    - Targeted self-study based on technical and research-based information
    - Incremental prototyping and testing: This approach involves
    - Validation of workflows in collaboration with medical professionals
- This is because the project plan has ensured adequate provision for learning, experimentation, and refinement in the early stages of development.

## 12. CONSULTING METHODS

- The project uses the continuous consultation method in order to deal with design-related problems and assumptions.
- Conversations with nurses and doctors are used to understand real-world hospital workflows, patient interaction patterns, and patient escalations.
- There is feedback provided by academic and technical mentors at critical design and integration activities.
- Research papers, technical documents, and healthcare guidelines relevant to the problem are reviewed to overcome conceptual and implementation issues.
- Findings and insights will be applied iteratively to the design of systems, agent behavior, and workflow orchestration.

## 13. SCOPE DEFINITION

### 13.1 Included Scope:

- Design and development of a voice-first patient assistance system for hospital ward environments.
- ESP32-based Implementation of a Bedside Device for Voice Input from Patients and Audio Output.
- Integration of AI4Bharath for Speech Recognition, Multilingual Translation, and Text-to-Speech Processing.
- Building a policy-guided agentic domain system for AI with CrewAI for Intent Understanding, Orchestration, and Response Generation.
- Simulated patient data: Medical record retrieval and summarization for context-aware interaction support.
- Handling of non-medical, medical, and emergency requests with proper escalation scenarios.
- Complete the tasks and respond to the queries of patients.
- An online nurse and doctor dashboard for monitoring, approvals, and alerts.
- Simulation of a hospital deployment on existing computer resources.

### 9.2 Excluded Scope:

- Regulatory certification or clinical validation.
- Integration with Wearable Devices or Real-Time Vital Sign Monitoring.
- Autonomous medical diagnoses or treatment suggestions.

## 14. COMPLETE ACTION PLAN WITH TIMELINES

In this section, an improved and staged action plan will be discussed according to the alignment of system development and technical dependencies.

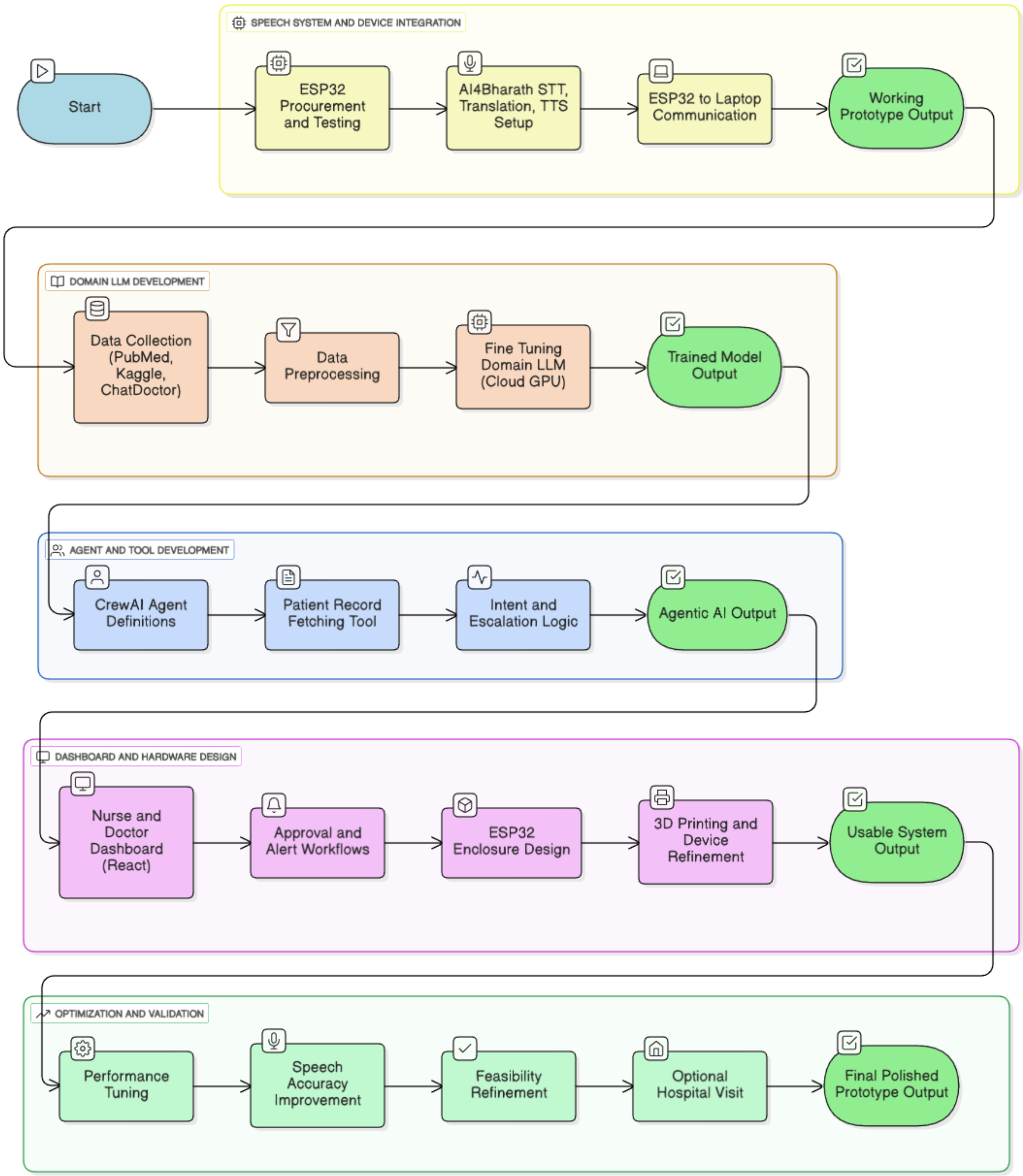


*Figure 3: Phase-wise development roadmap of the CareMate system outlining key activities and expected outcomes across the project timeline.*

