Our company distributes products through various channels like retail partnerships, vending machines, and online platforms. We're facing challenges with customer service due to limited staff and high inquiry volume across WhatsApp and email.

To optimize our customer service without increasing headcount, we're seeking a streamlined solution that can:

- **Centralize communication:** Manage all customer inquiries from WhatsApp and email in one platform.
- **Automate responses:** Provide instant answers to frequently asked questions, order status updates, and delivery information.
- **Personalize interactions:** Tailor responses based on customer history and preferences.
- **Offer self-service options:** Empower customers to find information and resolve issues independently through a knowledge base or chatbot.
- **Analyze data:** Track customer service metrics to identify areas for improvement and measure the effectiveness of our efforts.

By implementing such a solution, we aim to:

- **Enhance customer satisfaction:** Provide faster and more efficient support.
- **Reduce response times:** Address inquiries promptly and minimize wait times.
- **Improve efficiency:** Free up staff to focus on complex issues and strategic tasks.
- **Control costs:** Avoid the need to hire additional customer service representatives.

We believe that streamlining our customer service processes will not only improve customer experience but also contribute to our overall operational efficiency and growth.

AI Chat

This specification outlines the requirements and features for an interactive AI chat interface that enables users to engage in conversations with an AI through both voice and text. The interface will utilize advanced speech-to-text (STT) and text-to-speech (TTS) technologies to facilitate seamless communication.

Functional Requirements

User Interaction

- Voice Input:
  - Users can provide input via voice,
  - To be transcribed into text
  - Use a speech recognition engine.
- Text Input:
  - Users can also type their queries or commands into a text box.
- Multimodal Output:

  - System displays responses in clear, formatted text on screen
  - Provides natural-sounding synthesized speech output
  - Synchronizes text and voice for seamless interaction
  - Offers configurable output preferences (text-only, voice-only, or both)
  - Maintains consistent response timing for optimal conversation flow
  - Includes visual indicators for system processing and response status

Tech Engine
Tech Engine Core Components:
- Multimodal Input Processing
  - Text input handler for keyboard/touch interactions
  - Voice input processor with real-time STT capabilities
  - Image recognition module for visual inputs
- Speech Processing System
  - Real-time TTS engine for natural voice output
  - Voice quality optimization for clear audio delivery
  - Latency management for seamless interactions
- Contextual Management
  - Conversation state tracking and memory
  - Context-aware response generation
  - User session management
- LLM Integration
  - API connection to selected LLM provider
  - Token monitoring

Conversational Features
- Contextual Understanding:
  - The AI will maintain context throughout the conversation
  - Allowing for follow-up questions and more personalized interactions.
- Intent Recognition:
  - The system will accurately identify user intents
  - Input from both voice and text inputs to provide relevant responses.
- Scenario-Based Responses:
  - The AI will provide structured responses based on predefined scenarios (e.g., FAQs, customer service inquiries)
  - Allowing for open-ended conversation.

User Experience
- Feedback and Language Mechanism:
  - Users can provide feedback on the quality of responses
  - Be used to improve the AI's performance over time.
  - The system will support English.