**Phase 1: Speech System and Device Integration (Month 1):**
Activities:

- Finalize the design of the speech input/output pipeline.
- Acquire hardware parts:
  - ESP32 microcontroller
  - microphone module
  - Speaker module
  - Push button
- Establish communication between ESP32 and a local system (laptop simulating hospital edge server).
- Define and verify the data exchange schema between ESP32 and the server.
- Configure the AI4Bharath speech systems for:
  - Automatic Speech Recognition (Indic languages)
  - Language translation (Indic to English)
  - Text-to-Speech synthesis
- End-to-end voice interaction validation:
  - Patient speech is translated to English text.
  - English text that is received is transformed to Oral Response

Phase Output:

- An operational speech translation system with the capability to recognize Indic languages and transform them into English.
- Reliable interaction between the ESP32 and the laptop through a given communication protocol.

Time Management:

- Approximate time required: 12-14 hrs/week
- Available effort:
  - Monday to Friday: 1-2 hours/day
  - Weekends: Longer sessions
- Feasibility: Possible within focused daily effort

**Phase 2: Development of the Domain Language Model (Month 2):**
Activities:

- Dataset collection and curation:
  - MEDLINE search (PubMed)
  - Patient data samples (Kaggle)
  - Patient query datasets: These include datasets such as ChatDoctor
  - Emergency keywords and phrases

- Preprocessing and organizing data for training.
- Fine-tune a domain-specific open-source language model using Hugging Face tools.
- Testing the model for answering queries from patients.
- Responses must stay within the permissible boundaries of information provision.

Phase Output:

- An Articulate domain-specific language model for model response to patient queries in a generic and safe fashion.

Time Management:

- Approximately effort time: 14-16 hours/week
- Available effort:
    - Regular daily workouts and weekend training sessions
- Feasibility: Requires concentrated effort but is feasible within the timeframe.

**Phase 3: Agent and Tool Implementation (Month 3)**

Activities:

- Develop agentic architecture with the use of CrewAI.
- Define and develop agents for:
    - Patient context and memory management
    - Intent and emergency assessment
    - Authorization and Escalation Handling
    - Central control and response implementation
- Create tools for agents to use for support, such as:
    - Patient file retrieval
    - Context Summarization
    - Interaction logging
- The domain LLM is to be integrated within the agent system.
- Simulate overall request processing.

Phase Output:

- An agentic artificial intelligence system capable of reasoning, routing, and responding to patient queries.

Time Management:

- Estimated time required: 12-15 hours/week
- Available effort: Consistent with planned schedule
- Feasibility: Equitable development and testing workload

**Phase 4: Clinical Dashboard and Hardware Design (Month 4)**
Activities:

- Implement Nurse and Doctor dashboards using React.
- Enable:
  - Request Visualization
  - Approval and escalation flows
  - Emergency alert messages
- Integrate the dashboard with the domain AI backend.
- Design the physical enclosure for the ESP32 module.
- Develop enclosure designs utilizing elementary 3D design software (for example, Fusion360 or Tinkercad).
- Perform 3D printing or prototyping for the enclosure.
- Assess its usability and installability within a simulated hospital setting.
- Hardware design needs to occur for several reasons:
  - The core system behavior is already stable.
  - Hardware optimization is driven by real patterns of use.
  - Prevents enclosure design until functionality is validated.

Phase Output :

- Functional clinical dashboard.
- Bedside device prototype: enclosure-based improvement.

Time Management :

- Estimated time required: 10-12 hours/week
- Effort available: Adequate within the planned schedule

**Phase 5: Optimization, Field-Validation, and Refining Phase**
Activities:

- Improve system latency, speech recognition accuracy, and response flow.
- Enhance hardware casing and device ergonomics.
- Whenever possible, make a hospital visit to observe conditions on the ward and adjust assumptions accordingly.
- Feedback incorporated to enhance feasibility and usability.
- Complete documentation, illustrations, and demonstration resources.
- Complete final project submission and presentation.

Phase Output

- A refined, verifiable, and feasibility-tested prototype amenable to university assessment.

Time Management

- Approximate time required: 8-10 hrs/week
- Available effort: Consistent with final polishing phase

**Overall Time Management Summary**

- Total Project Duration: 5 months
- Average Weekly Effort Required: 12-15 hours
- Actual Hours Available:
  - Weekdays: Approximately 1-
  - Weekends & holidays: Long working hours

**Conclusion:**

The project can be realistically completed within the given time frame, leaving time for learning and improvements.

## 15. SUBMISSION

- However, the complete Action Plan with system architecture design, phased development, and appropriate time management will be formally prepared as a part of the project submission.
- Each element within the Action Plan strictly follows the ‘Action Plan Guidelines for FAER Scholar Awards,’ addressing activity components, definition of problems, analysis of requirements, assessment of capabilities, and implementation timelines.
- The submission will contain:
  - The final written Action Plan document
  - System architectures and workflow descriptions
  - Phase-wise development achievements and expected results
- The submission of the project will be done on or before the 31st of December, and this is based on the deadline set for the task.
- Any kind of figures or illustrations that may support the understanding of this text have been included in this submission.